A CORPUS-BASED COMPARATIVE STUDY OF PRAGMATIC MARKERS:

I MEAN AND *YOU KNOW* IN NATIVE AND NON-NATIVE CONVERSATION

Thesis submitted in accordance with the requirements the University of Liverpool for the degree of Doctor in Philosophy

Wensheng Mei

Ph.D

September 2012

A CORPUS-BASED COMPARATIVE STUDY OF PRAGMATIC MARKERS: *I MEAN* AND *YOU KNOW* IN NATIVE AND NON-NATIVE CONVERSATION

WENSHENG MEI

Abstract

This study investigates how the two pragmatic markers *I mean* and *you know* are used by Chinese EFL learners as compared to British speakers.

To describe how these two markers by used by the Chinese learners, this study first investigates how they are used in the British speakers' data. To obtain a finer picture of how they are used by the native speakers, the interpretation of their pragmatic meanings is open to all plausible explanations instead of being confined to one single theory or framework. As a result, this study sets up its own categories and comes up with much longer function lists than previous studies. In addition, a new framework is proposed.

Following the completion of a detailed description of the pragmatic functions of *I mean* and *you know*, a detailed and systemic comparison between them in terms of the specific functions they play and their positioning are carried out on the grounds that these two markers are analyzed by following the same approach in the same data set. By highlighting the similarities and differences between them and explaining why, the comparison improves our understanding how they relate to each other in conversation.

Compared to the British speakers, the Chinese learners show different patterns of *I mean* and *you know* in their L2 English. The main features of the learners' uses of *I mean* and *you know* are: firstly, *I mean* is markedly under-represented and less pragmatized while *you know* is markedly over-represented and more pragmatized; secondly, both *I mean* and *you know* are used in more restricted contexts; finally, the pragmatic functions of *I mean* are more evenly distributed while *you know* heavily depends on a very small number of functions.

Since *I mean* and *you know* are very unlikely to be taught in the classroom, the accounting for the patterns of them in the learners' data is approached from the perspective of second language acquisition. This study follows the assumption that learners' L2 production can be seen as the result of the interaction of all potential factors and the importance of a certain factor varies from one L2 phenomenon to another. The analysis seems to suggest that the learners' uses of *I mean* are greatly influenced by the congruence between the pragmatic meaning and semantic meaning of *I mean* while the learners' uses of *you know* are mainly affected by L1 influence. Other factors that seem to have impact on the production of both markers include the tasks performed by the learners and the learners' proficiency level.

List of Contents

List of Tables

Chapter One

Introduction

This thesis will investigate how two English expressions – *I mean* and *you know* – are used by Chinese EFL learners as compared to native speakers of English. These two expressions belong to a group of linguistic items which currently are most frequently termed either 'pragmatic markers' (adopted in this study) or 'discourse markers'. Other typical members of this group include expressions such as *oh, well, like, actually, because, and, but* and *so*.

One main motivation for choosing an investigation of pragmatic markers as the topic of this study is that linguistic expressions of this group are full of contrasts. On the one hand, they are just simple-looking words or constructions i.e. they are very poor in semantic meaning and are syntactically peripheral. Therefore, traditional grammar cannot say much about them. For example, in Quirk et al.'s (1972) description of English grammar *oh* is classified as an interjection which is regarded as a minor category describing "something of a museum of oddments" (Quirk et al., 1972: 411). Moreover, because of the trivialness of their appearances, they have acquired titles such as fillers, verbal fillers and gambits and are often stigmatized by native speakers. For instance, Watts (1989) carries out a study which investigates how native speakers perceive the use of pragmatic markers such as *you know*, *right*, *well* and *like*. His evidence shows that these markers are negatively evaluated. One of the speakers in his data thinks that *you know* has association with people from the north of Britain although he himself uses *you know* to preface his comment on the footballers.

For the same reason pragmatic markers would be very unlikely to attract Chinese EFL learners' attention because many of them believe only big and difficult words (words that few people know) are worth studying and knowledge of them is the key proof of one's language ability.

On the other hand, pragmatic markers are ubiquitous in conversation and a substantial number of linguistic studies on English pragmatic markers that have been conducted since 1980s have shown that in contrast to their simple appearances and social stigma, they are, in fact, pragmatically indispensible in spoken discourse.

Acting as signposts in real time communication, pragmatic markers can facilitate the hearer's understanding. According to Fox Tree and Schrock (1999), *oh* can help native speakers of English interpret information in spontaneous talk. Their findings show that recognition of words was faster after *oh* than when the *oh* was either excised and replaced by a pause or entirely excised, and semantic verification of words heard earlier in the discourse was faster after *oh* than when the *oh* was either excised and replaced by a pause or excised entirely. Flowerdew and Tauroza's (1995) test of the effect of discourse markers such as *so*, *right*, *well*, *ok* and *now* on Cantonese speakers' comprehension of English lectures shows a similar result that the subjects understood the lecture with discourse markers better than the lecture without them. Not surprisingly, a spoken discourse which lacks such markers would be detrimental to listeners' comprehension. Tyler (1992) finds that native speakers of English perceived the lecture delivered by the Chinese teaching assistant as incoherent and difficult to comprehend because it lacked discourse markers.

In addition to the function of helping the hearer process the information conveyed by utterances, Crystal (1988: 48) claims that pragmatic expressions such as *you know* can be seen as the oil which helps us "perform the complex task of spontaneous speech production and interaction smoothly and efficiently".

As opposed to native speakers' rather negative view on pragmatic markers, Olynak et al. (1990) claim that the use of pragmatic markers can be seen as a positive contributor in listeners' perception of non-native speakers' fluency. In their study of French speakers' L2 English, they compared subjects of high fluency with those of

low frequency. The result shows that more pragmatic markers were used by the high fluency group.

Similar findings are reported by Sankoff et al. (1997) who investigate how speakers of English use pragmatic markers in their L2 French. Their results indicate that "very low-frequency use of discourse markers is a signal of lesser overall linguistic competence, whereas higher frequency is the hallmark of the fluent speaker" (Sankoff et al., 1997: 204). The overall picture that emerges from their analysis suggests that "the ability to express oneself fluently and confidently in a second language entails the use of those discourse markers that native speakers produce so effortlessly" (Sankoff et al., 1997: 214).

Given the essential roles that pragmatic makers play in spontaneous talk in English and their contribution to fluency, it is self-evident that for learners of English mastering the use of them should be equally important to their acquisition of grammatical competence. Wierzbicka (1976:327) posits that learners' communicative competence would be greatly impaired if they did not understand the meaning of pragmatic markers. Svartvik (1980) even points out that learners' errors in pragmatic markers could be more problematic than their grammatical mistakes in oral communication because "if a foreign language learner says *five sheeps* or he *goed*, he can be corrected by practically every native speaker. If, on the other hand, he omits a *well*, the likely reaction will be that he is dogmatic, impolite, boring, awkward to talk to, etc., but a native speaker cannot pinpoint an 'error'" (Svartvik, 1980:171).

However, so far research in interlanguage pragmatics, which mainly focuses on the pragmatic aspects of learners' language, has not shown very much interest in how learners use these simple-looking words or constructions, which can potentially be an important indicator of learners' pragmatic competence, which is far more difficult to be assessed than grammatical skills. Most of the studies (e.g. Blum-Kulka et al.,1989;

Chen, Ye, & Zhang, 1995; Du, 1995; Liao & Bresnahan, 1996; Takahashi, 1996; Spencer-Oatey, 2000) have been almost exclusively about identifying strategies employed by learners when performing "certain functions, in particular interpersonal functions such as apologies, requests, or complaints" (Ellis, 1994: 159). The lack of studies on the use of pragmatic markers in learners' language is another important reason for carrying out the current study, which will compare the learners' performance with native speaker's, because it can be assumed that the higher pragmatic ability learners have, the closer their performance to native-speakers'.

Previous studies on learners' use of pragmatic markers also show that more in-depth studies are needed because most of them (Hays, 1992; Nikula, 1996; Romero Trillo, 2002; He, 2002; Hellermann & Vergun, 2007; Fung & Carter, 2007; Liao, 2009) only present a rather general picture of how pragmatic markers are used by learners, although learners with various L1 backgrounds are investigated. They mainly look at what pragmatic markers occur in their data and their overall frequencies. So their major findings include: compared to native speakers non-native speakers tend to 1) use pragmatic markers less frequently; 2) use a smaller number of pragmatic markers; 3) use pragmatic markers in more restricted contexts; and 4) use pragmatically transparent markers such as *and*, *because*, *but* and *so*. However, Müller's (2005) study is an exception. Instead of looking at all occurring pragmatic markers in her data, her study narrows down to four markers and her comparison between native speaker data and non-native speaker data is made on the basis of the specific pragmatic functions played by each marker. But she does not make an attempt to account for the characteristics of the non-native speakers' uses of the four pragmatic markers.

Following Müller's (2005) study, this study will aim to give an in-depth description of how pragmatic markers are used by Chinese EFL learners by focusing on two markers – *I mean* and *you know*. To give the best possible account, this study will

first investigate what exactly these two markers do in the baseline data i.e. the native speaker data, because a thorough investigation of the baseline data is essential for the following comparison between the Chinese EFL learners' data and the native speakers' data. To depict a more comprehensive picture of how *I mean* and *you know* are used by Chinese EFL learners, the comparison between the two data sets will be made not only in terms of pragmatic functions that *I mean* and *you know* play but also their positioning and co-occurring linguistic items. Although the description of the two markers will be mainly qualitative, statistical tests will be carried out where applicable. To take a step further than Müller (2005), this study will also attempt to account for the similarities and differences shown between the two groups.

There are four reasons why this study has chosen to investigate these two particular markers: *I mean* and *you know*. Firstly, they are among the most frequent pragmatic markers in English spoken discourse and they have been extensively discussed in previous studies. Secondly, they also frequently occur in the Chinese learners' data chosen by this study. Thirdly, as shown by previous studies, they are extremely elusive and to pin down their functions is very difficult because they can occur in a wide range of contexts. According to Müller (2005: 165), *you know* has been argued to play almost thirty functions in the literature and more than half of them are supported by at least two authors. In her data twelve functions have been identified.

Finally, considering *I mean* and *you know* together can sharpen our understanding of their pragmatic meanings. As shown by previous studies (Schiffrin, 1987; Brinton, 2007), the pragmatic functions of these two markers are to a greater or lesser extent influenced by their original semantic meanings, which are complementary to each other. Therefore, their overall pragmatic functions would be assumed to be complementary, as pointed out by Schiffrin (1987:309). However comparison between *I mean* and *you know* is either very brief (Schiffrin, 1987) or rather general (Fox Tree and Schrock, 2002).

When comparing their functions, Schiffrin (1987: 309) briefly mentions that "*I mean* focuses on the speaker's **own** adjustments in the **production** of his/her own talk" while "*y'know* proposes that a hearer adjust his/her orientation (specially knowledge and attention) toward the **reception** of **another**'s talk". Although Fox Tree and Schrock (2002) present a systematic comparison of *I mean* and *you know* by referring to the main functions that played by them in previous research, their comparison is carried out in only five categories: interpersonal, turn management, repair, monitoring and organizational. Since this study aims to give each marker an in-depth description, it would be a good opportunity to take a further step in their comparison i.e. to make an in-depth comparison of their pragmatic functions so as to show in what ways they are pragmatically complementary to each other. For instance, an attempt will be made to account for why there are contexts where both markers can occur. With respect to those contexts where only one of them can apply, explanations as to why *I mean* not *you know* or vice versa will be sought.

As mentioned earlier, *I mean* and *you know* have been extensively studied in the literature and a large number of proposals as to what functions they play have been made. However, instead of applying those previously identified functions directly to the chosen data by this study, this study will aim to set up its own categories on the basis of previous findings and of the uses that occur in the data analysed. The main reason for this decision is because the data that most previous studies work on tend to be rather small compared with the data of this study. For example, Őstman (1981:68) admits a small sample in his analysis of *you know*, which is based on 17 dinner conversations among which only 5 to 7 are studied in more detail. Schourup's (1985) data consist of 3 extended dyadic face to face conversations between 6 undergraduate volunteers. The data that Schriffin (1987) uses consist of sociolinguistic group interviews of 7 Jewish American speakers among whom there are 3 couples.

As the most recent study on *you know* available when the present study started, Müller's(2005) analysis uses a much bigger database, which comes from Gissen-Long Beach Chaplin Corpus. Her native speaker data consists of 53,023 words from 34 speakers while her non-native speaker data consists of 95,555 words from 77 German EFL learners. All speakers are university students. But her data only include one activity: students are asked to retell and discuss the film "Chaplin", which is seen as the limitation of the data by Müller herself as she points out that "it could not be expected that all discourse markers would be used, nor that those which occurred would be used in all their functions" (Müller, 2005: 34). Compared to Müller's data, the data for this study are not only bigger but also cover a wider range of topics. For example, the native speaker data of this study consist of 90 direct recorded conversations which took place in various venues such as homes, offices, parties and restaurants and were produced by not only university students (although majority of the speakers are) but also university staff, writer, journalist, musician, dentist, accountant, carpenter and housewife. Although it could not be expected that all functions of *I mean* and *you know* would occur in the native speaker data of this study either, different or more functions of them can be expected to emerge.

To achieve the goal of revealing how *I mean* and *you know* are used by Chinese EFL learners as compared to native speakers, the present study will address the following questions :

1) What pragmatic functions do *I mean* and *you know* play in L1 English conversation?
2) What are the similarities and differences between *I mean* and *you know* with regard to their pragmatic functions and why? What are the similarities and differences between them in terms of distributions and positioning?
3) What are the characteristics of the Chinese EFL learners' use of *I mean* as compared to the native speakers of English? What are the potential

factors that could contribute to the similarities and differences between them and why?

4) What are the characteristics of the Chinese EFL learners' use of *you know* as compared to the native speakers of English? What are the potential factors that could contribute to the similarities and differences between them and why?

The remaining chapters of this thesis will be organized as follows. Chapter 2 will aim to present a general picture of research in the field of pragmatic markers or discourse markers by giving an account of its historical development, main theoretical and methodological issues such as its terminology, classifications, linguistic theories and analytical models that have been used. Since this study, which is purely descriptive, has chosen two linguistic expressions which have the least contention as to their membership status, it will not attempt to make contributions in the discussion of theoretical issues. However, explanations will be given as to why certain positions on relevant theoretical and methodological issues have been taken in this study. In chapter 3, detailed information of the data used in this study will be provided. Also, analytical terms used in the analysis will be defined or clarified.

Chapter 4 will give a detailed account of how the pragmatic functions of *I mean* and *you know* are identified and how a modified analytical framework came into being by drawing upon previous findings including a pre-existing framework and by adapting this to cover all the instances which are found in my native speaker data. This chapter will, then, end up with an in-depth comparison between *I mean* and *you know*, which are analyzed under the same framework and on the basis of the same data set.

Chapter 5 will show the characteristics of *I mean* and *you know* in the Chinese EFL learners as compared to the British speakers in terms of statistical findings, which include comparisons of overall frequencies, the distributions of subfunctions and

distributions of positioning between the non-native speaker data and the native speaker data.

Chapter 6 will seek to account for the similarities and differences shown by Chapter 5. Among various potential variables that may have an impact on learners' L2 English, the features of how *I mean* is used by the Chinese learners will be interpreted mainly from the perspective of cognitive constraints on acquisition while explanations of those of how *you know* is used in the learners' data will focus on the learners' L1 influence. To know if learners' L1 influences their L2 English, we need to know how the Chinese equivalents of *I mean* and *you know – wo de yisi shi* and *ni zhidao* – are used in L1 Chinese. Since currently there has not much research done on them, this chapter will also contain a small-scale study of how *wo de yisi shi* and *ni zhidao* behave in L1 Chinese and their behavior will be compared to how *I mean* and *you know* are used by the Chinese EFL learners. Finally, chapter 7 will summarize the major findings of present study.

Chapter Two

Literature review

This chapter consists of three main sections. The first section will aim to give an overview of the historical development of the research on pragmatic markers/discourse markers. In section 2.2 theoretical and methodological issues in this research area will be reviewed. Since this study is purely descriptive and there is a general agreement among researchers as to the membership status of the two pragmatic markers chosen by this study, this study will not attempt to make contributions to theoretical issues such as definition and classification or to propose a new method of analysis. The main purpose of section 2.2 is to provide relevant background information so as to explain why certain theoretical or methodological positions have been taken in this study. Finally, section 2.3 will give a review of previous studies on L2 learners' use of pragmatic markers.

2.1. A panoramic view of research on pragmatic markers

Previous research of pragmatic markers or discourse markers seems to show that such research (e.g. Lakoff, 1973; Wierzbicka, 1976; Keller, 1979; Murray, 1979) started in the 1970s when the focus of linguistic study began to shift from grammaticality in terms of syntax to the appropriateness of a sentence in a given context (Lakoff , 1973: 453). Interestingly, Lakoff's (1973) study on *well*, the first study on *well*, does not aim to explore the pragmatic functions of *well*. Instead, he is interested in appropriateness of an answer to a question following Labov's (1970) study of appropriateness of a question. The reason that *well* appears in his observation seems to be because the injection – *well* – often occurs in the context where it is used to preface an answer. Lakoff (1973) notes that one constraint on *well* is that it cannot preface a direct answer to a question. As shown by example 2.1, which he quotes from Garner (1969), *well* cannot preface the first answer i.e. the direct answer – three o'clock.

2.1

Q What time is it?

A: Three o'clock.

My stomach says that it is nearly lunch time.

The sun just came up

(Lakoff, 1973: 455, but the list of answers is shortened)

1980s appears to be a very important period for research of pragmatic markers because it witnesses not only increasingly more attention from researchers but also a number of influential publications, which have paved the way for this subfield of linguistic research to flourish from the 1990s onwards. Schourup (1985) gives a book-length description of pragmatic markers such as *well*, *oh*, *like*, *well*, *you know*, *I mean*, *mind you* and *sort of* by applying his "tripartite model" (Schourup 1985:7). Schiffrin's (1987) study, another book-length account of *oh*, *well*, *and*, *but*, *or*, *so*, *because*, *now*, *you know* and *I mean*, has become one of the most often cited studies. Her model of five planes of talk has been applied by later studies (e.g. Salmons, 1990; Hays, 1992; Demirci & Kleiner, 1997; Kyratzis & Ervin-Tripp, 1999; Fung & Carter, 2007).

In addition to publications of descriptions of pragmatic markers, there are also a number of theories developed during this period which are extensively applied in later research. The first theory is Sperber & Wilson's (1986) Relevance Theory, which is further developed by Blakemore (1987, 1988, 1992). Blakemore's relevance-based model seems to attract a bigger number of users (Watts, 1988; Blass, 1990, 1992; Jucker, 1993; Lamiroy, 1994; Takahara, 1999; Andersen, 2000) than Schiffrin's (1987) model. Another important theory that is applied in the analysis of pragmatic markers is Brown & Levinson's (1987) Politeness Theory. All the major models or theories will be further reviewed later in section 2.3. Schiffrin (1987) and Fraser (1988) initiate the discussions on theoretical issues such as how to define and classify these problematic expressions, which tend to be kept in a ragbag by traditional grammar.

The booming of research on pragmatic markers since 1990s can be shown in the following aspects. Firstly, interest in this group of linguistic expressions has spread from the English language to a number of other languages such as Italian (Bazzanella, 1990), French (Olynak et al., 1990; Lamiroy, 1994; Sankoff et al., 1997; Hansen, 1998; Andersen et al. 1999), Spanish (De Fina, 1997), German (Meng & Schrabback, 1999), Latin (Kroon, 1995), Japanese (Takahara, 1999; Park, 1998), Korean (Park, 1998) and Chinese (Chen & He, 2001; Lee-Wong, 2001; Feng, 2008; Liu, 2009) although the English language is still the one that has received the most attention from researchers. As a result, the later discussion on theoretical issues of pragmatic markers will be mainly based on studies on English pragmatic markers.

Secondly, a range of variables that can affect the use of pragmatic markers have been investigated. Among them, age seems to be the one which has attracted the most attention. For instance, Stenström (1990) carries out a corpus-based comparative study of pragmatic markers used by teenagers and adults in London. Meng & Schrabback (1999) compare how children and adults use two German interjections – *hm* and *na*. Romero Trillo (2002) compares children and adults' use of English discourse markers. Hasund (2002) conducts a corpus-based comparative study of *like* used by teenagers between English and Norwegian.

Another variable that interests researchers seem to be the impact that speech context can have on use of pragmatic markers. Chen & He (2001) study how the Chinese pragmatic marker, *duibudui*, is used in the classroom. De Fina (1997) compares the use of an L1 Spanish pragmatic marker, *bien*, in the classroom and conversation. Fuller (2003) shows the differences in the use of discourse markers between two speech contexts: interviews and conversations. Fung & Carter (2007) study reveals features of discourse markers in the speech of pupils in secondary classrooms in Hong Kong.

Other variables that have been explored include social class and gender. For example, Huspek (1989) analyzes *you know* and *I think* variation in some American industrial workers' speech which tends to be associated with very high frequency of *you know* and relatively low frequency of *I think*. Stubbe and Holmes (1995) compare pragmatic markers between middle class and working class. Holmes' (1986) studies characteristics of women's use of *you know,* which tends to be regarded as a "women's language" form as compared to men's. Erman (1992) makes her comparison between female and male in two different interactions: same-sex and mixed-sex interactions.

Finally, pragmatic markers have become a platform where researchers from different research fields converge. One group that has been attracted by pragmatic markers is researchers whose interest is in language change. Their research on pragmatic markers has become a very important part of study of grammaticalization. The question they aim to address is what changes have taken place between pragmatic markers and their homomorphous forms. Thompson & Mulac (1991) carry out a study on the epistemic parenthetical – *I think*. They argue that since *I think* has taken on the pragmatic function of "expressing the degree of speaker commitment, functionally roughly as an epistemic adverb such as *maybe* with respect to the clause it is associated with" (Thompson & Mulac 199l: 313), *I think* should not be viewed as the "that-deletion" alternation of the construction *I think that* where *I* is the subject and *think* is the main verb. Traugott (1995) also observes that expressions such as *I think*, *even*, *actually*, and *really* have gone through "a pragmatic-semantic process whereby 'meanings' become increasingly based in the speaker's *subjective belief state/attitude toward the proposition*, in other words towards what the speaker is talking about" (Traugott, 1989: 35).

The search for evidence of grammaticalization of lexical items has been carried out not only synchronically but also diachronically. Brinton has done several studies on

looking for evidence of when a lexical item begins to take on pragmatic meanings from Old English to Modern English. For instance, Brinton (1998) presents diachronic evidence of grammaticalization of *only* from free adverb to adversative conjunction in English while Brinton (2007) gives evidence of the historical development of the pragmatic marker – *I mean*.

Researchers who are specialized in language contact also show their interest in pragmatic markers. What they aim to reveal is how pragmatic markers are used by bilingual speakers. There are quite a few studies (Salmons, 1990; Goss & Salmons, 2000; Fuller, 2001) which are about German-English bilinguals. They seem to have rather similar findings i.e. in German discourse the system of German discourse markers appears to have been replaced by the English system (Salmons, 1990). The English discourse markers seem to be first borrowed into the German language and then replace the German markers (Goss & Salmons, 2001; Fuller, 2001).

Maschler's (1994) investigation of Hebrew-English bilinguals shows a different and interesting picture. She notes that a recurring pattern in the English conversations of the Hebrew-English bilingual speakers i.e. they often switched to Hebrew when they needed to use discourse markers. Sankoff et al. (1997) also reports that the French-English bilingual speakers switched to English discourse markers in French discourse, though to a very slight degree.

It is not surprising at all that pragmatic markers have become a topic for the next group which consists of researchers who show interests in contrastive or cross-linguistic studies and translation because pragmatic markers as a small subsystem of language would be an ideal candidate for comparative studies across languages. As Wierzbicka (1976: 366) points out languages are too complex to be compared as wholes. What should be compared are subsystems "which are small enough to be examined in their entirety, and which at the same time promise to be

particularly revealing" rather than "a random selection of isolated features of grammatical and lexical systems".

Ariel (1994: 3252) also contends that the subsystem of pragmatic markers appear to be universal because even unrelated languages can have rather similar pragmatic markers. But their universality differs in different functional levels. Markers, which mainly function as discourse organizers, will occur universally, "although their formalization as well as frequency of use may vary" while markers functioning on the interpersonal level will show more social or cultural variations, hence will be less universal.

English has been compared with quite a few other languages such as German (Fischer & Drescher, 1996), Scandinavian dialects (Andersen, 1997), Korean (Park 1998), Japanese (Takahara, 1998) and Chinese (Zhang, 2007; Chen, 2011). Fischer & Drescher (1996) conduct their comparison between two of Woody Allen's dramas – Death and God – and their German translations. The English pragmatic marker *well* has been found to have 10 equivalents (*also, naja, na, komm, ach, na also, ja, tja, schon, gut*) in German. This finding echoes Kroon's (1995:17) discussion of the English discourse marker *because,* which is again a multifunctional marker because it can occur in three types of utterance relations: cause, evidence and justification. But French needs to use three different markers i.e. the English marker *because* has three French equivalents: *parce que, puisque* and *car*. Fischer & Drescher's (1996) findings show that contrastive analysis, i.e. translation, is very helpful in terms of revealing functions of pragmatic markers.

Andersen's (1997) cross-linguistic study of comparing English pragmatic markers to their equivalents in three Scandinavian urban dialects aims to prove if universal features of teenage speech exist or not. This cross-linguistic survey of pragmatic markers seems to support the view that similarities of teenage speech across

languages do exist. For instance, the English pragmatic *like* has a parallel in all the three Scandinavian dialects (*liksom/typ* in Stockholm and Bergen, *lissom/ligesom* in Aarhus). All the equivalents have similar and pragmatic meanings of *like*.

As can be seen from the above contrastive studies of pragmatic markers, translation is a very useful technique. So it is not surprising that experts of translation have also noticed this linguistic phenomenon. Bazzanella & Morra (2000) consider the translation of pragmatic markers as a crucial and stimulating area for translation theory. They suggest that the priority of translation of pragmatic markers should be to preserve the functionality of the marker in question. Therefore, translators have to assign temporary and context-specific translations. Sometimes even mutually exclusive translations can be used for the same marker, for example, *well* can be used in either agreement or disagreement.

Interestingly, pragmatic expressions have even attracted attention from child language acquisition, a research area which seems to be even more dominated by studies of child's acquisition of morphology and syntax than SLA. Although study of how a child acquires pragmatic markers has just started, some interesting findings have been reported. Andersen et al. (1999) compare how American English-speaking, Lyonnais French-speaking and Chicano Spanish-speaking children use discourse markers in conversations where they role-play with puppets representing a variety roles such as father and mother in family settings, and doctor, nurse and patient. The comparison shows striking cross-linguistic parallels in the way children learn to use discourse markers. The three groups all show similar preferences in discourse markers. For example, when playing roles of higher status such as mum or doctor, English-speaking children use more *well, so, now, then*; French speaking-children use more *alors, bon (ben; et ben), maintenant, allez,* and Spanish-speaking children use more *bueno, pues, ahora/ahorita, entonces/luego*. When playing lower-status roles such as baby and patient, *uh* (English), *euh* (French), or *eh* (Spanish) show

higher frequencies.

Montes (1999) reports a longitudinal case study of how a Spanish-speaking child (1;7 to 3;0) acquires discourse markers such as *ah, oh, ay, oy, uy and eh*. The observation shows that the child started from the more literal, contextual uses to the more elaborated discursive functions which emerged later. For example, the meaning of *I see* as *I understand* emerged later than did the literal meaning of *I see* (an object).

Kyratzis & Ervin-Tripp (1999) investigate the development of acquisition of discourse markers by having four- and seven-year children paired in best friend dyads perform role play and story-telling. One of the differences between the two age groups suggests that different functions of discourse markers may be acquired at different ages. The following two examples show that function of justification of *because* may be acquired earlier than cause.

2.2

(4-year-old girl) Jan

12 okay [puts the other human figure aside]
13 this is me. so I'm supposed to act her
14 and you're supposed to act her
15 *because* you're the mom/[hums] (p1333)
(Kyratzis and Ervin-Tripp 1999: 1333, quoted as example (12))

2.3

(7-year-old girl) Lyn

5 you know, she cries
6 *because* her brother turned into a little bear (p1331)
(Kyratzis and Ervin-Tripp 1999:1331, quoted as example (11))

Lastly, even forensic linguistics has shown its interest in pragmatic makers. Blackwell's (2000) analysis shows that discourse markers can play a role in forensic purposes which heavily depend on analysis of content words. She points out that "discourse markers such as *well*, *honest* and *look* are an essential feature of natural spoken language, and their patterns of occurrence in a disputed text (e.g. transcribed

police interview) may aid the forensic linguist in evaluating conflicting claims about the text's origin" (Blackwell, 2000: 13).

2.2 Theoretical and methodological issues in the research of pragmatic markers

This section consists of 4 subsections. Section 2.1 will look at the issue of terminology, which "presents a particular difficulty" (Schourup, 1999:228). A review of the various proposals as to which term should be used in the literature will be presented and explanation as to why the term "pragmatic marker" has been chosen by the current study will be given. Section 2.2 will discuss the main features and functions of pragmatic markers and classification of pragmatic markers. Section 2.3 will be a review of the main theories or models used previous studies of pragmatic markers. Section 2.4 will review various methods used to establish pragmatic functions of pragmatic markers in previous research and explain how this study has made its decision as to what methods should be used.

2.2.1 Terminology

By searching the references of pragmatic markers for this study more than 10 candidates have been found, but currently only a number of them have remained in use. Those terms that seem to have died out include gambit (Keller, 1979), speech marker (Olynak et al., 1990), pragmatic particle (Őstman, 1981, 1982, 1995; Foolen, 1997), phatic connective (Bazzanella, 1990), pragmatic connective (Lamiroy, 1994), discourse connective (Blakemore, 1987; Unger, 1996), discourse item (Stenström, 1990); pragmatic force modifier (Nikula, 1996); discourse operator (Redeker, 1991) and pragmatic operator (Ariel, 1994). One main reason for these terms being abandoned might be that they are either too general or too narrow. For example "speech marker" sounds a very general term because it does not necessarily have to be confined to linguistic expressions. "Discourse item" seems to be general as well because it could include any item in a discourse. On the other hand, "discourse connective" would be too narrow because it highlights the linking function, which is

only one of the main functions of these small words. The word ‘operator’ in ‘discourse operator’ or ‘pragmatic operator’ can be associated with conjunctions, which are represented by logical truth-tables such as &, V, >, →, : and ↔ in formal semantics. The reason for dropping the term “gambit” is obviously due to its negative connotation, which suggests that this group of expressions does not have any important function and is not worth studying.

The most frequently used terms are discourse marker (Schiffrin, 1985, 1987; Watts, 1989; Brinton, 1990; Salmons, 1990, Blakemore, 2002; Fuller, 2003; Müller, 2005; Fung & Carter, 2007), pragmatic marker (Watts, 1988; Redeker, 1990; Andersen, 1997, 1998, 2000, 2001; Aijmer & Simon-Vandenbergen, 2004; Feng, 2008; Norrick, 2009) and discourse particle/particle (Wierzbicka, 1976; Svartvik, 1980; Blass, 1990; Abraham, 1991; Kryk, 1992; Fischer & Drescher 1996; Van Barr, 1996; Schourup, 1985; Bolinger, 1989; Hasen, 1998; Lee-Wong, 2001; Aijmer, 2002). Among them ‘discourse marker’ is the most preferred candidate while ‘discourse particle/particle’ and ‘pragmatic marker’ show a similar frequency. However, discourse particle/particle shows a tendency of dying out. The main reason for objections to ‘discourse particle/ particle’ is because ‘particle’ has traditionally been used to refer to a unified grammatical category (e.g. modal particles in German and Chinese) while the linguistic expressions under study in this group either cannot fit into any well-established category i.e. they tend to be put in a waste basket or come from various categories such verbs, conjunctions, prepositions, adverbials and interjections. Most researchers who use ‘discourse particle/particle’ tend to study those languages which are rich in particles such as German (Abraham, 1991; Fischer & Drescher, 1996), Polish (Kryk, 1992), Latin (Kroon, 1995); French (Hasen, 1998) and Chinese (Chappell, 1991; Lee-Wong, 2001).

The vast number of competitors is not the only reason that contributes to the complicated picture of terminology. Another problem is that the same author would

use different terms while the same term used by different authors may not refer to the same group of linguistic items. For instance in her earlier studies (Erman, 1986, 1987, 1992) Erman uses 'pragmatic expression' but changes to 'pragmatic marker' in her later study in 2001. Schourup (1985) first uses 'discourse particle' and switches to 'discourse marker' in 1999. Aijmer changes her term from 'discourse particle' (Aijmer, 2002) to 'pragmatic marker' (Aijmer & Simon-Vandenbergen, 2004). Although 'discourse marker' has become the most popular term, it does not always have the same reference. Some researchers follow Fraser's (1988, 1990, 1996, 1998, 1999) definition which treats 'pragmatic marker' as the cover term and 'discourse marker' is one type of 'pragmatic marker' while many more others follow Schiffrin (1985, 1987) who uses 'discourse marker' to cover a wide range of linguistic expressions including *you know* and *I mean*, which are not considered as discourse markers but another subcategory of pragmatic markers – parallel markers (for details of Fraser's classification of pragmatic markers, see section 2.2).

Despite the fact that discourse marker is the most preferred term, this study has chosen pragmatic marker because firstly this study aims to reveal the pragmatic functions of *I mean* and *you know* and secondly it best catches the main features of *I mean* and *you know*, which do not contribute to propositions they work with either semantically or syntactically. Like Erman (1986, 1987), I take the view that discourse marker has the problem of confining *I mean* and *you know* to a single function i.e. organizing discourse. Using the term pragmatic marker will allow me to examine their functions from as many angles as possible.

2.2.2 Features, functions and classification of pragmatic markers

Like the confusing terminology, the definition of pragmatic markers is not straightforward either. There is a substantial variation as to the number of linguistic expressions that should be classified as pragmatic markers. According to Brinton (1998:11) the number ranges from a dozen to five hundred and the most frequently

discussed pragmatic markers include *actually, anyway, I mean, now, so, then, well, you know* and *you see*. To define pragmatic markers, two questions need to be answered: 1) what characteristics do pragmatic markers have? 2) what functions do pragmatic markers play? Section 2.2.2.1 and section 2.2.2.2 will look at the answers to the above two questions proposed in the literature respectively. Section 2.2.2.3 will review the issue of the classification of pragmatic markers.

2.2.2.1 Features of pragmatic markers

A number of researchers have contributed to the description of characteristics of pragmatic markers. Svartvik (1980) summarizes the features of *well* as a pragmatic marker as compared to *well* as a manner adverb or as a degree word.

1) it is very difficult to decide which function or word class it should belong to
2) it is very difficult for lexicographers to pin down what it means
3) it is very difficult to translate it idiomatically into another language.
4) it only occurs in spoken language with a very high frequency

(A summary of the features of *well* as a pragmatic marker based on Svartvik (1980: 168-169))

Őstman (1982, 1995) proposes features of pragmatic markers on the basis of structural and functional criteria. Unlike Svartvik (1980) whose observation is made on one English marker, *well*, Őstman (1982:149) has a group of cross-linguistic expressions in mind because he believes that pragmatic markers are a universal phenomenon.

1) they are short
2) they are prosodically subordinated to another word
3) they do not contribute to the propositional content of the sentence
4) they tend to occur outside an utterance and modify that utterance as a whole

(A summary of features of pragmatic markers based on Őstman (1982: 150))

He points out that criterion 3) – propositional emptiness criterion – should be more important than the others (Őstman, 1982: 150), which explains Svartvik's (1980)

observation that describing what *well* as a pragmatic marker means is a difficult job.

Brinton (1990: 46-47) presents a longer list of the features of pragmatic markers. She includes features of high frequency, orality and akwardness of fitting into any word-class on Svartvik's (1980) list and features of being phonologically short and unstressed, semantically empty and syntactically optional in Őstman's (1982, 1995) proposal. An additional feature in her list is the positioning of pragmatic markers: they are very flexible although sentence-initial is the preferred position.

1) they predominantly occur in spoken discourse
2) they are very frequent
3) they are short
4) they are phonologically reduced or unstressed
5) they occur either outside the syntactic structure or loosely attached to it
6) they are very flexible in terms of positioning but generally sentence-initial
7) they are difficult to place within a traditional word class

(A summary of features of pragmatic markers based on Brinton (1990: 46-47))

As can be seen from the following list proposed by Hansen (1998), the additional feature is the observation that pragmatic markers tend to cluster. By comparing with the above lists, we can conclude that researchers seem to agree that being semantically empty and syntactically optional are the two key characteristics of pragmatic markers.

1) they take the entire host sentences, proposition, or utterance in their scope
2) they normally do not carry stress
3) they cannot be coordinated but can and often do cluster
4) they do not contribute to propositional meaning
5) they are grammatically optional

(A summary of features of pragmatic functions based on Hansen (1998: 42))

Like Őstman (1982, 1995) and Jucker & Ziv (1998), these features are viewed as prototypical rather than defining in this study. Those expressions which display more of the features can be considered as prototypical members while those which show fewer characteristics may be considered more peripheral. According to Jucker & Ziv (1998:3), one advantage of this scalar conception is that it can better account for "the range of items displaying partially overlapping characteristics across a variety of languages." For instance, Őstman (1982: 99) notes that English pragmatic markers do not follow his criterion 2) ("pragmatic markers are prosodically subordinate to another word").

Semantic emptiness and syntactical optionality are considered as the most essential features of pragmatic markers in this study. They are used as the key criteria to distinguish pragmatic use from non-pragmatic of *I mean* and *you know* by this study.

2.2.2.2 Functions of pragmatic markers

From the following review of the functions of pragmatic markers in the literature, we can see another feature of pragmatic markers emerge: they are multifunctional. Őstman (1982: 152) claims that pragmatic markers perform two types of tasks: interactional and attitudinal. The interactional functions are either sociological or discourse-functional, while the attitudinal functions mainly focus on cognitive and psychological aspects. Brinton (1990) gives a list of functions that pragmatic markers play, shown below. Unlike Őstman (1982) who describes functions of pragmatic markers in very broad categories, Brinton (1990) lists rather specific functions. Most functions on her list can actually fit into Őstman's (1982) interactional category except functions 4) and 8). They appear to fit into the attitudinal category because there is a cognitive motivation behind using a pragmatic marker to buy more time or to mark foregrounded or backgrounded information. The former benefits the speaker by winning more time to organize his/her thoughts while the latter benefits the hearer by making the information processing easier.

1) initiate discourse
2) to mark a boundary in discourse, i.e. to indicate a shift or partial shift in topic
3) to preface a response or a reaction
4) to serve as a filler or delaying tactic
5) to aid the speaker in holding the floor
6) to effect an interaction or sharing between speaker and hearer
7) to bracket the discourse either cataphorically or anaphorically
8) to mark either foregrounded or backgrounded information

(A summary of functions of pragmatic markers based on Brinton (1990: 47-48))

Brinton (1998) makes another proposal of functions of pragmatic makers. Compared to her earlier list (Brinton, 1990), this proposal includes even more specific functions, which are classified into two broad categories in Halliday & Hasan's (1976: 26-28) terms – textual and interpersonal.

Textual functions

1) getting the hearer's attention,
2) initiating and ending discourse,
3) sustaining discourse, marking boundaries (topic shifts and episode boundaries),
4) constraining the relevance of adjoining clauses,
5) repairing discourse

Interpersonal functions

1) subjective functions such as expressing response, reactions, attitudes, understanding, tentativeness, or continued attention
2) interactive functions such as expressing intimacy, cooperation, shared knowledge, deference, or face-saving (politeness).

(A summary of functions of pragmatic functions based on Brinton (1998: 12))

2.2.2.3 Classification of pragmatic markers

Bruce Fraser (1988, 1990, 1996, 1998, 1999) has contributed a great deal to theoretical issues of pragmatic markers, such as how to define, classify and approach pragmatic markers. Fraser (1996) gives a very detailed account of his classification of pragmatic markers. Before discussing his classification in detail, it should be pointed out that in his system "pragmatic marker" and "discourse marker" are hyponymous i.e. discourse marker is one type of pragmatic marker while many other researchers treat these two terms as synonyms. The broadness of "pragmatic marker" in Fraser's system is reflected in its definition, which is based on the assumption that sentence meaning can be divided into two separate and distinct parts: propositional meaning and non-propositional meaning. Fraser (1996:167) has called the non-propositional part of sentence meaning Pragmatic Markers. Table 2.1 below is a summary of Fraser's (1996) taxonomy of pragmatic markers.

Table 2.1 Taxonomy of pragmatic markers (Fraser 1996)

Main categories	Subcategories	Examples	
Basic markers: *contributing conceptual information over and above that of the propositional meaning*	Structural basic markers	mood (declarative, interrogative or imperative)	
	Lexical basic markers	Performative expressions	***I promise*… *I (hereby) apologize…***
		Pragmatic expressions	***please (kindly), perhaps (maybe)***
	Hybrid basic markers	Declarative-based hybrids	Tag questions, positive tag questions,
		Interrogative-based hybrids	***Can (can't) you...?*** ***Why not…?*** ***May I…?*** ***I suggest that*…**
		Imperative-based hybrids	***Talk, or I'll shoot.*** ***Wash, and I'll dry.***
Commentary pragmatic markers: *having propositional meaning over the entire message and procedural meaning signaling the message as a comment*	Assessment markers	***Amazingly***, Derrick passed the exam.	
	Manner-of-speaking markers	***Frankly***, you need to stop now	
	Evidential markers	***Conceivably***, Tim is right.	
	Consequent-effect markers	***To clarify***, no one is permitted to smoke in this building!	
	Hearsay markers	***Reportedly***, the game was postponed because of rain.	
	Mitigation markers	***If you don't mind***, bring it to me about 7 this evening.	
Parallel markers: *signaling an entire message in addition to the basic message*	Vocative markers	Standard titles e.g. ***Mr President*** Occupation names: ***doctor, waiter*** General nouns: ***Brother, boy, guys*** Pronominal forms***: you, everyone***	
	Speaker displeasure markers	***Damned, damn well, right now, the hell***	
	Solidarity markers	***My friend***, we simply have to act together and face this problem.	
Discourse markers:	Topic change	***back to my original point, before I forget, by***	

signaling the relationship of the basic message to the foregoing discourse and providing instructions to the addressee on how the utterance to which the discourse marker is attached is to be interpreted.	markers	***the way***
	Contrastive markers	***Anyway, but, however, on the other hand***
	Elaborative markers	***above all, in other words, what is more, in addition…***
	Inferential markers	***after all, so, as a consequence, therefore, as a result, because of this/that …***

Obviously discourse marker is a very important category among the four types of pragmatic markers in Fraser's classification because most of his publications focus on discourse markers e.g. how to define discourse markers (Fraser, 1999), how to approach discourse markers (Fraser, 1990), and how to classify English discourse markers (Fraser, 1998). Although discourse marker is only one group of pragmatic markers in Fraser's classification, there is a key overlapping between the main features of Fraser's discourse markers and those of pragmatic makers or discourse markers discussed in section 2.2.1 where discourse markers, pragmatic markers or many other terms are seen as competing candidates for the same group of linguistic expressions. Fraser (1990: 387-390) lists the following features of discourse markers when discussing how to study them. Among them, characteristic 4), i.e. zero semantic and syntactic participation, is among the most frequently proposed features in the literature, hence the most essential characteristic.

1) they are very difficult to analyze by traditional grammatical categories
2) they will not serve both roles (discourse marker and its homophonous form) in the same sentence.
3) they typically occur only in utterance-initial position but they do occur in utterance-internal or utterance-final position.
4) their presence or absence does not alter the potential discourse relationship between the message which follows and the foregoing discourse.

Instead of grouping pragmatic markers on the basis of the functions they play as Fraser (1990) does, Ariel's (1994: 3251) classification is made on the correlation between form and function. The first category is termed 'transparent operators' ('pragmatic operator' is used instead of 'pragmatic marker' by Ariel) because their pragmatic meaning can be directly inferred from their semantic meaning. Examples of this category include *but, or, so* and *because*. The second category is called 'intermediate operators' because although their function can be related to their semantic import, it cannot be derived automatically. Pragmatic markers such as *you know, I mean, after all, of course,* and *or something* are under this category. The last category is labeled 'opaque operators' because there is almost no link shown between their function and semantic meaning. Pragmatic markers such as *oh* and *well* can be seen as members of this group. Ariel (1994) claims that most pragmatic markers are in "somewhere on a continuum ranging from transparent linguistic expressions to completely opaque expressions interpreted by reference to their idiosyncratic rules of use" (Ariel 1994: 3251).

Ariel's (1994) classification has been found very helpful in accounting for acquisition of pragmatic markers. Hays (1992) observes that the most frequent pragmatic markers used by the Japanese EFL learners are *and, but* and *so,* which are on the transparent end of the continuum. Nikula's (1996: 89) comparison between advanced Finnish EFL learners and British speakers reveals that the non-native speakers find it easier to resort to explicit markers, which have relatively transparent meanings. Fung & Carter (2007) report similar findings. The Hong Kong learners in their study are found to show a liberal use of pragmatic markers such as *and, but, because,* and *so* but a relatively restricted use of markers such as *yeah, really, say, sort of, I see, you see, well, right, actually, cos,* and *you know.* The reason that learners tend to use more transparent markers than intermediate or opaque markers could be that the more transparent the markers are the less difficult for learners to acquire. Since *I mean* and *you know* are both under the intermediate category, this

study will look at the correlation between the specific functions and the semantic meaning when accounting for the characteristics of learners' use of these two markers. However, a similar continuum of the congruence between pragmatic meaning and semantic meaning is used.

2.2.3 Main theories or models used in previous studies of pragmatic markers

In this section, Schourup's (1985) Tripartite Model, Schiffrin's (1987) Discourse Coherence Model – Five Planes of Talk, Blakemore's (1987, 1992) Relevance Model and Brown and Levinson's (1987) Politeness Theory will be discussed. As can be seen from the following review, Schourup's and Schiffrin's models are designed for analyzing functions of pragmatic markers while the Relevance Model and Politeness Theory are not. The Relevance Model is a pragmatic model developed to "account for the psychological constraints that govern the hearer's choice of context for the interpretation of a given utterance" (Blakemore, 1987: 12). It has been applied to this research field because pragmatic markers used by the speaker can help accommodate the hearer's cognitive need of real time processing of the utterance. Politeness Theory is developed to account for the redressing strategies used by the speaker in performing face-threatening acts so as to minimize the detriment to both speaker and hearer's face in social interaction. It has been applied in studies of pragmatic markers because it can account for the speaker's social motivation in using pragmatic markers in conversation.

2.2.3.1 Schourup's (1985) Tripartite Model

The core of Schourup's model is his proposal of considering unexpressed thinking of conversants in describing conversation because conversants are not only talking but also thinking and their unexpressed thoughts have a direct impact on what they are saying. Although unexpressed thinking is inaccessible to direct observation, its existence should not be denied.

Schourup (1985:4) posits that unexpressed thinking is a very important aspect of conversation because conversants do not say everything that crosses their mind "but judiciously retain, shape, and sequentially place them in ways that often display considerable gamesmanship". Given the importance of the unexpressed thinking of conversants in conversation, he proposes that the situation of a participant in a conversation should consist of three worlds:

> The covert thinking of the speaker, what that speaker has presently in mind but has not disclosed, will be referred to below as the private world; what is on display as talk and other behavior on the part of conversants and is thus available to both the speaker and any other(s) will be called the shared world; and the covert thinking of other conversants, which is inaccessible to the speaker, will be called the other world.
>
> (Schourup, 1985: 7)

Schourup (1985:3) also proposes that the role played by pragmatic markers (discourse particle is used in Schourup's study) is mediating in a specific way between the private world, the other world and the shared world.

There does not seem any application of Schourup's model in the studies of pragmatic markers reviewed in this study except one study – Salmons' (1990) study of bilingual discourse markers. He compares the applications of Schiffrin's (1987) Coherence Model and Schourup's model in his German-American corpus. However, the Tripartite Model does attain its goal of identifying a core use for the pragmatic markers considered in Schourup's study. Given that the role of pragmatic markers is defined as the way that pragmatic markers mediate between the private world, the other world and the shared world, the core meaning of *I mean* in Schourup's framework is to indicate what the speaker has said and what the speaker has in mind are not well matched i.e. *I mean* indicates a nonequivalence of what is in the shared world and what is in the private world. Interestingly, under this framework the pragmatic marker *like* is found to play the same role as *I mean*, i.e. it mediates between the shared world and the private in the same way as *I mean* does. The explanation of the difference between *like* and *I mean* offered by Schourup (1985: 147-148) is: *like* indicates a minor nonequivalence which is unlikely to cause

misunderstanding of the intended meaning while *I mean* indicates a substantial nonequivalence which requires correction or would lead to misunderstanding. Such a difference does provide a plausible explanation as to why *I mean* is often used to preface corrections while *like* is not.

The core function of *you know* in this Tripartite Model is to indicate that "the speaker expects that there is no communicatively significant discrepancy between what is now in the private world and what is now in the other world, with respect to what is now in the shared world" (Schourup, 1985:102). So one obvious difference between *I mean* and *you know* under this framework is that *you know* involves the hearer while *I mean* does not, which does match to the general impression described by other researchers i.e. *I mean* is more speaker-oriented while *you know* is more hearer-oriented. But the core meanings of *I mean* and *you know* identified by this model do not seem to fit into the description that the pragmatic functions of these two markers are complementary. In addition, this model does not seem to be able to reveal any more specific differences between *I mean* and *you know*. Since one main aim of this study is to reveal the specific contexts of these two markers, Schourup's Tripartite Model will not be applied in this study.

2.2.3.2 Schiffrin's (1987) Discourse Coherence Model

Schiffrin's (1987) investigation of a number of pragmatic markers is one of the most important studies in the literature. Redeker (1991:1139) points out that Schiffrin's (1987) study is "an ambitious and valuable attempt to integrate meticulous empirical analysis of the functions of discourse markers with a theory of discourse coherence".

According to Schiffrin (1987), the goal of developing the Coherence Model is to provide not only a theoretical background for her study of discourse markers but also a model on which both her analysis of specific markers and her general conclusion are based. Schiffrin (1987:3) maintains that the key assumptions of language central

to her analysis are: language always has a context, is context sensitive, always communicative and designed for communication. She continues to propose that discourse has three main properties: structure, meaning and action. Structure and meaning are more about discourse "as extended sequences of smaller units, e.g. sentences, propositions and utterances while action is "more concerned with language as it is used in a social interaction" (Schiffrin, 1987: 7). These three properties of discourse interact with each other because language has structure, conveys meaning, and is used to perform actions.

Schiffrin (1987) sees coherence as crucial to discourse for it distinguishes discourse from random collections of smaller units. Like speakers and hearers who make great efforts to construct coherence in conversation, discourse analysts strive for a description of how coherence is constructed. She (Schiffrin, 1987:24) points out that her coherence model focuses on local coherence, which is constructed through relations between adjacent units in discourse but can be expanded to more global coherence. She further proposes that discourse markers in her model are viewed as "indicators of the location of utterances within the emerging structures, meanings, and actions of discourse" (Schiffrin, 1987: 24).

The main idea of Schiffrin's Coherence Model ("five planes of talk" in Schiffrin's term) is that discourse markers can function in five different components of coherence, namely exchange structure, action structure, ideational structure, participation framework and information state. Each discourse marker has one primary plane but can function on more than one level, either separately or simultaneously. The following is the definitions of the five planes summarized by Fraser (1999).

> **Exchange Structure**, which reflects the mechanics of the conversational interchange (ethnomethodology) and shows the result of the participant turn-taking and how these alternations are related to each other;
> **Action Structure,** which reflects the sequence of speech acts which occur within the discourse;

> **Ideational Structure**, which reflects certain relationships between the ideas (propositions) found within the discourse, including cohesive relations, topic relations, and functional relations;
> **Participation Framework,** which reflects the ways in which the speakers and hearers can relate to one another as well as orientation toward utterances; and
> **Information State,** which reflects the ongoing organization and management of knowledge and metaknowledge as it evolves over the course of the discourse.
>
> (Fraser, 1999: 934)

As mentioned in the earlier review, Schiffrin's model has influenced a number of later studies of pragmatic markers. Not surprising, her model has received a lot of criticism (e.g. Jucker, 1988; Redeker, 1991). The main criticism seems to be about whether it is necessary to propose five different planes. For example, Jucker (1988) criticizes the five planes model by pointing out that the "information state" would actually include all the other levels. Redeker (1991) proposes a three-component coherence model instead, namely ideational structure, rhetorical structure, and sequential structure, which are roughly equivalent to ideational, action and an extended variant of exchange structures in Schiffrin's model. Another problem with this model is its applicability. Jucker (1988:221) points out *well* is classified as a marker which primarily functions at the plane of participation framework, which reflects speaker and hearer relationship and orientation towards utterances, but the core function of *well* proposed by Schiffrin does not seem to be influenced by her five planes of talk. Schiffrin (1987: 126) defines *well* as "one device used by speakers in their attempts to build coherence in the face of multiple options: *well* anchors the speaker into a conversation precisely at those points where upcoming coherence is not guaranteed". Jucker (1988: 224) further points out that Schiffrin's model "makes it almost impossible to discern the common core underlying all the seemingly disparate uses of *well*."

The main reason that Schiffrin's model is not directly applied to the analysis of this study is because it only reveals three main functions of *I mean* (Schiffrin, 1987: 300) while Brinton's (2007) framework (for details see section 4.1.1), which is used as a

starting point in this study, presents a more detailed description of *I mean*.

2.2.3.3 Blakemore's (1987, 1992) Relevance Model

Blakemore's Relevance Model is another influential theory in research of pragmatic markers. As mentioned in the earlier review, there seem to be more advocates of this theory than those of Schiffrin's Coherence Model. As just mentioned in the above subsection, Jucker (1988) points out that Schiffrin's Coherence Model has failed to account for the use of *well*. In his later study, he even claims that Blakemore's Relevance Theory is "the only theory that can account for all the uses of *well* on the basis of a general theory of human communication on cognitive principle" (Jucker, 1993: 438).

Blakemore's Relevance Model can be seen as a product of the interaction between the assumptions she has formed on what makes successful communication and Sperber & Wilson's (1986) Relevance Theory. Blakemore (1992:27) claims that Sperber & Wilson's Relevance Theory could be regarded as an attempt to develop the notion of relevance, which is left undefined in Grice's (1975, 1989) account of the Co-operative Principle.

In Blakemore's model, there is a distinction between what the speaker means and what his/her words mean. Every utterance produced on a particular occasion by the speaker has a variety of possible interpretations. Successful communication requires the hearer to recover the message intended by the speaker. The success of narrowing all the possible interpretations down to the intended one is because the hearer uses the Relevance Principle as the criterion to evaluate all the possibilities i.e. the hearer assumes that the first interpretation that satisfies the criterion is the only one. The Relevance Principle claims that every act of communication creates a presumption of its own optimal relevance. Therefore an utterance will reach its optimal relevance if it has enough contextual effects and if it does not cause the hearer unjustifiable

processing efforts.

Since success of communication is achieved by cooperation between the speaker and hearer, what the speaker should aim to do is to make the interpretation of his/her utterance easier. For example, the speaker can help to constrain the interpretation recovered by the hearer by constraining the hearer's choice of context. Pragmatic markers ('discourse connective' is used by Blakemore) used by the speaker in conversation can be a very useful device, signaling to the hearer the most likely way that the speaker wishes his/her utterances to be interpreted.

There are three ways identified by Blakemore in which information conveyed by an utterance can be relevant:

1)It may allow the derivation of a contextual implication

2)It may strengthen an existing assumption (by providing better evidence for it)

3)It may contradict an existing assumption

(Blakemore, 1992:138)

Accordingly, three types of discourse connectives have been identified (Blakemore 1992: 138-142):

1) discourse connectives which introduce contextual implications (*so*, *therefore*)
2) discourse connectives concerned with strengthening (*after all*, *besides*, *moreover, furthermore*, utterance-initial *also*, *indeed*)
3) discourse connectives which introduce denials (*however, still*, *nevertheless*, *but*)

Before moving on to the next subsection, I should point out that in Blakemore's relevance-based account of interpretation of utterances, the two concepts – context and coherence – are defined differently from those in coherence-based accounts. As stressed by Blakemore, context is not confined to "either to the immediate physical environment or to the immediate preceding text or discourse" (Blakemore, 1992:18).

Instead, context is defined as a subset of the hearer's beliefs and assumptions about the world. In coherence-based accounts of utterance interpretation the role of context is restricted to establishing coherence relations (Blakemore, 1987: 110). But it is very common to see discourse consisting of utterances that do not show coherence relations but can still be comprehensible as shown by example (2.5).

(2.5)

A What did Susan say?

B You've dropped your purse.

(Blakemore 1987: 110 quoted as example (6))

The relevance-based account also claims that the right context for interpretation of an utterance will not be given to the hearer in advance of the utterance, but will be constructed in the course of interpreting the utterance by the hearer. Therefore, shared knowledge is viewed as "a result of, rather than a prerequisite for, successful communication" (Blakemore 1992: 21).

Unlike the coherence-based model, which views coherence as the evidence or characteristic which distinguishes discourse from a random collection of sentences, a relevance-based account sees coherence "as a result of the way in which hearers use contextual information in their search for relevance" (Blakemore, 1992:2). Therefore, the same utterance can be perceived as coherent by some hearers but incoherent by others.

Blakemore's Model will not be followed in this study because its advantage does not lie in revealing how pragmatic markers behave in specific contexts. As shown in Andersen's (1998: 148) study of the pragmatic marker *like*, the advantage of the framework of relevance theory is that it can provide a unitary account of *like*.

2.2.3.4 Brown and Levinson's (1987) Politeness Theory

The concept of 'politeness' in Brown and Levinson's (1987) theory is much more abstract than those specific polite rules such as opening doors for others, saying

please or *thank you*. As Bloomer (2005: 112) points out, Holmes's version, where 'politeness' refers to "behavior which actively expresses positive concerns for others, as well as non-imposing distancing" (Holmes, 1995:5) is very close to what is meant by 'politeness' in Brown and Levinson. The core of Brown and Levinson's (1987) theory is the notion of 'face', which refers to the public self-image that every competent adult member of a society wants to have. It consists of two related components:

> (a) negative face: the basic claim to territories, personal preserves, rights to non-distraction – i.e. to freedom of action and freedom from imposition.
> (b) positive face: the positive consistent self-image or 'personality' (crucially including the desire that this self-image be accepted and approved of) claimed by interactants.
>
> (Brown & Levinson, 1987: 61)

To ensure smooth social interactions, participants work together to maintain each other's face. However, in actual communication, there are certain kinds of acts which inherently run contrary to the face wants of interactants. These acts are termed face-threatening acts (FTAs). For instance, the act of giving advice will threaten the hearer's negative face because it can pose pressure on the hearer to either perform or not perform, while the act of disagreeing will damage hearer's positive face because it shows the speaker does not want what the hearer wants. Of course, to avoid face loss, the best strategy for interactants is "don't do the FTA". But unfortunately, FTAs such as request, apology, disagreement and refusal are part of our communicative needs. What interactants can do is to formulate their requests, apologies or refusals in a way that they can protect their face to the maximum degree.

Brown and Levinson identify four types of politeness strategies, namely 1) without redressive action, baldly 2) positive politeness 3) negative politeness and 4) off record. The first politeness strategy is to do the FTA directly without any actions of redressing being made. This strategy is often used in situations where interactants have a very close relationship or the FTA is done in an emergency or in the interests

of the hearer. Positive politeness strategies are used to minimize the detriment to the hearer's positive face. For instance, the speaker can use compliments to make the hearer feel good about him/herself or build up solidarity by claiming common ground. Negative politeness strategies are used to protect the hearer's negative face i.e. his/her freedom from action and imposition. Using hedges is one of the ways that the speaker can decrease the imposition on the hearer. The last strategy is indirect strategy. For instance, instead of directly declining the hearer's invitation, the speaker should indirectly state the reasons why s/he cannot accept the invitation. Brown and Levinson also predict that the more a FTA threatens the speaker's or hearer's face, the more the speaker will want to choose a higher-numbered strategy.

Brown & Levinson's Politeness Theory has been related to research of pragmatic marker because pragmatic markers are part of the linguistic efforts made by the speaker to attend to the hearer's face. Brown and Levinson observe that *you know* can be used as a positive politeness strategy. By using *you know*, the speaker shows his/her interest in the hearer by his/her effort to "draw the hearer as a participant into the conversation" (Brown & Levinson 1987: 107). *You know* is found to be used to mitigate FTAs such as disagreeing and giving suggestion in the present study. So in Brown and Levinson's term, *you know* is also used to attend to the hearer's negative face.

The Politeness Theory does help to account for the cases where *I mean* and *you know* serve the purpose of maintaining interactants' negative or positive face. But it cannot account for those cases where the two markers play ideational or textual functions. So the application of this model is confined to cases where *I mean* or *you know* is used to preface FTAs in this study.

As will be seen later in this thesis, no pre-existing model has been applied directly to the analysis of the chosen data. The reason for this is that a bottom-up approach

would suit this study better than a top-down approach does, given that the aim of the present study is to give a detailed account of the two pragmatic markers considered. A top-down approach starts with a pre-existing model and the analysis will be carried out under the model. A bottom-up approach, on the other hand, starts with investigation of the contexts where pragmatic markers are used, then assigns various functions to them and finally more often than not sets up its own framework. As Erman (1987: 114) points out the main problem with a top-down approach is that priority will be given to the model and those functions identified in the analyzed data which cannot fit the model will have to be disregarded.

The application of bottom-up analysis seems to be a rather common practice in previous research. For instance, Holmes (1986) investigates what functions are played by *you know* in her data and comes up with her classification which is made over a continuum of certainty and uncertainty. Müller (2005) gives a very detailed description of a number of pragmatic markers without using a previous framework. The pragmatic functions identified in her data are classified into two major categories: textual level and interpersonal level. In the present study, the identification of pragmatic function of *I mean* and *you know* starts with investigating its own data by drawing upon previous findings on these two markers and then all identified functions are classified over a continuum of speaker-orientation and hearer-orientation to meet the need of carrying out a detailed comparison between *I mean* and *you know*.

Since there are various underlying motivations as to why the speaker uses pragmatic markers, this study will seek any plausible explanations rather than focus on one single explanation. As Aijmer (2002:11) points out, focusing on the Relevance Theory, which accounts for the speaker's cognitive motivation, means politeness i.e. the social motivation may be neglected.

2.2.4 Methods used in establishing pragmatic functions of pragmatic markers in the literature and in the present study

In this subsection, four main issues of establishing pragmatic meanings of pragmatic markers will be reviewed. Firstly, what other aspects of description of markers should be included in addition to their functions? Secondly, how many functions can a pragmatic marker have? Thirdly, how to define functions of pragmatic markers? Finally, why is paraphrase or translation a useful technique in pinning down the meaning of pragmatic markers?

Although it seems that there have not been explicit discussions as to what should be included in the description of a pragmatic marker in the literature, previous research shows that in addition to pragmatic functions, the core of the description, other often described aspects include prosodic features and positioning. Svartvik's (1980:169) description of *well* is made from seven perspectives: "position in discourse, collocation, prosody, pauses, incomplete surrounding structure, functions and Swedish equivalents". Erman (1986:131)'s analysis of *you know*, *you see* and *I mean* is also made from seven aspects:

1) position in utterance/turn
2) syntactic environment
3) semantic relationship between clauses connected by them
4) position and use in information structure
5) phonological properties
6) listener reaction to them
7) interaction with pause

The commonality between Svartvik (1980) and Erman (1986) is that both descriptions include functions, prosodic features and positioning. For the present study, description of functions is no doubt the main concern. Other aspects include positioning and co-occurring linguistic elements. This study does not include

prosodic features because the data chosen in this study do not carry prosodic information. It is true that prosodic information such as intonation plays a very important role in interpreting what pragmatic markers mean. Certainly more research needs to be done in this aspect as Bolinger (1989:300) comments in his study of intonation features of *well* that "aside from an occasional acknowledge that intonation plays a role, few details have been offered." However, according to Svartvik (1980: 172) there is no distinct correlation found between prosody and meaning. Bolinger (1989:338) does agree with Svartvik's criticism of "Crystal & Davy's attempt to assign meanings to *well* that belong rather to the intonation".

This study includes the description of positioning of *I mean* and *you know* because flexibility of position is a key feature which differentiates their status as pragmatic markers vs. their use as fully lexical verbs. Although "there is no one-to-one correspondence between position and function" (Erman, 1987:206), a certain function might have its preferred positions or individual markers might have different preferred positions. This study also includes the description of co-occurring elements of *I mean* and *you know* because one feature of pragmatic markers is that they tend to cluster. Co-occurring expressions sometimes can even be linguistic clues to the interpretation of the functions of pragmatic markers (Aijmer, 2002: 30).

As discussed earlier, to pin down what pragmatic markers mean is a difficult job because they tend to occur in various contexts. There are three approaches as to how many functions a pragmatic marker can have. They are the maximalist, minimalist and polysemy approaches. Maximalist is also termed 'homonymy' approach because it allows sometimes a very large number of different senses assigned to the marker while minimalist is also called 'monosemy' approach because it aims to "isolate a unitary core meaning, usually of a highly abstract and schematic nature, from which all uses of a given item can be derived" (Hansen, 1998: 86).

In the earlier studies (Murray, 1979; Svartvik, 1980; Őstman, 1981, 1982, 1995; Carson, 1984; Schourup, 1985; Schiffrin, 1985; Fraser, 1990) of pragmatic markers a minimalist approach is rather prevalent because one main concern for them is that the pragmatic meaning of a marker should be able to explain why marker A rather than marker B is used in a given context. Murray (1979:728) proposes to give "a unified account of *well*, narrow enough to distinguish it from *oh*, broad enough to capture the diversity of its use". Caron (1984:26) attempts to identify the core use of *well* which can be applied to all instances of *well* and can provide a plausible explanation as to why *well* appears in those particular contexts. Őstman (1995:97) claims that "research on pragmatic particles should aim to answer question why X was used here, and not Y". Fraser (1990:395) claims that "a core meaning must be found and evaluated for the other markers as well".

The minimalist approach is certainly very useful for this study because one of its aims is to explain why certain functions can only be played by *I mean* but not *you know* or vice versa. However, core meaning tends to be very abstract or general, and therefore it is almost impossible to see how it can help to achieve another key goal of this study i.e. to compare *I mean* and *you know* in terms of specific functions they play. Therefore, like Hansen (1998) and Aijmer (2002), in this study I have decided to take the middle way approach i.e. the polysemous approach in which pragmatic markers are viewed as having different functions, which are related to the core meaning in a polysemous way.

As to the establishment of functions of pragmatic markers, one has to look into the contexts where a pragmatic marker occurs, because as Erman (1987:114) points out the context and function should be considered as one entity, so if we state one we have to state the other. Both Erman (1987) and Schiffrin (1985) establish the functions of pragmatic markers by the coherence relations between two adjacent utterances or propositions linked by the pragmatic markers in question. Erman

(1987:114) said she had to postulate the existence of a covert semantic link between adjacent propositions so as to have a meaningful discourse analysis and two main types of coherence relations were identified – causal and concessive. Schiffrin's (1985:643) analysis of *well* is carried out by looking paired utterances such as question and answer; request and compliance sequences. Schiffrin (1987:24) claims that her coherence model focuses on local coherence, which is constructed through relations between adjacent units in discourse.

However, the two utterances connected by pragmatic markers do not always display coherence relations. For instance, pragmatic markers can preface utterances functioning in digressions from a topic or topic changes (Lenk, 1998:19). Schiffrin (1985:643) also reports that there are occurrences where *well* cannot be explained by a pairwise view of coherence in her study.

However, Redeker's (1990) taxonomy of coherence relations includes adjacent units which do not show the coherence relations mentioned by Schiffrin (1985) and Lenk (1998) above. They are termed sequential relations. The table below is the summary of Redeker's (1990:369) taxonomy of coherence relations.

Table 2.2 Taxonomy of coherence relations (Redeker 1990)

Coherence relations	Ideational relations	temporal sequence, elaboration, cause, reason, consequence		
	Pragmatic relations	Rhetorical relations	antithesis, concession, evidence, justification, conclusion	
		Sequential relations	Paratactic	transitions to next topic or next point
			Hypotactic	into or out of a commentary, correction, paraphrase, aside, digression, or interruption segment

Although Redeker's (1990) classification of coherence relations can help to identify more functions than Erman (1987) and Schiffrin (1985, 1987), it fails to account for cases where pragmatic markers are used to preface FTAs, thus functioning in

mitigating FTAs. In the present study, the labeling of pragmatic functions will give priority to coherence relations. For instance, both *I mean* and *you know* are found to preface speaker's attitude in this study, but they will be coded as 'indicating speaker attitude' only under the condition that the utterance which expresses the speaker attitude and its adjacent utterance do not display any coherence relations. In other words, the establishment of functions of pragmatic markers will consider the coherence relations between the adjacent utterances linked by the markers in question before sequential relations and other interpretations such as the function of softening FTAs.

The last issue discussed in this subsection is about the technique of pinning down pragmatic meanings of pragmatic markers – paraphrase or translation – a very useful tool which the present study has inherited from previous research of pragmatic markers. Although it has been widely used by researchers (Wierzbicka, 1976, 2003; Svartvik, 1980; Carlson, 1984; Bolinger, 1989; Aijmer & Simon-Vandenbergen, 2004; Brinton, 2007), Wierzbicka (1976:328) seems to be the one who explicitly argues that paraphrase is the only possible way to accurately decompose the meanings embodied in pragmatic markers because pragmatic markers play the role of "illocutionary forces' which are "bundles of assumptions, intentions and other more or less elementary "postures" and "turns" of the mind". She further proposes that since the meaning of a sentence is compressed into a marker the right way of stating its meaning should be to reconstruct the sentence (Wierzbicka, 1976: 328). The following two sentences are the paraphrases of the pragmatic marker *well* offered by Wierzbicka (1976: 360):

1) I don't want more time to pass like this.
2) I say: this is well; something else has to be said.

According to Wierzbicka (1976: 360), sentence 1) could account for *well* in examples 2.3 and 2.4 while sentence 2) could be the paraphrase of *well* in example 2.5.

2.3

– ***Well***, did Harry capture the aardvark? (The speaker is waiting for the addressee to say

something)

2.4

– ***Well***, to make a long story short… (The speaker is reluctantly permitting.)

2.5

– You'll find yourself in the Fourth Square in no time. ***Well***, that square belongs to Tweedledum and Tweedledee…

However, researchers who have used paraphrase all seem to point out the benefits it has brought to their research. Svartvik (1980) used paraphrase to establish the functions of *well*. His paraphrases of *well* were done in both English and Swedish. According to Svartvik (1980:172) paraphrasing in both English and Swedish not only helps identify the meanings of *well* but also divide his data into functional categories. He further points out that the Swedish translation "can be seen as serving a dual purpose: to provide contrastive statements as well as to further highlight the meanings of *well*" (Svartvik, 1980:172).

Carlson (1984:3) claims that his intuitions of *well* are sharpened by its English paraphrase *oh* and one of its Finnish counterparts. Aijmer & Simon-Vandenbergen (2004) propose that translations can be used as a heuristic device for setting up lexical fields. To establish meaning relations between markers of expectation, they look at Swedish and Dutch equivalents of expectation markers – *actually, in fact* and *really* – in translation corpora (including English into Swedish and Dutch and their translations back into English). The advantage of the translation method is "it provides rich details about both the source and the target items involved" (Aijmer & Simon-Vandenbergen 2004: 1797).

Translation technique is also used in Brinton's (2007) analysis of *I mean*, which helps start the analysis of *I mean* in this study. Unlike most of the studies where the pragmatic markers considered are either paraphrased in a sentence or other pragmatic markers whose meanings may be equally complicated, *I mean* in Brinton's (2007) analysis is translated into rather transparent markers such as *because*, *in other words*,

namely and *for example*. Undoubtedly, this has made the application of her findings to the data of this study much more feasible. As will be seen later in this thesis, this translation technique has also been applied in the analysis of *you know*.

To sum up, the positions that I have taken in terms of the above four issues of how to establish pragmatic functions of pragmatic markers are: firstly the present study will include co-occurring linguistic expressions and position of pragmatic markers; secondly the study will take the polysemous approach i.e. the identified functions of a marker relate to its core meaning in a polysemous way; thirdly it will give priority to the labeling of pragmatic meaning of a marker by the coherence relations between two adjacent utterances.

Finally following Brinton's (2007) study, this study will use the translation technique in its analysis. For example, it will translate *I mean* or *you know* into markers which tend to have a transparent form and function correlation to establish the broad categories. As to the identification of the subcategories, other criteria like the lexical-grammatical features that may consistently co-occur with the specific sub-types of *I mean* and *you know* will also be considered. For instance, one of the main functions of *I mean* is to repair. Under this category, *I mean* can be roughly translated into ' what I meant to say'. One subcategory of repair identified in this study is assumption correction (for details see section 4.1.2.1.1). Identification of this subcategory is strengthened by the fact that *I mean* tends to co-occur with negation.

2.3 Previous studies on L2 learners' use of pragmatic markers

As mentioned earlier, pragmatic markers in the English language have received the most attention among all languages. Similarly, most of the previous studies of pragmatic markers in SLA are about L2 English produced by learners with various L1 backgrounds.

Romero Trillo (2002) investigates 6 pragmatic markers: *look, listen, you know, I mean, well* and *you see* in his Spanish EFL learners' data. Both his native speaker and non-native speaker data consist of two data sets – children and adults. The comparison between children and adults in the native speaker data shows that there is an increase in the use of pragmatic markers in the adults. Although the English-speaking children and the Spanish EFL children show a similar pattern in their use of the markers, the adult non-native speakers fail to match their native speaker counterparts because "the appearance of these markers in non-native adults is even more limited than in the native children corpora" (Romero Trillo, 2002: 779).

Fung & Carter (2007) look at what discourse markers Chinese EFL learners use in the context of classroom as compared to British speakers. On the basis of the main functions of discourse markers, the discourse markers occurring in their data are classified into 4 categories: interpersonal category, referential category, structural category and cognitive category. Their findings show that the Chinese EFL learners are found to display a liberal use of referentially functional discourse markers (*and, but, because, ok, so*, etc) but a relatively restricted use of other markers (*yeah, really, say, sort of, I see, you see, well, right, actually, cos, you know)* while native speakers are found to use discourse markers in a wider variety of contexts.

Hays (1992) records what discourse markers are used by the Japanese EFL learners. The frequencies of the used markers shows that the most frequent ones are *and, but* and *so*, which carry more referential meaning. He argues that the preferences shown by the Japanese learners could be that markers with more referential meaning might be easier to be acquired.

Nikula (1996) investigates how advanced Finnish EFL learners use pragmatic markers in their L2 conversation. She notes that compared to the native speakers the learners showed higher frequencies in those markers (e.g. *I think, I suppose, I don't*

know, maybe, really) which have a close translational equivalent in Finnish. By comparing the speakers' L2 English and L1 Finnish, she argues that such preference is a sign of positive transfer i.e. the learners' use of these markers is facilitated by the learners' L1.

Demirci & Kleiner (1997) also suggest that L1 could account for the characteristics of the Turkish learners' use of pragmatic markers. Their findings show that Turkish learners used more markers which have equivalents in Turkish. For example, they used more *but, because,* and *so* but less *well, oh* and *anyway*. They also note that *but* played two main functions – contrastive and resumptive – in the native speaker data while only the function of contrastive was used in the non-native speaker data. They argue that the absence of the resumptive use of *but* in the non-native speaker data is because the Turkish equivalent of *but* does not function in resumption in L1 Turkish.

Given that most studies on describing characteristics of learners' use of pragmatic markers in the literature are done by comparing native speaker and non-native speaker performance at a very general level, Müller (2005) decides to make her comparison between German EFL learners and native speakers of English at the level of individual functions because she thinks that " a sound comparison between native speaker and non-native speaker discourse markers use has to be carried out at the level of individual functions and not just of the discourse markers in general" (Müller 2005: 242).

Müller (2005) looks at four pragmatic markers *well, you know, like* and *so* in German EFL learners' conversation as compared to how they are used by native speaker speakers of English. In addition to statistical results, the bulk of her study is devoted to a detailed description of what specific functions those markers play in her learners' data. As a result, the similarities and difference between native and non-native speakers are not only shown through frequencies and distributions but also individual

functions of each pragmatic marker. Undoubtedly, such a description improves our understanding as to how pragmatic markers are used by non-native speakers. However, Müller does point out that she has not been able to identify reasons for the difference frequencies shown between the two groups. Although she notes that there are some functions only occurring in the native speaker data, she does not attempt to account for their absence in the non-native speaker data. For instance, her explanation for the absence of the quotative *you know* (i.e. when *you know* is used to introduce a quotation) is simply that the Germans seem do not know this subfunction.

As can be seen from the above review, previous research on L2 learners' use of pragmatic markers seems to come up with a similar pattern regardless of their L1 backgrounds i.e. L2 learners tend to use pragmatic markers less frequently than do native speakers and they prefer to more transparent markers. But to better understand how pragmatic markers are used by L2 learners, we cannot ignore the variations among the individual markers in terms of their frequencies and more importantly in terms of their specific functions. Obviously more in-depth corpus-based studies like Müller's (2005) study are needed in this branch of research.

In this chapter I first presented a broad picture of research of pragmatic markers by looking at the historical development of this branch of linguistic study and how pragmatic markers have been approached by researchers with a wide range of research interests. Then, I reviewed the main theoretical and methodological issues in this research field and explained why certain positions in those relevant theoretical and methodological issues have been taken in this study so as to achieve the best possible description of *I mean* and *you know* in L1 English. Finally, I reviewed studies on L2 learners' use of pragmatic markers and explained why the present study will follow Müller's (2005) study by comparing *I mean* and *you know* in native speaker data and non-native speaker data on the level of specific functions they play.

Chapter Three

Methodology

This chapter consists of 5 sections. Section 3.1 will give a detailed description of the analyzing materials used by this study. Section 3.2 will show how the native speaker data and non-native speaker data are transcribed. Section 3.3 will discuss which cases of *I mean* and *you know* are not considered in this study. Section 3.4 will show how reliability of data coding is achieved in this study. Finally, section 3.5 will clarify some terms used in describing the positioning of *I mean* and *you know*.

3.1 Data

This study is based on the analysis of spoken texts from published corpora of both native speakers and EFL learners.

As baseline data, I selected the face-to-face conversations of private dialogue in ICE-GB (International Corpus of English-Great Britain) "where nearly one-third of the samples consist of spontaneous dialogues. This is the most common type of speech that English speakers engage in" (Nelson et al., 2006). These conversations were all recorded non-surreptitiously between 1990 and 1993 mainly either at home or at work. This subcorpus consists of 185,208 words and involved 338 speakers whose age ranges from 18 to 92. Its main discourse types are asking and providing information, describing one's experience either of one's work or personal life, and giving opinions.

The Chinese EFL learners' data come from two subcorpora: SECCL 1.0 (Spoken English Corpus of Chinese Learners) from SWECCL 1.0 (Spoken and Written English Corpus of Chinese Learners) and SECCL 2.0 from SWECCL 2.0 respectively. Both SWECCL1.0 and SWECCL2.0 are major projects accomplished by teams of Chinese researchers. SWECCL1.0, led by Nanjing University, was completed by cooperation among eight universities and research institutes**,** while

SWECCL2.0 had as many as thirty-four universities as its participants.

Both SECCL1.0 and SECCL 2.0 are recordings of spoken tests of a national English test for English major students in China, which is entitled TEM4 (Test for English Majors, Band 4). Examinees are second year university students, who are about 20 years old. One of the main differences between these two spoken corpora is that SECCL1.0 collected data from 1996 to 2002 while SECCL2.0 covered tests from 2003 to 2006.

The oral test consists of three tasks, but only data from task 3 is used in my analysis because task 3 is dialogues while both task 1 and task 2 are monologues. In task 3, examinees are divided into pairs. Each of them is given a cue card telling them their role in the conversation and also the situations in which their conversation takes place. Then they are given three minutes to prepare and four minutes to talk. In order to ensure that all examinees have something to talk about, the topics of task 3 have to be relevant to university students' daily life or study. Therefore, what they are mainly asked to do is to give opinions, discuss or argue with their friends or school fellows on certain issues such as education and jobs. Since the ranking of examinees' performance is provided in both SECLL 1.0 and 2.0, I decided to compare a group of best-performers of Chinese EFL learners, who I would define as intermediate learners, with native speakers. Because only one topic is given for task 3 each year, I included conversations from 1996 to 2006 so as to maximize the variety of topics in the EFL learners' data. In order to ensure that the size of Chinese EFL learners' data is comparable to that of British speakers, Wordsmith 5.0 was used to calculate the total number of words of selected EFL conversations, because the number of words each file contains is not available. Table 3.1 provides a summary overview of the data.

Table 3.1 Summary of the two sub-corpora chosen by this study

British speakers	Chinese EFL learners
ICE-GB	SECLL1.0&2.0
Private direct conversations between 1990 and 1993	Role-played conversations of intermediate learners between 1996 and 2006
Dyadic or group conversations recorded either at home or work	Dyadic conversations recorded during a national English test
Main discourse types are describing experiences, asking for and providing information and giving opinions.	Main discourse types are seeking advice, discussing or arguing with friends or school fellows on given topics.
185,024 words	185,480 words

Since the Chinese EFL learners' data are role plays, they are admittedly not the same as the British speakers' data which are authentic. But the role-played dialogues in the learners' data can actually generate interactions which resemble natural conversations because they are "spontaneous and open role-plays" (Kasper 2000: 323) where participants do not take on social roles different from their own and the course and outcome of the conversation are not predetermined. However, the native speaker data do cover a wider range of topics, discourse types, and age groups than do the learners' data. Therefore, the variety of topics, discourse types and age may be intervening variables in this study.

3.2 Transcription of data

The transcriptions of both data are orthographic rather than prosodic, but there are some slight differences between them. For example, in the native speaker data, only pauses are encoded and a binary system of long and short pauses is used (<,,> and <,> stand for long pause and short pause respectively) while in the non-native speaker data punctuation marks of comma, full stop and question mark are used and pauses are also encoded but no distinction between long and short pauses is made. Another difference is that the conversations in the native speaker data are broken into

lines and numbered while those in the learners' data are not. As a result, when examples from the native speaker data are referred to, the number of the conversation and the number of the line are used while when examples from non-native speaker data are cited, only the file name of the conversation is mentioned.

3.3 Cases of *I mean* and *you know* that are not considered in this study

Since this study focuses on pragmatic use of *I mean* and *you know*, those cases where they are not considered as pragmatic markers are excluded. Like Müller (2005), being syntactically optional is regarded as the essential feature among the properties of pragmatic markers in this study. Therefore, for *I mean*, instances where it takes as complement either a clause as in example a) or a nominal group as in example b), or a to-infinitive as in example c) are excluded.

a) *I mean* that nothing is totally safe.

b) I don't mean the teacher. *I mean* the student.

c) *I mean* to hold a feast for my guests.

(Longman Online Dictionary)

For *you know*, the following examples are excluded in this study: those where it is part of an interrogative form in examples d) and e), takes a complement, either a clause as in example f) or a nominal group as in example g), or is used in a more or less fixed phraseology as in examples h) and i).

d) Do *you know*…?

e) How do *you know*…?

f) *You know* where Lord Street is.

g) *You know* the kind of thing.

h) *You know* what I mean.

i) As *you know*…

Also, in this study those instances where the pragmatic functions cannot be established due to the insufficient context are excluded because the identification of the pragmatic functions of *I mean* and *you know* is based on the interpretation of the context in which they occur: as Brinton summarises “For many scholars, the central function of pragmatic markers is to express the relation or relevance of an utterance to the proceeding utterance or to the context” (Brinton 1996:30). They are put under the unidentified category.

Fortunately, only a small number of unidentified cases are caused by unclear words in the recording or the absence of intonation marking in the transcription, while majority of them are like examples (3.1) to (3.4) where the current speaker either abandons his/her message and restarts another message or has not completed his/her message before the next speaker takes his/her turn.

(3.1)

175B Perhaps he was twenty-three and been divorced twenty-three three times ***I mean*** <,> <laughter>
176B It seems a <unclear-word>
177B Because he was obviously an innocent you know abroad literally
(S1A-014-ICE-GB)

(3.2)

28D Are we insured <,>
29C At last
30C ***I mean*** I was
31A It seems very expensive
(S1A-008-ICE-GB)

(3.3)

255B I 've I 've read masses of stuff but I mean y you can only draw on things that are appropriate to what you 're doing can't you
256B I mean I could write a whole sequence of stuff based on ***you know***
257B How does he
258B I mean how does he
(SLA-045-ICE-GB)

(3.4)

164B I was wondering <,,>
165B would you like to do some sight singing
166A Mm
167A Yes
168A I'd love to

169B Just ***you know***
170A Would you like me to use your this book or
171B No
(SLA-045-ICE-GB)

3.4 Data-coding

To ensure that the reliability of the analysis was as high as possible, the following measures were taken to strive for consistent interpretation of the pragmatic meanings of *I mean* and *you know*. Firstly, one-third of the British conversations were taken as samples to set up the categories of the pragmatic functions before the full study. All samples were checked by another analyst and disagreements resolved through discussion. The rest of the data, then, were coded twice and there were a few weeks between each coding. Again, controversial cases were resolved through discussion with another analyst. The agreement rate between the two codings was about 80% while 70% is considered to be sufficient to ensure reliable results according to Müller (2005:28).

3.5 The positioning of *I mean* and *you know*

In the analysis of the positioning of *I mean* and *you know*, this study not only investigates their turn position but also their proposition position so as to give a detailed account of their behaviour when they appear in turn medial position. There are 5 terms which need to be clarified when describing their positioning. They are proposition, linker, pre-linker, medial-linker and post-linker.

In this study, a proposition refers to the part of an utterance which plays a part in the transitivity of a clause (Halliday 1994). A general term of linker is used for the sake of convenience to refer to those linguistic expressions which are not included in transitivity analysis, although it is not ideal. They include pragmatic markers such as *you know*, *I mean*, *well*, conjunctions such as *so*, *and*, *because* and interpersonal expressions such as *basically*, *personally* and *I think*.

Example (3.5) is a typical example where *you know* or *I mean* appears in proposition initial position.

(3.5)

326A Did you catch that word
327B And the others <laugh>
328B Actually it 's strange
329B ***you know*** so many people just don't bother to shop anywhere else for those
(SLA-017-ICE-GB)

However, there are cases where the marker is not at the very beginning of a proposition but co-occurs with other linkers. The proposition positions of *I mean* or *you know* in examples (3.6) and (3.7) are all coded as proposition initial, but *I mean* in example (3.6) and (3.7) are termed pre-linker and post-linker respectively and *you know* in example (3.7) is termed medial-linker because of their spatial relationship with their co-occurring linkers. This classification of linkers also applies to turn initial position when the pragmatic mark in question is with other linkers.

(3.6)

141B Why do why do you think he doesn't write then
142B Does he not have the time
143B or do you just think he 's not <unclear-syllables>
144A I don't know
145A ***I mean*** *I think really uhm* <,,> it 's very difficult to to to produce any form of art unless you are driven <,>
(SLA-015-ICE-GB)

(3.7)

20A I wanted
21B About you wanted to keep the fea uh
22A Oh yeah
23A *And I said well* ***you know I mean*** a lot of people wouldn't be necessarily be interested in uh <,,> element theory but if they if they thought that it c it could be handled in feature geometry <,> then they would probably read the element bit as well and they might ultimately become interested in it <,>
(SLA-005-ICE-GB)

All ranking clauses are treated as independent propositions. For instance, both main clauses and subordinate clauses are considered as separate propositions rather than as parts of propositions. In example (3.8), the proposition position of *I mean*, which

occurs in the middle of a clause complex but at the beginning of a main clause, is coded as proposition initial. *I mean* in example (3.9) is also in proposition initial position because it appears at the beginning of a subordinate clause, which is introduced by ‘because’.

(3.8)

113B Because uh I think <,> some able-bodied people <,> may think we can be too fragile to touch <,> to push over <,> to tip out of our chairs

114B but <,> if we say what we're what we're capable of and people are careful ***I mean***<,> I don't think it's that risky

115B I think it's exciting <,>

(3.9)

209B I I thought it was probably better to actually write the thing first and send it to them than than to just contact them <,,> ***I mean*** because they always want proof of one 's work

(SLA-066-ICE-GB)

Most proposition positions in the ICE-GB data can be easily decided by the transcription i.e. whether *I mean* or *you know* is transcribed in the beginning, middle or end of a line. However, there are instances where the proposition position cannot be identified automatically. Example (3.10) is a typical example where *you know* is transcribed on a separate line, which means it can either be seen as the end of the previous line or the beginning of the next line.

(3.10)

158B The first thing she said was feline

159B She was the most feline person I'd seen all year <,>

160B ***you know***

161B And people do people do tend to describe themselves

162B and also <,> if people have a very limited vocabulary which a lot of people do <,> it's interesting looking at them and seeing <,> in what ways it 's limited <,,>

(SLA-037-ICE-GB)

You know is coded as proposition final here because it is more likely that *you know* relates back to line 159 because of the conjunction ‘and’ in line 161.

Example (3.11) is a case where *I mean* appears at the end of a line, but is coded as proposition initial given the typical use of *I mean*.

(3.11)

96B Mhm
97B Yeah
98B I I really don't know ***I mean***
99B The way that I <,> would approach th those sort of things would be just to do them as an interest as a hobby
100B and uhm if they if you got somewhere then that's great
101B but n not to sort of pin too much on it
(SLA-033-ICE-GB)

My analysis suggests that *I mean* is most often used to connect two separate propositions usually appearing in two successive lines, with *I mean* starting the second proposition and linking back. There is a plausible link between line 98 and line 99 in example (3.11). Therefore, it is less likely that *I mean* here relates back only to line 98. The more plausible assumption of *I mean* being transcribed in the end of line 98 would be that B might pause slightly after 'I mean' to think about exactly how to express the next point, but intend to carry on (it is unfortunate that the technical information about the ICE-GB corpus gives no indication of how decisions were made as to how lines are broken). On the other hand, if it is *you know* that appears at the end of a line, it would typically relate back only to the line it is in and not connect to the next.

As to proposition medial, a further distinction is made as to whether the pragmatic marker in question occurs inside a constituent such as NP, VP and PP or between constituents. In example (3.12), *you know* is inside the noun phrase 'a feature theory', so its position is termed 'within a constituent' while in example (3.13), the position of *you know* is termed 'between constituents' because it appears between the verb *be* and its complement.

(3.12)

13A What Bruce Hayes was saying
14A Well I mean he he is a bit weird to have a to combine <,> particl par particle <,> theory with a ***you know*** feature theory
(SLA-005-ICE-GB)

(3.13)

48A There were two hour practicals on Wednesdays which took me ten minutes because I wasn't bothered with the accuracy of my <,> results

49A The idea was ***you know*** we were supposed to do all these graphs and stuff (SLA-008-ICE-GB)

In this chapter I firstly gave a detailed account of the analyzing materials used in this study. Although the native speaker data chosen by this study are bigger and more representative than those used in previous studies, I am aware of the differences between the native speaker data and the non-speaker data in this study. For instance, as will be seen in the following discussion, when accounting for the features of the Chinese learners' use of *I mean* and *you know*, discourse type is considered as one of the factors that affect the learners' production of these two markers. Then, I explained how pragmatic and non-pragmatic cases of *I mean* and *you know* are distinguished in this study. I also mentioned the procedures that this study has followed so as to improve the reliability of data-coding. Finally, some terms used in describing the positioning of *I mean* and *you know* have been clarified with illustrations.

Chapter Four

Functions of *I mean* and *you know* in the native speaker data

This chapter consists of 3 main sections. To begin with, section 4.1 will show how pragmatic functions of *I mean* are identified in the British speakers' data with illustrations and how my own framework is formed. Section 4.2 will, then, demonstrate how categories of *you know* in the native speaker data are set up by following the same approach used in analyzing *I mean* and the identified categories will be classified under the same broad framework as *I mean*. By doing so, another objective of this study, a detailed and systematic comparison of *I mean* and *you know*, can be achieved. Thus in the final section such a comparison will be presented.

To ensure the consistency of the data-coding of this study, the following three guiding principles will be followed. Firstly, labelling priority will be given to local contexts in the identification of the pragmatic functions of the two markers. For example, wider contexts (which can cover a few turn exchanges) will be considered only when no obvious functions can be identified via the two messages connected by the two markers. Typical examples are functions such as summarization (for details see sections 4.1.2.1.8 and 4.2.2.1.10) and resumption (for details see section 4.1.2.3.4).

Secondly, as to those functions that can be identified through investigating local contexts, labelling priority will be given to cases where coherence relations can be identified by the two linked messages. As mentioned in Chapter 2 (p 43), the pragmatic function of the two markers will be coded as ' indicating speaker attitude' only when the message that expresses the speaker's attitude and its adjacent message do not show any coherence relations.

Finally, with regard to those cases where the markers in question play more than one function, labelling priority will be given to the more important function (for details see Chapter 4, p 86) or the more typical function (for details see Chapter 4,

p131-132).

As to the terms used to label the pragmatic functions of the two markers, one point needs to be clarified here. Like Erman (1987:114) the identification of the functions examines in what context the marker in question occurs. However, for the sake of convenience, instead of labelling a function as 'occurring with certain context'', all the functions are named after the contexts where *I mean* or *you know* occurs. For instance, when the two markers are considered as playing the role of justification, it actually means that they co-occur with the context of justification i.e. the justification is realized through the two messages linked by the markers rather than the markers themselves.

4.1 Functions of *I mean* in the native speaker data

As a very frequent phrase in English conversation, *I mean* has received a lot of attention. Although various functions of *I mean* have been proposed, there seems to be a consensus among researchers (Crystal & Davy, 1975:97; Swan, 1992:377) that it is extremely difficult to pin down what exactly *I mean* means.

Brinton's framework (2007) has been chosen as the starting point for this study because her corpus-based study gives a more detailed account of *I mean* than other previous studies, most of which are rather brief, ranging from a short paragraph (Swan,1994) to a few pages (Crystal & Davy, 1975; Schourup, 1985). Although there are a few book-length studies (Goldberg, 1980; Schiffrin, 1987), they are based on very limited data and only focus on the general functions of *I mean*. Schiffrin (1987:300) claims that the three broad functions that *I mean* plays in her analysis are modifier of the speaker's ideas, modifier of the speaker's intentions and replacement repair given the semantic meanings of *I mean*. However, to answer the question of what exactly *I mean* means, a further investigation of the specific contexts in which these general functions are realized is essential because *I mean* "is tied to its moment of utterance and indicates but does not itself specify the nature of the nonequivalence

the speaker finds to exist between what is said and what is meant" (Schourup 1985:148). More specific functions such as explicitness, exemplification and cause identified in Brinton's analysis certainly help to reveal that the speaker's ideas are actually expanded by giving more details, providing further examples and giving explanations.

However, since her study focuses on the historical development of *I mean*, her examples are from quoted speech of written English corpora to "provide a better comparison with the (necessarily) written corpora of Middle and Early Modern English" (Brinton, 2007:41). Therefore, to allow new functions to emerge from the spoken data of this study and avoid the problem of forcing the data to suit a pre-existing framework, her framework was treated as the starting point and was deliberately left open to not only new functions but also reclassification of the functions.

This section will be organized as follows. In section 4.1.1, Brinton's framework will be introduced. Section 4.1.2 will give a detailed account of functions of *I mean* identified in this study and my own classification.

4.1.1 Brinton's framework (2007) of *I mean*

As can be seen from Table 4.1, the four main categories that Brinton identifies in her findings of the pragmatic meanings *I mean* are appositional meanings, cause, speaker attitude and interpersonal meaning. Following Schiffrin's classification (Schiffrin, 1987:300), Brinton also believes that the pragmatic functions of *I mean* can fall into two broader categories, metalinguistic (modifying the speaker's ideas) and metacommunicative (modifying the speakers' intentions). The category of appositional meanings is metalinguistic because all its subcategories "focus on code, on the particular expressions used" (Brinton, 2007:45) while the other three categories are metacommunicative because these usages of *I mean* are intersubjective (Brinton, 2007: 50, 53).

Table 4.1 Summary of pragmatic functions of *I mean* in Brinton's (2007) study

Functions	Examples
Appositional meanings	
Repair (=what I meant to say)	(4.1) "I'll see you in the morning." She laughed. "***I mean***, afternoon." (4.2) "How many…***I mean***, how long is it since you got the first of these?"
Reformulation(=in other words)	(4.3) I just want to look at the stuff. ***I mean***, examine it physically, not experience it emotionally.
explicitness(=namely, that is)	(4.4) It could be embarrassing, you see. Politically, ***I mean***.
Exemplification(=for instance)	(4.5) Miranda was a star; I was space dust. ***I mean***, when she made cheerleader our sophomore year, I got elected treasure of the Latin Club.[sic]
Cause (=because or I'm saying it because)	(4.6) Don't you think it's time you put that thing away. ***I mean***, look at it, it's antique; you could hurt yourself with it.
Speaker attitude	
Express emphasis or assert the veracity of an utterance	(4.7) But Cousin Alexander is rich. Really rich, ***I mean***.
Express evaluation or judgment	(4.8) ***I mean***, it's humiliating to be beaten by someone who doesn't even walk properly.
Express sincerity(=I'm serious when I say)	(4.9) I would never pick up the phone and call him; ***I mean***, I wouldn't do that.
Interpersonal meaning	(4.10) It is because she isn't that she is successful… ***if you understand what I mean.*** (4.11) "If it was, then conceivably Congressman Metcalf resented that and –Well. ***You see what I mean*** (4.12) Or the paper does rather. ***Know what I mean***?

However, there are some slight differences between them with regard to under which of these two categories certain functions should be subsumed. Example (4.13) is an example cited by Schiffrin where *I mean* is used to preface expansion of the speaker's prior ideas. Therefore, it is metalinguistic.

(4.13)

a But I think um ten years from now,
b it's going to be much more liberal.
c I could see it in my own job
d ***I mean***, when I started working for the government, there were no colored

people.

e And today eh…uh…twenty five, thirty percent, forty percent of the people I work with are—are colored.
(Schiffrin 1987:296)

According to Schiffrin's interpretation, *I mean* here prefaces the speaker's explanation of why he believes racial integration is increasing. However, in Brinton's analysis, *I mean* indicates that "the speaker is being attentive to the hearer's need for explanation" (Brinton, 2007: 50) and can be paraphrased by 'I am saying this because'. Therefore, it should be coded as cause, which is metacommunicative in Brinton's classification.

Brinton's study is concerned with the historical development of *I mean*, so she includes all cases of *I mean*. To show the diachronic change of the usage of *I mean*, she distinguishes between the semantic meaning ("full meaning" in Brinton's (2007: 43) term) and pragmatic meaning of *I mean*. But her category of interpersonal meaning, which is considered pragmatic in her framework, will be excluded in this study because *I mean* in those examples (4.10-4.12) is part of more or less fixed phrases in which it has its full lexical meaning.

In sum, Brinton's framework excluding the category of interpersonal meaning will be used as the starting point for this study. The next section will show how a finer description of *I mean* in this study has been developed on the basis of Brinton's framework.

4.1.2 Pragmatic functions of *I mean* in this study

As shown by Table 4.2, this study came up with a longer list of the pragmatic functions played by *I mean* than does Brinton's analysis. One of the reasons for Brinton's shorter list might be due to her data, which consist of quoted speech from written corpora because it would reduce the chances of including those instances which tend to occur in authentic conversations and those instances which show very low frequencies. Summarization and resumption are two typical examples. They do

not appear on Brinton's list because identification of them would involve a rather long turn or a couple of turn exchanges. Also, both of them show very low frequencies (resumption 2% and summarization 1.4%, respectively).

Table 4.2 Classification of pragmatic functions of *I mean* in this study

Hearer-oriented	**Interactant-relationship-oriented**	**Speaker-oriented**
Assumption-correction	Softener of FTA	Indicating speaker attitude
Exemplification	Interactional repair	Hesitation marker
Explicitness	Justification	Quotative
Reformulation	Conclusion	Restart
Cause		Resumption
Reason		Transactional repair
Result		
Summarization		

It should be noted that the majority of the newly identified functions are the result of further categorization of Brinton's categories. Take the subfunction of repair in her category appositional meanings as an example. There are no subcategories of repair identified in her analysis, but it seems that repair has further been categorized into vocabulary repair and grammatical repair in previous studies. Crystal & Davy (1975:97) posit that the speaker uses *I mean* to repair either syntactically awkward utterance or a wrongly chosen word. Goldberg notes that *I mean* "tends to repair a prior phonological or grammatical error" (1980:244) when it is in utterance medial position. In the present study four subtypes of repair are identified. Brinton's examples (4.1) and (4.2) would both be coded as transactional repair because in both cases *I mean* marks the correction of wrong information. As opposed to transactional repair, another subtype of repair, interactional repair, is identified in cases where what the speaker is correcting is a socially inappropriate message rather than wrong information (for details see section 4.1.2.2.1). The reason for making a distinction between these two kinds of repairs is because transactional repair is about modification of messages while interactional repair handles interactants' relationship. The third type of repair is called restart because *I mean* is used to reorganize an incomplete utterance (for details see section 4.1.2.3.2) while the fourth subtype is

coded as assumption-correction because what the speaker is trying to repair is the gap between the hearer's potential assumptions on the basis of what has just been said and what the speaker has intended to say (for details see section 4.1.2.1.1).

Another typical example is Brinton's category of cause under which her two interpretations of *I mean* 'I am saying it because' and 'this happens because' are subsumed. The rationale for coding these two interpretations separately in this study is that, when *I mean* is read as 'This is my reason for saying x', the upcoming message prefaced by *I mean* serves to justify the previous message which is, in most cases, face-threatening so as to maintain the relationship between the speaker and hearer, while, when *I mean* is read as 'This is the reason why x happens', *I mean* is used to indicate how the upcoming message semantically relates to the prior one. As a result, the former interpretation was recoded as justification while the latter one was recoded either as cause or reason in this study (for details see sections 4.1.2.2.2, 4.1.2.1.5 & 4.1.2.1.6). Interestingly, *I mean* also occurs in the context where it can be paraphrased as 'so' in this study. But it shows rather low frequency, which could be the reason for its absence on Brinton's list. Again, the two subcategories of result and conclusion (for details see sections 4.1.2.1.7 & 4.1.2.2.3) are set up for the same reason as the distinction being made between cause, reason and justification. The category of result is set up as opposed to categories of cause and reason because the message prefaced *I mean* is the consequence of the previous message while the category of conclusion is set up as opposed to the category of justification because the upcoming message is the speaker's inference made on the basis of the previous message.

Not only do Brinton's categories provide a very good start for the analysis of this study, her technique of identifying the pragmatic functions played by *I mean* has also been applied because of its practicality and consistency, which supports her claim that:

> But often the semantic-pragmatic interpretation of *I mean* in any given context

> rested on translation equivalents, such as 'namely' or 'in other words'. This offered the most reliable and transparent method for teasing out pragmatic meaning in these contexts and provided the greatest possibility for replicability. (Brinton 2007: 41)

As can be seen from Table 4.3 below, which shows the translation equivalents of *I mean*, *I mean* in the majority categories can have a rough translation.

Table 4.3 Translation equivalents of *I mean*

Hearer-oriented		**Interactant-relationship-oriented**		**Speaker-oriented**	
	Translation equivalents		Translation equivalents		Translation equivalents
Assumption-correction	what I meant to say	Softener of FTA	well (as a pragmatic marker)	Indicating speaker attitude	I'm serious when I say
Exemplification	for instance	Interactional repair	what I meant to say	Hesitation marker	
Explicitness	namely, that is	Justification	I'm saying it because	Restart	
Reformulation	in other words	Conclusion	So	Resumption	as I'm saying this
Cause	because			Transactional repair	what I meant to say
Reason	because				
Result	so				
Summarization	so				
Quotative					

However, a translation equivalent of a category does not necessarily mean that it can equally apply to all cases of that category. Both examples (4.14) and (4.15) can be categorised as explicitness. In example (4.14), line 71 specifies what exactly the speaker refers to by 'big cracks in his face' and *I mean* here can be roughly translated as 'namely or that is' while in example (4.15) by 'her home' Speaker B means 'her own home', but it seems that the same equivalent does not fit in line 278 as well as it does in line 71 of example (4.14).

(4.14)

68A Yeah he had some sort of disease <.>undiag

69A undiagnosed I I believe

70A So he had these great big big cracks in his face <,,>

71A ***I mean*** extremely deep wrinkles <,>

(S1A-008-ICE-GB)

(4.15)

277B They sent one to my mother after she died or something

278B Well<,> it was sent to her home and then it was sent ***I mean*** to her own home uh and she'd moved from there of course and was in the nursing home and they sent it there

(S1A-007-ICE-GB)

It could be argued that in (4.15) the speaker breaks off after 'it was sent' (having intended to say 'then it was sent to the nursing home') and restarts 'I mean to her own home' (referring back to 'to her home'). In that case, 'that is' would actually sound ok. But without intonation it's not possible to be sure.

Also, different categories can share the same translation. For example, transactional repair and interactional repair can both be replaced by 'what I meant to say'. To further distinguish them, a closer look at the specific contexts is needed so as to see what is being repaired, the information or an inappropriate message.

In sum, starting from Brinton's framework, this study has developed a much finer description of *I mean* by looking closer at the specific contexts in which *I mean* is used. Among the functions of *I mean* identified in this study, exemplification, explicitness and reformulation are the only functions from Brinton's framework which are directly applied to the data analysis of this study. The category of 'speaker attitude' is kept but renamed as 'indicating speaker attitude' because it seems to better capture the function of *I mean* in this category where the speaker's attitude is expressed by the message either prefaced or followed by *I mean* rather than by *I mean* itself (for details see section 4.1.2.3.5). Although cause is still used, it only applies to cases where *I mean* is used to indicate the semantic relationship between the linked messages. The remaining functions are either newly identified or the products of further categorizations of Brinton's categories of repair and cause.

In addition to newly identified functions, this study also attempts to reclassify the pragmatic functions of *I mean* from a different perspective. Instead of viewing them in terms of what type of modification (modification of the speaker's expressions vs. modification of the speaker's intentions) is made, the classification of this study was made on the basis of for whose benefit the modification is made. As can be seen from Table 4.2, there are three categories: hearer-oriented, interactant-relationship-oriented and speaker-oriented. The first category is termed hearer-oriented because the overall function of *I mean* in this category shows the speaker's effort to help the hearer interpret the messages more efficiently by marking the adjustment of the prior talk taking into account the hearer's likely interpretations and processing difficulties or the logical connection between messages. The second category is called interactant-relationship-oriented because *I mean* marks the face-work done by the speaker to maintain a smooth relationship between the interactants. Since a socially inappropriate message or a face-threatening act can actually threaten both the speaker and the hearer's faces, some face work such as repairing or justifying needs to be done to save both sides' face so that the relationship between them can be maintained. The reason that the final category is labeled speaker-oriented is that the linguistic efforts marked by *I mean* are made for the purpose of accommodating the speaker's needs. For example, the speaker uses *I mean* to correct his/her own mistakes, or indicate his/her attitude.

In the following sections, a detailed account of the pragmatic functions played by *I mean* in this study will be presented.

4.1.2.1 Hearer-oriented

4.1.2.1.1 Assumption-correction

This is one of the subtypes of repair identified in this study. It is termed assumption-correction because what the speaker is trying to repair is the hearer's potential assumptions derived from what he/she has just said.

(4.16)

58B But he's a perfectionist
60B ***I mean*** there 's a
61B I 'm not saying you 're not
(S1A-008-ICE-GB)

The upcoming message in example (4.16) serves to guide the hearer to interpret the previous message in the way that the speaker intends. Since line 58 could be interpreted as 'what you are saying is he is a perfectionist while I am not', by telling the hearer explicitly in line 61 that such an assumption should be ruled out, the speaker can limit the hearer's interpretations so as to save the hearer's processing time and avoid unnecessary misunderstanding. *I mean* here can be interpreted as 'I guess this could be one of your interpretations of my message. But what I want to tell you is this is definitely not what I mean'. Not surprisingly, negation frequently appears in the upcoming utterance although not necessarily.

(4.17)

47C But they make sure they sort of they seem to make sure that the processes take long enough for you to miss your train
48C So you have to wait for hours <,,>
49C A sort of manip
50C Well this was in Timisoara actually where we had to wait
51A In where
52C Timisoara where we had to
53B <unclear-words> to dilettante
54C And sort of riots started there
55C But ***I mean*** my only recollection of it is sleeping in a wood for about four or five hours
56C Rather idyllic
(S1A-014-ICE-GB)

Extract (4.17) is an example of assumption-correction without negation. C is telling his/her travelling experiences in Timisoara. Hearing that riots started in Timisoara in line 54, the listener could be expecting a story of a horrible experience. But what C can remember is actually something idyllic. Although negation does not appear, with the help of 'but' in line 55, *I mean* signals to the listener that the upcoming message is contrary to his/her potential expectation.

4.1.2.1.2 Exemplification

This function is called exemplification because the upcoming message either serves as the backup of the previous message or as further example(s) when the previous message is about something rather broad. In the latter case, *I mean* works with expressions such as *something like*, *for example*, and *things like*.

(4.18)

189F I'd be far more upset if somebody say scratched one of my records <,> than tore one of my books
190F I mean I seem I don't read books for pleasure, at all
191F ***I mean*** I think nineteen eighty-two was last time I read a book
(S1A-013-ICE-GB)

In line 190, F claims that he/she does not enjoy reading at all and continues telling the hearer in line 191 that the last time he/she read a book was about ten years ago (the conversation was recorded in 1991). Such a detail is designed to give the listener a clearer idea of how much the speaker dislikes reading books.

(4.19)

78B D' you mean having the space there and uh
79A Yes
80A Absolutely
81A Uhm or whatever they provide
82A Uhm <,> well they ***I mean*** they they do <,> things like aerobics and <,> basketball <,> uhm and they have a a whe a wheelchair <,> team of basketball players uhm
83A But th I don't know <,>
(S1A-003-ICE-GB)

In line 82, 'aerobics and basketball' are some details of 'whatever they provide' (line 81) and *things like* helps signal that A is giving examples.

4.1.2.1.3 Explicitness

(4.20)

107B So we agreed that there was independence competitiveness <,> uhm <,> there was a third one I said before
108A Well the whole the whole drivenness of working hard
109B I know
110B All right drivenness
111B Uhm <,,> secondly as I mentioned before more difficulty than I suspect I would have had in dealing with men of my father 's age just not knowing how to relate to them <,>
112B Uhm <,> thirdly uhm <,,> mm let me think <,,>

113B Are we talking about my views now ***I mean*** now at my age or my views then <,>
114A OK
115A Well both
(SLA-075-ICE-GB)

In example (4.20), speaker B uses the message prefaced by *I mean* to specify whether A wants him/her to talk about his/her view 'now' or 'then'.

4.1.2.1.4 Reformulation

In this function, *I mean* can be paraphrased as 'in other words'. Although this function is identified in Schiffrin (1987: 302), she considers it as a replacement repair. In this study, this function is coded as reformulation as it is in Brinton's analysis.

(4.21)

18A Will that be full-time
19B No <,,> but it pays what I would call a part-time wage
20A Uhm <,> and can you do your own hours
21A ***I mean*** you can be there whenever you want
(S1A-011-ICE-GB)

In line 21, the speaker actually repeats what he/she says in line 20 but puts in a slightly different wording. It seems that the repeated information in line 21 is redundant, but the repetition should make it easier for the hearer to process the message.

4.1.2.1.5 Cause

(4.22)

97A Didn't like going to school
98B Uhm <,,> uh what I didn't like uh <,> was leaving my mum really
99B I think that was it
100B I was I was insecure
101B Uhm <,,> I can remember that ***I mean*** it was it was particularly bad at primary school
102B I can I can remember I can remember the first day <,>
103B I can remember vividly my first day at primary school
104B I cried a lot
(S1A-076-ICE-GB)

In example (4.22), the fact that the speaker can still remember that he/she felt very insecure in school is caused by his/her particularly bad experience of primary school.

I mean here can be paraphrased as ‘because’, indicating a cause-effect link between the messages linked by *I mean.*

4.1.2.1.6 Reason

(4.23)

135B Well like I sold a cheapo ten pound Walkman and I 'm ashamed of this
136B It's the only black marketing I did for sixty quid sixty roubles really you know
137B But ***I mean*** it's because I needed to go on a trip for a week in Leningrad
(S1A-014-ICE-GB)

Although *I mean* in (4.23) can also be paraphrased as ‘because’ because the upcoming message in line 137 explains why B sold his walkman cheaply in the black market, it should be distinguished from the *I mean* in example (4.22). In example (4.23), it was the speaker’s decision to do black marketing for the reason that he/she needed the money for a trip to Leningrad, while in (4.22) there is no speaker’s volition involved.

4.1.2.1.7 Result

As opposed to cause and reason where *I mean* can be interpreted as ‘because’, *I mean* in this category can be roughly translatable as ‘so’. Therefore, this function is called result.

(4.24)

52A Uhm <,,> and you say here writing B B C <,,> museums <,> and so on
53B Yeah <,>
54A Right <,,>
55A Do you have the background information on all or any of these
56B I mean<,,> museums I have uh some kind of background in just because the work that I was doing was quite closely affiliated with museums and stuff
57B Uhm <,> ***I mean*** I 'm applying this week for a job at the V and A
58B and I've already been told there's a strong internal candidate <,,> in the India department

In example (4.24), for speaker B, the fact that he/she has some background in museum leads to his/her action of applying for a job at the V and A (Vitoria and Albert Museum).

4.1.2.1.8 Summarization

(4.25)

1A <unclear-words> So it was OK in fact <unclear-word>
2A And it went all right
3A so I've decided that uhm I'm as far as that bit of the paper goes <,> I'm just going to modify a little bit the bit that goes on about element theory <,> rather than not element theory sorry feature geometry rather than develop it there <,>
4A Because me and John said
5A well he said they they they could definitely handle it <,> but it because they'd get round it by underspecifying this that and the other and having tricks where the underspecified bits got put in <,>
6A but there's so many different forms of it that they 'd easily manage it <,>
7A So basically ***I mean*** I said to him <,,> well the rea I want to keep element th uh feature geometry in <,> I said because not <,,> because <,,> you know I'm pandering to people who support it

(S1A-005-ICE-GB)

To identify the function that *I mean* plays in (4.25), we need to look at the relationship between the upcoming message line 7 and all the previous messages from line 1 to 6. A is recounting what has happened to his/her paper, a very complicated theoretical issue which would obviously need a longer turn. After 6 lines of talk, A summarizes what he/she has said so far for the listener because this summarization should make it easier for the listener to understand and remember the main points of the story so that the conversation can move forward smoothly. It is clear that another two pragmatic markers *so* and *basically* before *I mean* contribute to the identification of this function.

4.1.2.1.9 Quotative

This is another category which is very rarely played by *I mean*. It only occurs twice in this study and in both cases *I mean* co-occurs with another pragmatic marker *you know,* which frequently plays the role of indicating the coming message is a quotation (for details see section 4.2.2.1.11).

(4.26)

18A What was I saying
19A God I've lost my thread <,,>
20A I wanted
21B About you wanted to keep the <.>fea uh

22A Oh yeah

23A And I said *well you know* ***I mean*** a lot of people wouldn't be necessarily be interested in uh <,,> element theory but if they if they thought that it <.>c it could be handled in feature geometry <,> then they would probably read the element bit as well and they might ultimately become interested in it <,>

(S1A-005-ICE-GB)

Here *I mean*, with *well* and *you know*, signals to the hearer that the coming message is a direct quote of what the speaker said.

4.1.2.2 Interactant-relationship-oriented

4.1.2.2.1 Interactional repair

(4.27)

94B and people now do things like art therapy and dance therapy <,> uh which is great

95B It's very good work but I think <,> that art and dance actually include <,> the idea of therapy within them <,>

96A When you say recovering the whole person it suggests that there is something lost

97A ***I mean*** you know is there something incomplete

98A ***I mean*** <,> uh what's w what's what's incomplete <unclear-words> <,>

(S1A-004-ICE-GB)

The informational content in line 96 is not wrong, what A is repairing in line 97 and line 98 is the social inappropriateness that the message conveys. Line 96 sounds as if A is questioning the validity of what B has just said about dance therapy, but it seems that A is not in a position to do so because in this conversation B is a professional dancer. As can be seen from example (4.27), A has actually utilized two *I mean*s to complete his/her social repair from a very face-threatening act stating A's own belief about something that concerns B (line 96) to a less face-threatening question inviting B to state his/her opinion (a yes-no question in line 97) and then to an inquiry (a wh-question in line 98).

There are a few cases in this category where the interactional repair is done through change of speaker key. For instance, in example (4.28) speaker B is joking in line 120 about the value of 'this work' (here referring to speaker B's work of teaching a dancing group which involves disabled people). It does not seem very appropriate to

talk about such a serious topic in a joking way, so B switches to a serious speaker key in line 121.

(4.28)

117A Is is there anything else that <,> is there some is there anything else that's that you found particularly valuable about <,> this work <,,>

118B I think yeah I think uh <,,> I think it 's it 's a way of giving <,,> as well

119B I mean<,> I think <,> we're all used to taking <,,>

120B I mean it it sounds <laugh> <,> sounds a bit holy doesn't it really <laugh>

121B No but ***I mean*** it's it's very good to to share <,> things with people <,> and to share what you know<,> and to <,> to learn what <,> other people can give to you and what you can give to them <,>

(S1A-003-ICE-GB)

Cases similar to example (4.28) have also been observed in Schiffrin's analysis (1987). But they are put under the category of modifier of speaker's intentions. Example (4.29) is cited by Schiffrin to show that "*I mean* is also used to reestablish the tone of a conversation by establishing a serious speaker key" (Schiffrin 1987: 298). Although there is a change from joking tone to a serious tone in both example (4.28) and example (4.29), the difference between them is that in example (4.28) it is speaker B who changes his/her own tone while in example (4.29) what Irene is trying to do is to lead the conversation to a serious tone because it is not Irene but Henry that is joking in the prior talk.

(4.29)

Irene a. Henry, if they wanna do it, they'll go away and do it!

b. What'm I gonna do, take a gun and KILL'em?

Henry c. I hate you! Hhhhhhhhhhhhhhhhhhhh

Debby d. I didn't mean t'start thishhhhh

Irene e. No, we argue about this all the time.=

Henry f. hhhhhhhhhhhhhhhhhhhhhhhh

Irene g. =All the time. Always. Because I really don't-

Henry h. If I wanna get my adrenalin worked up we- she Comes in hhhhhhhhhhh

Debby i. Keeps y'young!

Irene j. ***I mean*** I've seen girlfriends of mines parents, sit shiva For them, [a mourning ritual]

k. because they were marrying out of their religion.

(Schiffrin, 1987: 298)

3.2.2.2 Justification

This function is termed justification because *I mean* can be interpreted as 'I am saying this because' i.e. the upcoming utterance explains why the previous message has been said. The speaker's view plus further explanation of why the speaker has come up with such a view would more likely to win the hearer's agreement. More importantly, when the explanation is prefaced *I mean*, it could convey a message to the listener that 'You do not have to agree with me. Feel free to disagree with me if you want to'. Thus, *I mean* decreases the potential imposition on the listener when functioning in justification. The three subtypes of justification identified in the native speaker data of this study are epistemic justification, justification of evaluation and justification of speech act (these terms are taken from Thompson, forthcoming).

(4.30)

284A You'd be able to afford it up there wouldn't you
285C Yeah <,>
286C I mean we d take out a little mortgage
287F Yeah
288A ***I mean*** you wouldn't necessarily need a house
(S1A-019-ICE-GB)

(4.31)

28A Well that's all right
29A you're in good company
30A ***I mean*** with you with your Oxford degree are now going to be photocopying and Sue with her degree is actually putting things in alphabetical order
31A that should make you both feel a lot better
32B <laugh>
(S1A-011-ICE-GB)

Example (4.30) is coded as epistemic justification because the reason why A thinks C would be able to afford an accommodation up there is that A thinks C wouldn't necessarily need a house while example (4.31) is coded as justification of evaluation because the reason for A's assessment that 'you're in good company' is that both the listener and Sue are degree-holders and are doing similar simple jobs. However, the distinction between epistemic justification and justification of evaluation is not always clear-cut. Example (4.32) is a case of justification which lies between these two subtypes of justification.

(4.32)

141B Why do why do you think he doesn't write then
142B Does he not have the time
143B or do you just think he 's not
144A I don't know
145A ***I mean*** I think really uhm <,,> it 's very difficult to to to produce any form of art unless you are driven <,>
146B Yeah
147A Uhm
148B But does he does he not just like write for himself
149A Internally driven I meant
(S1A-015-ICE-GB)

The coding of example (4.32) depends on how A's reply 'I don't know' (line 144) is interpreted. *I mean* would be considered as a marker of justification of evaluation if 'I don't know' is read as A does not think that there is an easy answer like 'he doesn't have the time' to the question 'Why doesn't he write?'. But 'I don't know' here can also be read as A not being sure about the two choices given by B. In this second reading, *I mean* is coded as epistemic justification.

(4.33)

154A Well do it somewhere else
155A ***I mean*** look
156A there's plenty of other places to put it
157A How about here
(S1A-010-ICE-GB)

Example (4.33) is called justification of speech act because what the speaker is justifying here is a speech act of giving a command or suggestion, an intrinsically fact-threatening act. To save both the speaker's and the hearer's faces, the simplest strategy is to avoid any FTA. The speaker, however, has chosen to carry out this FTA because it can be justified i.e. although it is an FTA, it is done for the good of the hearer.

As can be seen from the later discussion, this function of *I mean* shows the highest frequency, which means that *I mean* is mainly used to justify what the speaker has just said. This seems to be supported by the fact that most of the previous studies include examples of justification and they can fit into the above three subtypes of

justification identified in this study although they are assigned various terms rather than justification. For instance, Schiffrin's example of modification of speaker's ideas, which was quoted as example (4.13) above (repeated here for convenience),can be coded as epistemic justification because the message that the speaker's experience of working with more and more colored people is used to explain why the speaker thinks that the racial integration is increasing.

(4.13)

a But I think um ten years from now,
b it's going to be much more liberal.
C I could see it in my own job
D ***I mean***, when I started working for the government, there were no colored people.
e And today eh…uh…twenty five, thirty percent, forty percent of the people I work with are—are colored.
(Schiffrin 1987:296)

One of Brinton's examples of cause, which was quoted as example (4.6) above (repeated here for convenience), can be coded as justification of speech act. Again, the prior message in example (4.6) is an FTA, telling people what they should do, but is carried out for the sake of the hearer's benefit. The speaker tells the hearer to 'put that thing away' because he/she thinks 'that thing' could potentially hurt the hearer if it is not.

(4.6)

Don't you think it's time you put that thing away. ***I mean***, look at it, it's antique; you could hurt yourself with it.
(Brinton 2007: 50)

Interestingly, there are a few cases of *I mean* in this category where the upcoming message is used to justify what has just been done rather than said by the speaker.

(4.34)

3B How does it compare in size to Martin's hall
4B You've seen Martin's hall haven't you
5C Well it's about the s
6C Yes
7C It's comparable I would think but with a lower ceiling
8C <unclear>
9B Lower

10B Lower than Martin's
11B Gosh
12B I think ***I mean*** his ceiling's not not high
(S1A-073-ICE-GB)

In example (4.34), 'Gosh' itself in line 11 does not qualify as a message but is more like an action indicating that the speaker is surprised. The message prefaced by *I mean* explains why speaker B is surprised.

4.1.2.2.3 Conclusion

(4.35)

181B How did Peggy get on with her guitar
182B I 'm not sure I ever got round to asking her
183A Well I don't know is the short answer because she uh never actually answered the question when I asked her
184B Ah
185B I asked her but she never answered the question when I asked her
186A uhm <,> ***I mean*** I always get the feeling it was a grave disappointment in some way but uh whether that's
(S1A-023-ICE-GB)

As opposed to the justification category, where it is the upcoming message that justifies the previous message, in example (4.35) the previous message 'Peggy never answered the question' represents the reason for the upcoming message 'I always get the feeling it was a grave disappointment in some way'. *I mean* can be paraphrased as 'so'. Therefore, this example is coded as conclusion.

4.1.2.2.4 Softener of FTA

To be more accurate, this category should be better termed disagreeing because *I mean* is almost exclusively used to mitigate one FTA of disagreeing as shown in example (4.33). The reason of choosing a more general term softener of FTA is simply for the convenience of comparing with *you know*, which is observed to mitigate other FTAs such as breaking bad news or giving a command (for details see section 4.2.2.2.3).

(4.36)

326B Mrs Thatcher had elocution lessons didn't she yeah yeah
327B Nothing could make that voice apart from artificial training

328A No ***I mean*** I think <unclear-word> is quite useful cos it can help you use your voice

(S1A-018-ICE-GB)

As can be seen from example (4.36), interpretation of this function has to go beyond the same speaking turn because we usually disagree with other people's opinions. The comments on elocution made by B sound negative while A's message after *I mean* shows his/her disagreement that actually elocution can be good because it can help people use their voice properly. Since disagreeing is an intrinsically face-threatening act, efforts to mitigate it are likely to be made by the speaker. In example (4.36), the strategy that the speaker employs is to agree, albeit minimally, with the hearer first and then to use *I mean* to introduce his/her disagreement. This "agreement followed by disagreement prefaced by *I mean*" pattern applies to most instances of disagreeing in this study.

Example (4.37) is the only example found in the data of this study, where *I mean* prefaces another FTA of giving a command.

(4.37)

191B Uhm so that's quite a good thing but I think she's <,> much too scared to confide in them <,,> you know <unclear-words>

192A So she's playing a part

193B Yes she's playing a part

194A Yes

195B And I don't think that's healthy at all

196A Very tiring

197A could be tiring

198B Yes yes

199B Oh she is she is punishing herself

200A Yes

201B So I don't know

202B I 'm I 'm not quite sure

203A But you see ***I mean*** you mustn't flagellate yourself because her back thing <,,> proved that she was never going to be the easy adolescent was she

204A and it might have been something else

(S1A-054-ICE-GB)

As mentioned earlier, those functions identified in this study but absent in Brinton's analysis tend to show low frequency. But softener of FTA is an exception because it

occurs at a fairly substantial frequency of 6% in this study. However, a similar function is mentioned in the previous studies. For instance, Coates (1998:143) points out that *I mean*, as an epistemic modal form, can be used to hedge assertions so as to protect both speaker and hearer's faces.

4.1.2.3 Speaker-oriented

4.1.2.3.1 Transactional repair

(4.38)

125 B I I look forward to <,> to working each week
126 B ***I mean*** I'd like to do it every day <,,> but you know such is life <,,>
(S1A-003-ICE-GB)

What B is doing in line 126 is correcting the wrong information given in line 125. *I mean* here can be interpreted as 'no, what I meant to say was …'.

4.1.2.3.2 Restart

(4.39)

57A I would because to me <,> it seems
58A ***I mean*** I'd go to that and I'd go to the Palmer one if I was you
(S1A-005-ICE-GB)

To identify this subfunction, the prior message has to be an incomplete one followed by another utterance prefaced by *I mean*. The incomplete utterance could either be the result of the speaker's failure to organize his/her thoughts or the speaker's decision to abort the current message.

4.1.2.3.3 Hesitation marker

In the data of this study, *I mean* is very rarely used as a hesitation marker. When it plays this function, it typically occurs with another pragmatic marker *you know*, which is more often used as a hesitation marker (for details see section 4.2.2.3.1) or co-occurs with another marker of hesitation repetition.

(4.40)

127A And the physical contact out there is is the only<unclear-words>problem of being isolated in your <,> wheelchair
128B Uhm
129B Uh uh in a in a chair

130A Uhm
131B I think so
132B Yes
133B I mean I think uhm space is you know ***I mean*** <,> you know just the obstacles that you have in a room
(S1A-003-ICE-GB)

I mean here is coded as a hesitation marker because *I mean*, coupled with two *you know*s and one pause, actually helps the speaker stall for time to figure out how to define space.

(4.41)
101A And for the future would you want to continue <,,> with this group
102?
103C Yes
104B Well yes certainly
105C Uhm uhm
106B For <,,> however long it
107B Well we'll see what happens really
108B It's all ***I mean*** it's all very exciting really
109B We don't know <,> what will happen but <,> we can only sort of work and see what happens

Example (4.41) is also considered as a case of hesitation marker because it occurs with repetition of 'It's all'. Unlike *I mean* in the category of restart where the speaker does not finish the prior message and restarts another one, the upcoming message prefaced by *I mean* in example (4.41) begins with the repetition of the incomplete previous message.

4.2.3.4 Resumption

(4.42)
160A Now that's why you shouldn't worry about Tamsin and Damian coming together cos from Rebecca's point of view <,> it would be a godsend
161B Oh it would uh it would be a great blessing if they can s they can speak for a half an hour or something
162B it would be
163C Oh good
164A Cos they can make her feel easier because I think she feels she's being rather an intrusion
165C <unclear-words>
166C Well will you tell her that you might save Rebecca from complete despair

because being exposed twice within a month would be rather awful for her
167C ***I mean*** who did you say you s found in Scotland
168B Yeah
169B ***I mean*** it would be nice <„>
170B Just have somebody else to talk to really <,>
171B Have a game of table tennis or something
(S1A-021-ICE-GB)

Resumption means the speaker picks up his/her topic which has been interrupted in the previous talk. Interpretation of this function usually needs a wider context covering a few turn exchanges as shown in example (4.42). Interestingly, in this example both speakers B and C use *I mean* to direct the conversation to their previously interrupted topics. In line 167, speaker C intends to continue a previous topic, which is about somebody B met in Scotland. Instead of carrying on C's topic by telling C who he/she met in Scotland, in line 169 speaker B continues his/her own topic of line 162, which has been interrupted by C in line 163.

4.1.2.3.5 Indicating speaker attitude

In Brinton's framework, *I mean* expresses three types of speaker attitudes as shown by Table 4.1. In this study, most cases fall into her second type of speaker attitude i.e. evaluation or judgment. In line with Brinton's (2007:51) claim that when functioning in this speaker attitude *I mean* tends to occur in the context of an evaluative adjective, nearly half cases of this type speaker attitude in this study are expressed by an evaluative adjective. Examples (4.43) and (4.44) are two typical examples in which *I mean* signals that what follows is the speaker's evaluation or judgment. Example (4.43) contains an evaluative adjective while example (4.44) does not.

(4.43)
105A And I've got masses of vases I could have lent her
106E Yeah we used to buy Mum a vase every year for her birthday
107A Yes
108E So we got uh inundated with them
109E We never use any of them hardly
110E ***I mean*** it's just amazing
(S1A-019-ICE-GB)

In line 109, E tells the listener how he/she feels about the fact that they have got lots of vases but have never used any of them through the use of the adjective 'amazing'.

(4.44)

157B and then hopefully <,> we will get some more pieces together <,> and uhm then see where we can perform <,>
158B Uhm <,> personally I I would like to <,> to do that very much <,>
159B Uhm <,> I always keep saying I wish it had start
160B I wish I'd got involved <,> ten years earlier because you know <,> I 'm getting old
161B <laugh> and uhm <,> you know ***I mean*** <,>
162B There you go

(S1A-003-ICE-GB)

In line 162, B expresses his/her attitude that there is nothing people can do about the natural process of getting old.

Instances of the other two types of speaker attitudes identified by Brinton are rather rare in this study. Example (4.45) is an occurrence of expressing emphasis while example (4.46) is a case of expressing sincerity.

(4.45)

290D This looks delicious
291A Don't think makes that much difference <unclear-words>
292D Very on time <,,>
293B Well <,> one of us was
294C Mm
295D ***I mean*** <,> you were very on time
296D You were <,>
297D Absolutely on time

(S1A-022-ICE-GB)

D's repetition of 'very on time' prefaced by *I mean* and 'absolutely in line 297 contribute to the emphasis being expressed in example (4.45).

(4.46)

136B I shall have the dry first liquid second
137B I I I shall try that
138B Have a go at this <,,> whi which has a kind of colour that I was born to appreciate
139B ***I mean*** I love it

(S1A-056-ICE-GB)

I mean in this example can be paraphrased as 'I'm serious when I say'.

In addition to signaling various speaker attitudes, *I mean* has also been observed to mark the peak of a series of comments made by the speaker in this study. In both examples (4.47) and (4.48), speakers are making positive comments. Speaker A in example (4.47) tells the addressee how good a school is while Speaker B in example (4.48) describes how wonderful the performance of the best singer he/she has ever seen. Both speakers make a series of comments on the same subject with the one prefaced by *I mean* more dramatic than all the other evaluations made in previous lines.

(4.47)

100A Rugby <,,> the girls are just treated like a few honorary girls but they're not integrated
101A At King's Canterbury they are integrated but it isn't too free
102A it's still quite academic
103A Also the other thing is <,,> the school itself is so beautiful <,>
104A ***I mean*** it's such a wonderful ambience to be in and so that's quite nice uhm <,,>

(S1A-054-ICE-GB)

(4.48)

114B The b The best singer is this Olaf Bergh that I've seen
115B There's actual
116B He has total concentration
117B He seems not to move a muscle in a whole evening <,>
118B And it's totally relaxed and concentrated
119B And ***I mean*** it it gets to the whole audience and the audience doesn't do all this coughing and spluttering that they do between songs

(S1A-045-ICE-GB)

Due to the composite meaning of *I mean*, it is natural that it frequently occurs in the context where it signals the speaker's attitudes. Gerhardt and Stinson (1994) investigate this particular context where *I mean* functions to "convey the speaker's attitudes and evaluative stance toward the content of the discourse" (Gerhardt and Stinson 1994:151) by analyzing therapeutic discourse where the patient gives accounts of the self. However, not all cases where *I mean* occurs in the context of expressing speaker's attitude are put under this category in this study because identification of the pragmatic function depends on the relationship between the prior message and upcoming message. In example (4.49), it is clear that upcoming

message is about the speaker's evaluation or stance towards "it" (here referring to someone's marriage). But C's attitude towards the marriage in lines 290 and 291 is used to justify why he/she agree with D who holds a skeptical view toward it. So *I mean* in example (4.49) is coded as justification rather than indicating speaker attitude.

(4.49)

284D Well I hope it works
285D Wish them the best
286B Mm
287D I remain sceptical from my own personal experiences
288C Yeah yeah
289C ***I mean*** <,>
290C Just seems a little strange <,> sort of
291C It s uh pretty quick
(S1A-071-ICE-GB)

Brinton's example of expressing evaluation or judgment, which was quoted as example (4.8) above (repeated here for convenience), may not be coded as indicating speaker attitude in this study because this example does not include the prior message although there is no doubt that the message prefaced by *I mean* is about how the speaker feels about "to be beaten by someone who doesn't even walk properly".

(4.8)

I mean, it's humiliating to be beaten by someone who doesn't even walk properly.

The above is the description of all the categories of the pragmatic functions of *I mean* identified in the native speaker data of this study, which will be used as the baseline for the following comparative study between native and non-native speakers. But before moving to the comparison section, it is important to bear in mind that identification of these categories is not always clear-cut because pragmatic markers may play more than one function in one situation. Both Erman (2001) and Müller (2005) propose that when a pragmatic marker fulfils several roles in one instance, the predominant one should be used to label the function of that marker. One example of exemplification in this study, which was quoted as example (4.19) above (repeated here for convenience), is a case where *I mean* could be argued to play two functions.

It is coded as exemplification because the upcoming message gives details of what they provide, but *I mean* here may also be considered as a hesitation marker because it co-occurs with repetition of 'they' and other markers of hesitation such as *uhm* and pause. The function of exemplification is treated as the more important one as Stubbe and Holmes claim that "At general level, pragmatic devices undoubtedly do provide verbal planning time for speakers, and they often have an important function at another level as hedges expressing epistemic modal meaning"(1995: 63). The other reason for treating this case as exemplification is that *I mean* in example (4.19) can be translated as 'for instance'.

(4.19)

78B D' you mean having the space there and uh
79A Yes
80A Absolutely
81A Uhm or whatever they provide
82A Uhm <,> well they ***I mean*** they they do <,> things like aerobics and <,> basketball <,> uhm and they have a a whe a wheelchair <,> team of basketball players uhm
83A But th I don't know <,>
(S1A-003-ICE-GB)

4.2 Functions of *you know*

Like *I mean*, *you know* is also very frequently used in English conversation. But *you know* enjoys so much more freedom in terms of positioning than does *I mean* that it may seem that it can randomly appear anywhere in conversation, which probably gives an impression that it does not really play any roles in conversation and is only used "when one has nothing to say, or when one cannot, or will not bother to, find the proper words to express something" (Schourup, 1985). A letter quoted in Stubbe and Holmes (1995) shows how an ordinary native speaker perceives the use of *you know*.

> The phrase 'you know' is used with monotonous regularity when a person is being interviewed on TV or radio—to commence a sentence, be interspersed throughout, and even to conclude the same sentence. Let's hope 'you know' will soon die a natural death, although another exasperating expression will probably replace it—to ruin my listening enjoyment (New Zealand Listener, 16-22 April 1994).
>
> (Stubbe and Holmes, 1995: 63)

Although *you know* is considered as a stigmatized linguistic expression by ordinary native speakers, it has attracted a lot of attention from researchers. Compared to the studies on *I mean*, *you know* is more thoroughly investigated in previous research. The majority of the previous studies on *you know* are either journal articles (e.g. Holmes, 1986; Huspek, 1989; Stubbe and Holmes, 1995; Erman, 2001) or book-length studies (e.g. Östman, 1981; Schourup, 1985; Schiffrin, 1987; Müller, 2005). In addition to investigation of general roles that *you know* plays in conversation, researchers also show great interest in finding out how social variables such as gender (Östman, 1981; Holmes, 1986), age (Erman, 2001) and social class (Östman, 1981; Stubbe and Holmes, 1995) affect the use of *you know*. Due to the abundant previous research carried out on *you know*, it is not surprising that various functions and frameworks have been proposed. But as mentioned earlier this study aims to draw as delicate a picture of how *you know* is used in its chosen British data as possible, a direct application of a pre-existing framework will be avoided. Instead, the approach and framework of *I mean* in section 4.1 will be applied and findings of previous studies will be modified to fit in the analysis in the present study.

This section will be organized as follows. Section 4.2.1 will explain how the identification of the pragmatic functions of *you know* in this study is influenced by studies of *you know* in the literature. Section 4.2.2 will present a detailed description of pragmatic functions of *you know* identified in this study.

4.2.1 Functions of *you know* in the literature

The abundant studies on *you know* provide a very good foundation for this study. First this study has adopted an approach from earlier studies (e.g. Östman, 1981; Schourup, 1985; Holmes, 1986) which not only identifies specific functions of *you know* but also proposes a core meaning of *you know* i.e. a general function that can be extracted from all instances of *you know*, because the combination of individual functions and their underlying unity of *you know* would contribute to a "balanced and complete account" (Schourup, 1985: 139) of *you know*. More importantly, the

extracted core meaning of *you know* will provide a clue as to why it is *you know* rather than another pragmatic marker that is used in a certain context (Schourup, 1985: 139), since this study also aims to compare *I mean* and *you know*.

Although several versions of the core meaning of *you know* are proposed, they roughly mean the same. Östman (1981: 17) defines the core meaning of *you know* as "The speaker strives towards getting the addressee to cooperate and/or to accept the propositional content of his utterance as mutual background knowledge". Schourup (1985:102) claims that *you know* "indicates that the speaker expects that there is no communicatively significant discrepancy between what is now in the private world and what is now in the other world, with respect to what is now in the shared world". Schourup points out that his proposal is fairly close to a rough synonym 'you know what I'm talking about', which is suggested by Goldberg (1976: 42). Jucker and Smith (1998:194) claim that the general function of *you know* is to invite "the addressee to recognize both the relevance and the implications of the utterance marked with *you know*". The key message shared by the above proposals is that *you know* indicates the speaker's effort to get the addressee to participate in his/her talk. So it is not surprising that the function of appealing is most frequently mentioned by researchers (e.g. Östman, 1981; Edmonson, 1981; Jucker & Smith, 1998) although it is treated as one of the specific functions rather than as the general function of *you know* in some studies (e.g. Holmes, 1986; Erman 1992, 2001; Müller, 2005). The position that this study takes is to treat appealing as a general function of *you know* because it can be applied to all instances of *you know* and it will be helpful in accounting for not only cases where *you know* and *I mean* are not interchangeable but also cases where both *you know* and *I mean* can be applied.

The other support that this study has obtained from previous studies is the description of specific functions of *you know*. The findings from more recent studies (e.g. Erman, 2001; Müller, 2005) will be used as the main references for this study because they present a more detailed picture than the earlier studies. It seems that one important

reason that earlier studies come up with a rather sketchy picture of *you know* is because they tend to generalize the functions of *you know* on the basis of its intonation and utterance position. In Östman's study (1981), the only two subfunctions identified are 'as you know' and 'don't you know'. As Östman points out, these two subfunctions can be easily predicted from the core meaning of *you know*, the different intonation contours that *you know* can have and the utterance position of *you know* (1981:21). The subfunction of 'as you know' is played by utterance initial *you know*, which tends to have a Declarative contour while the subfunction of 'don't you know' is played by utterance final *you know*, which can have either a Declarative or an Interrogative contour. Östman thinks that the two different contour *you know*s in utterance final position have different readings. The Interrogative utterance final *you know* can be paraphrased as ' are you attending, 'do you agree', or 'do you see what I mean' while the Declarative utterance final *you know* can be read as 'I'm not going to say anything more about this', 'don't ask me anymore', or even 'it's obvious' (1981:23). But he decides to subsume these two interpretations under the same subfunction of 'don't you know'. The reason for this decision is they are not easily distinguishable because "the immediate response from the addressee is optional (though it gets more obligatory the more interrogative contour is)" (Östman, 1981:23).

Following Östman's study, Holmes (1986) presents a more detailed picture of *you know*, which is shown in Table 4.4. Like Östman, there is a distinction made between general function and subfunction of *you know,* but more than one general function is proposed. Holmes also keeps Östman's two subfunctions as her two broad categories. The broad category of expressing certainty is correspondent to Östman's subfunction of 'as you know' where *you know* is used to "express (presumed) certainty" (Östman, 1981: 22). The other broad category of expressing uncertainty is correspondent to Östman's other subfunction 'don't you know' where *you know* "suggests more uncertainty on the part of the speaker" (Östman, 1981:23).

Table 4.4 Summary of Holmes's (1986) findings

<table>
<tr><td>General functions of you know</td><td colspan="3">a. serve an intratextual coherence function, tying participants' turns together in a variety of ways</td></tr>
<tr><td></td><td colspan="3">b. function as verbal fillers</td></tr>
<tr><td></td><td colspan="3">c. allude specifically to the relevant knowledge of the addressee in the context of utterance.</td></tr>
<tr><td colspan="4">Classification of subfunctions of you know</td></tr>
<tr><td>Expressing certainty</td><td>Conjoint knowledge</td><td colspan="2">introducing what the speaker regards as incontestable mutual knowledge and refers to the fact that the speaker knows the addressee already knows the information being asserted in the proposition.</td></tr>
<tr><td></td><td>Emphatic</td><td colspan="2">emphasizing, intensifying, or boosting the strength of the speech act, to stress the speaker's confidence and hence reassuring the addressee concerning the validity of the proposition asserted.</td></tr>
<tr><td></td><td>Attributive</td><td colspan="2">expressing the speaker's certainty concerning the validity of the proposition and also express the speaker's confidence that the addressee knows</td></tr>
<tr><td>Expressing uncertainty</td><td>Appealing</td><td colspan="2">serving as an appeal for reassurance from the addressee</td></tr>
<tr><td></td><td rowspan="3">Linguistic imprecision</td><td rowspan="3">expressing the speaker's uncertainty concerning aspects of the linguistic expression of the proposition.</td><td>signaling lexical imprecision</td></tr>
<tr><td>introducing qualifying information</td></tr>
<tr><td>indicating false start</td></tr>
</table>

As can be seen from Table 4.4, subcategories are set up under each broad category. Unlike the establishment of the broad categories, the identification of those subcategories takes other contextual clues into account apart from intonation and utterance position of *you know*. Under the broad category of expressing certainty, Holmes further differentiates three subcategories by looking at whether the speaker is certain about the fact that the hearer knows the proposition or the speaker is certain about the validity of the proposition.

Under the other broad category of expressing uncertainty, Holmes sets up two subcategories and further categorizes the subcategory of linguistic imprecision into three subfunctions. Interestingly, the context where *you know* plays the subfunction

of ‘indicating false start’ is termed ‘specific verbal environments or ‘parameters’ by Östman (1981: 72) when he investigates how gender influences the usage of *you know*. The following are the five verbal environments:

a) use between obligatory constituents in an utterance
b) use before optional elements, at transition-relevant places in an utterance
c) use between repetitions of a word or a phrase
d) use after false starts
e) use as an attention-getting, or topic-changing device

(Östman, 1981: 72)

As can be seen from Holmes’s examples of ‘indicating false start’ shown by examples (4.50) and (4.51), the function of *you know* in Östman’s verbal environment d) would be coded as ‘indicating false start’ in Holmes’s analysis.

(4.50)

Young man to friends at dinner party
and I've been on this bloody speed reading course which is / ***you know*** so one / one notices
(Holmes, 1986: 11)

(4.51)

Female radio interviewee
but fortunately w- ***you know*** there's been more recent research
(Holmes, 1986: 11)

Obviously, to identify more specific functions of *you know* needs an in-depth investigation of the verbal environments of *you know*. For this study, these verbal environments are particularly important because the prosodic information is not available in the chosen data. In this study, *you know*s in Östman’s environments c), d) and e) would be coded as hesitation marker (for details see section 4.2.3.1), restart (for details see section 4.2.3.2) and ‘introducing a new topic’ (for details see section 4.2..3.4) respectively while the two other verbal environments need to be further specified for the roles of *you know* to be identified.

Although Holmes's model does produce a fairly comprehensive description of *you know*, some modifications need to be made so as to suit the aims of this study. Firstly, the identified subfunctions of *you know* will be classified on a scale of interactant-orientation, i.e. whether a subfunction is speaker-oriented or hearer-oriented instead of on a scale of speaker-certainty. Although both interactant-orientation (Östman, 198l; Edmonson, 1981; Holmes, 1986; Schiffrin, 1987) and speaker-certainty are mentioned in the literature, interactant-orientation is the scale shared by *I mean* and *you know* due to their semantic meanings. Since the subfunctions of *I mean* were classified on the scale of interactant-orientation, to serve the purpose of a comparison between *I mean* and you *know*, the same scale needs to be applied to *you know.*

Moreover, it seems that *you know* does not necessarily express the speaker's certainty when it introduces a message that the speaker is sure that the hearer knows. Example (4.52) is cited by Holmes as an instance of *you know* where the speaker is very certain that the hearer knows the proposition introduced by *you know* because what *you know* introduces in example in (4.52) is "incontestable mutual knowledge"(Holmes, 1986: 8).

(4.52)

Woman to husband introducing a narrative at dinner party
well ***you know*** we went to Sally's that night
(Holmes, 1986: 8)

For this study, to pin down what role *you know* plays here needs to know what has been said before *you know*. If this "incontestable mutual knowledge" introduced by *you know* in (4.52) is used to support the speaker's view expressed by the prior message, it could be coded as justification (for details see section 4.2.2.1). Obviously, it would be easier for the speaker to win over the hearer to his/her side if his/her opinion is justified by shared knowledge or experience. Another function of bringing up something known to both the speaker and hearer in conversation is to claim common ground so that positive politeness can be achieved. But since Holmes does

not include the message prior to *you know*, it would be difficult to pin down the function played by *you know* in (4.52).

Also, the three subfunctions under the subcategory of linguistic imprecision do not appear to indicate that the speaker is uncertain or unconfident about what he/she is going to say. When *you know* is used to signal lexical imprecision or indicate false start it appears to be more about accommodating the need of real time communication where the speaker needs to search for appropriate words or repair unsuccessful utterances. When *you know* is used to introduce qualifying information, it seems to be more about the speaker's efforts to accommodate the hearer's potential need of a "more precision or some clarification of the propositional content of the previous utterance" (Holmes, 1986: 11).

The other subcategory of appealing under the broad category of expressing uncertainty is treated as a general function in this study because as mentioned earlier this core meaning of appealing can be applied to all instances of *you know* and its main function seems to be more about inviting the hearer to contribute to the speaker's talk at least mentally if not verbally than signaling the speaker's uncertainty and lack of confidence.

With regard to the general functions of *you know* proposed by Holmes, 'serve an intratextual coherence function, tying participants' turns together in a variety of ways' and 'function as verbal fillers' can actually be seen as general features shared by pragmatic markers. So they would be too general to provide an answer to the question why it is *you know* rather than *I mean* that is used in certain contexts. Although the general function of 'allude specifically to the relevant knowledge of the addressee in the context of utterance' is more specific than the above two general functions, like Östman (1981) and Jucker & Smith (1998) this study assumes that by using *you know* the speaker appeals to the addressee to see the relevance or implication of the utterance marked by *you know* rather than only alludes to the

relevant knowledge of the addressee.

Compared to Holmes's study (1986), Erman (2001) and Müller (2005) give much more detailed descriptions of *you know* by focusing more on the specific verbal contexts of *you know*. As shown by Table 4.5 and Table 4.6, both studies come up with a rather long list of specific functions of *you know*, which provides a very good start for this study. In addition, most examples given in these two studies include wider contexts than those in Östman (1981) and Holmes (1986), which makes the application of them to the data analysis of this study more feasible.

Table 4.5 Summary of Erman's (2001) findings

Textual monitors	Clause	1) mark/highlight certain elements in the thematic structure
		2) introduce a change of information content, frequently correcting or modifying previous discourse
	Discourse	3) introduce propositions
		4) appear between the speaker's position and backing up of it
		5) mark transition between states and events
		6) mark transitions between direct and reported speech
		7) mark inserts of parenthetic comments containing information that speaker assume the addressee needs to know in order to be able to follow
	Editing	8) hesitation marker
		9) repair
Social monitors		10) turn-taking
		11) turn-yielding
		12) comprehension-securing function
Metalinguistic monitors		13) emphatic function
		14) approximator
		15) appealing

However, neither Erman (2001) nor Müller (2005) treats the function of appealing as the general function or core meaning of *you know*. Erman regards appealing as one of her subfunctions although when she comments on the subfunction of approximator she points out that approximators "can be used with an appealing function, the speaker appealing to shared knowledge of the world, general truth or otherwise

controversial issues" (Erman, 2001: 1348). Also in her earlier study, she claims that when *you know* used in the subfunction of hesitation marker, the speaker uses it "as a staller for time and as an appeal to the listener to have patience" (Erman, 1987: 137). In Müller's study, it seems that appealing is treated as a general function in her subfunctions at textual level because they are labeled with this general function being extracted i.e. they are mainly named after the specific contexts where *you know* occurs . However, her subfunctions at interactional level can actually be seen as the result of blending the appealing function into the specific contexts because they are various kinds of appealing identified in her study. According to Müller, in 'imagine the scene', 'see the implication' and 'reference to the shared knowledge' the appeal is made on behalf of the narrative while in 'appeal for understanding' and 'acknowledge that the speaker is right' the appeal is made on behalf of the speaker (Müller, 2005: 181).

Table 4.6 Summary of Müller's(2005) findings

Textual Level:
1) marking lexical or content search
2) marking false start and repair
3) marking approximation
4) introducing an explanation
5) quotative
Interactional Level
6) imagine the scene
7) see the implication
8) reference to shared knowledge
9) appealing for understanding
10) acknowledge that the speaker is right
Various functions (cases occur less than three times)
a) introduce relevant background information, in a parenthetic comment
b) introduce a new topic/digression
c) introduce general knowledge
Unidentified cases

Before moving to next section, which will present the subfunctions identified in this study, there is a need to have a brief discussion of the two subfunctions turn-taking and turn-yielding on Erman's list (2001) because this turn management role of *you*

know is very frequently mentioned in the literature but there is not a consensus among the researchers as to whether this role should be listed separately or even whether *you know* plays this role or not.

Like Erman (2001), Müller (2005) treats this turn-management role of *you know* as a separate function in her study. But it is recoded as 'imagine the scene' because of her corpus, which consists of movie narratives. According to Müller, when functioning in 'imagine the scene', *you know* can be paraphrased as 'you can imagine the scene, can't you?', 'I'm sure you can imagine the scene', or 'Please imagine the scene!', depending on the intonation contour" (Müller, 2005 :171).

However, Östman (1981) claims that although utterance-initial *you know* can play a turn-taking function while utterance-final *you know* can serve a floor-yielding function, the turn-switching function should not treated as another function of *you know* because it can blend into the two subfunctions 'as you know' and 'don't you know' and they can operate simultaneously within one and the same instance of *you know*.

Schourup (1985) even proposes that *you know* cannot function in turn management. He claims that when Declarative *you know* occurs in turn-final position it should not be considered as a facilitator of turn-taking because it is not *you know* but the utterance coming to its terminating point. In addition, half instances of Interrogative turn final *you know* in his data are actually followed by backchannel responses rather than by a full change of turn. This study will take Schourup's position because firstly it seems very difficult to prove that it is *you know* that leads to the turn-exchange. Secondly, identification of this role involves prosodic information of *you know*, which is not available in this study.

To sum up, in order to achieve the most delicate possible picture of *you know*, the descriptions of specific functions of *you know* by Erman (2001) and Müller (2005)

will be used as the main references but their function lists will be, again, left open to allow new subfunctions emerge from the data from this study because different and bigger samples are used in this study. To present a better comparative picture of *I mean* and *you know*, the function of appealing will be extracted from all instances of *you know* as a general function and the identified subfunctions will be classified over the scale of interactant-orientation. Since prosodic information for *you know* is not available, the identification in this study will mainly rely on the verbal environments of *you know*.

4.2.2 Functions of *you know* in this study

As can be seen from Table 4.7 below, a larger number of specific functions of *you know* were identified in this study than in either Erman or Müller's study. Although there is an overlap between the function list of this study and Erman and Müller's lists, more than half of the subfunctions are only identified in this study. It is not surprising that the overlapping subfunctions tend to be those cases where the verbal environments of *you know* appear to have rather distinctive syntactic features and offer less room for alternative interpretation. For instance, when *you know* plays the role of being a hesitation marker (for details see section 4.2.2.3.1), it usually appears "after function words, within the phrase after a determiner, the speaker obviously doing lexical search, or after a con/disjunct at the beginning of the clause for the sake of planning the overall continuation of it" (Erman, 2001: 1344). When *you know* is coded as approximator (for details see section 4.2.2.3.3), it nearly invariably co-occurs with other approximators such as *sort of* or *whatever*. Quotative *you know* (for details see section 4.2.2.3.5) is usually found to co-occur with quotative verbs such as *say* and *go*.

As to the newly identified subfunctions in this study, some of them are due to the fact that the data were interpreted from different angles. For instance, this study attempted to define the pragmatic functions of *you know* by the potential logical relationship between the messages linked by *you know*. Like *I mean*, *you know* was

also found to link messages which had a cause-effect relationship. Subfunctions of justification, reason and cause (for details see sections 4.2.2.2.1, 4.2.2.1.4 and 4.2.2.1.3) were identified. By applying the paraphrase technique used in the analysis of *I mean*, *you know* in justification can be translated as 'I'm saying this because…' while *you know* in reason and cause can be interpreted as 'X happens because of Y'. Subfunctions of result and conclusion (for details see sections 4.2.2.1.5 and 4.2.2.2.2) were set up because there was also a cause-effect relationship between the messages connected by *you know* but in a reversed order i.e. cause is followed by effect. *You know* in both cases can be paraphrased as 'so'.

Table 4.7 Classification of pragmatic functions of *you know* in this study

Hearer-oriented	**Interactant-relationship-oriented**	**Speaker-oriented**
1) Assumption-correction	12) Justification	20) Hesitation marker
2) Introducing background information	13) Conclusion	21) Restart
3) Cause	14) Softener of FTA	22) Approximator
4) Reason	15) Interactional repair	23) Introducing a new topic
5) Result	16) Indicating marked expressions	24) Indicating speaker attitude
6) Explicitness	17) Indicating the most likely event	
7) Reformulation	18) Indicating the unspoken message to be completed by the hearer	
8) Exemplification	19) Indicating the coming message is meant to be evaluated	
9) Seeking confirmation		
10) Summarization		
11) Quotative		

Like the analysis of *I mean*, this study also attempted to define the pragmatic functions of *you know* by the speech act with which it co-occurs. Softener of FTA (for details see section 4.2.2.2.3) was set up because *you know* was found to mark a number of face-threatening acts such as giving a suggestion, disagreeing and breaking bad news.

Some new subfunctions are the result of further categorization of functions identified in previous studies. One of the most often mentioned function in the literature is that

you know is used to introduce modifying information (Crystal, 1998: 47; Holmes, 1986: 1986:11; Erman, 2001: 1342; Müller, 2005: 164). In this study, this function was further categorized into exemplification, explicitness and reformulation (for details see sections 4.2.2.1.8, 4.2.2.1.6 and 4.2.2.1.7) because they actually specify the ways in which the previous discourse is modified.

A few researchers (e.g. Östman, 1981; Schourup, 1985; Jucker & Smith, 1998) argue that *you know* can play the role of claiming common ground between the speaker and hearer so that intimacy can be achieved. In this study, two specific ways of claiming common ground were identified. In both ways *you know* is used to introduce shared knowledge or experience. One was termed 'indicating the unspoken message to be completed by the hearer' and the other was called 'indicating the most likely event'. Both of them appeal to the shared experience or general knowledge. The difference between them is in the former case the speaker leaves the shared knowledge unfinished and invites the listener to complete it mentally while in the latter case the speaker spells out the shared knowledge.

Holmes (1986) points out that the type of discourse could be a potential factor for the frequency of *you know* because she finds that in her corpus "*you know* seems to occur most frequently in sections of relatively sustained narrative or accounts of the speaker' personal experiences intended to amuse, amaze, or, at least, retain the interest of the addressee" (Holmes, 1986: 15). The new subfunctions 'indicating marked expressions' (for details see section 4.2.2.2.5) and 'indicating the coming message meant to be evaluated' (for details see section 4.2.2.2.8) can actually explain how the speaker makes his/her talk more interesting and engaging to the hearer.

Finally, the reason for the absence of subfunctions such as assumption correction and interactional repair in either Erman or Müller's lists could be because they fail to be sampled in Erman and Müller's studies due to their very low frequencies. There is a need to point out here that this study includes all identified subfunctions regardless of

their frequencies, while Müller (2005) puts those subfunctions which occur less than three times under the category of 'various functions' as shown by Table 4.6. For example, she further categorizes the function of 'introducing an explanation' into exemplification, clarification and amplification. But she decides not to include them on her function list. One of her reasons for this is that the three subtypes are too infrequent to form their own categories (Müller, 2005: 167).

As mentioned earlier, the specific functions of *you know* in this study will be classified over the interactant-orientation scale, which was applied in the analysis of *I mean*, for the sake of a better comparison between *I mean* and *you know*. There are 11 (out of 24) subfunctions of *you know* under the hearer-oriented category. They are classified as hearer-oriented because they are used to mainly serve the hearer's needs by making the previous message easier to be understood. When playing the role of assumption correction, *you know* signals to the hearer that the coming message is the unintended interpretation of the prior message. When *you know* is used in cause, reason and result, it indicates to the hearer the cause-effect relationship between the linked messages. *You know* in subfunctions such as explicitness, reformulation and exemplification is used to introduce qualifying information which clarifies the previous messages. When functioning in 'introducing background information', *you know* introduces information which the speaker assumes the hearer needs to know so as to understand the prior message. *You know* in 'seeking confirmation' is used to ensure that the hearer "has correctly understood specific references in the text" (Erman, 2001: 1346). In summarization, *you know* introduces a summary of the previous talk so that the hearer can have a clearer idea about what has been said. Quotative *you know* signals to the hearer that the coming message is a quotation.

Subfunctions under the interactant-relationship-oriented category share the role of smoothing the relationship between the speaker and hearer. For instance, a socially inappropriate message certainly needs to be repaired so as to maintain the interactant-relationship. What *you know* does in softener of FTA is to minimize the

face damage caused by the FTA that the speaker has to carry out in conversation. By justifying what has been said, the speaker would be more likely to win the hearer over. In addition, face-threatening speech acts such as giving commands and disagreeing need to be justified. 'Indicating marked expressions', and 'indicating the coming message meant to be evaluated' are considered interactant-relationship oriented because their overall function is to make the conversation more interesting and engaging so that the social interaction can be maintained. The reason for 'indicating the unspoken message to be completed by the hearer' and 'indicating the most likely event' being classified as interactant-relationship-oriented is because they can bring interactants closer by claiming common ground between them.

The subfunctions under the speaker-oriented category all mainly serve the speaker' own purposes. For example, *you know* is used to indicate the speaker's own attitude. With *you know*, the speaker manages to edit his/her online communication by restarting or stalling more time. When *you know* is used as an approximator, it indicates that the speaker cannot find the exact word(s) to describe what he/she is talking about at that point and asks the hearer to get on with the inaccurate language. When functioning in 'introducing a new topic', *you know* is used by the speaker to shift to another topic.

In the following subsections, a detailed account of how the specific functions of *you know* were identified in this study will be presented with illustrations.

4.2.2.1 Hearer-oriented

4.2.2.1.1 Assumption-correction

This subfunction was identified as one of the subtypes of repair done by *I mean*. As mentioned in section 3.2.1.1 this subfunction is called assumption-correction because what the speaker is trying to repair is the hearer's potential assumptions derived from what he/she has just said. In example (4.53), speaker B recounts his unpleasant experience of working as a professional photographer. What speaker B is particularly

unhappy about is it is not the quality of the photo that matters. Although the record company could end up using his picture if he was willing to compromise himself, he simply thinks that it is not worth it. In line 164, speaker B says such a situation lasted about a few years in his life, which might well give the hearer the impression that the speaker is rather stubborn. Apparently, this is not the assumption that the speaker wants from the hearer. So, in the next line speaker B makes it very explicit that he is actually a person who is open to change.

(4.53)

143A: Uhm so you shot a job that he did that they didn't like then
144B: They didn't like his pictures yeah but then I did the pictures
145B: but the record it was a record company
146B: and the record company said why're you doing it
147B: Terry O'Neill has done it you know
148A: Mmm oh you mean the pop group didn't like it
149B: No no they they didn't they did Terry O'Neill's session and it was such garbage they got him to reshoot
150B: But like the record company then refused to pay me without having seen the pictures
151B: They said Terry O'Neill's done it you know <,>
152B: As if that that would sort of uh
153A: So have they used your pictures in the end of the day <,>
154B: No they didn't actually <laughter> which is uh
155B: but don't mind that's uh that that's that that's a part
156B: that's part of
157B: I mean uh that that being a professional in that profession is like it's the bottom
158B: It's the absolute
159A: They've used his picture yeah
160C: That was probably the company's decision
161A: It's not worth arse-licking all round
162A: I mean it's the big arse-lickers uhm
163B: It's
164B: I mean I I I realise that like in some respects it's the only way of quite a few years of my life
165B: but *I mean* <,> ***you know*** I 'm open to change
166B: and so that's
167A: Well you will
168A: How old are you
169A: you're not very old anyway

(S1A-52-ICE-GB)

However, this subfunction seems to be only a minor role played by *you know* because example (4.53) is the only instance found in this study. In addition, this role is co-played by *I mean*, which rather frequently functions in assumption-correction on its own. But the extra work that *you know* does here is its appealing function. By using *you know*, B appeals to the hearer to acknowledge that he is actually open-minded.

4.2.2.1.2 Introducing background information

This is a subfunction which is identified by both Erman (2001) and Müller (2005) although Müller does not include it on her list because it occurs less than three times in her data. It is termed 'introducing background information' because *you know* is used to "mark inserts of parenthetic comments containing information that the speaker assumes the addressee needs to know in order to be able to follow" (Erman, 2001: 1344).

(4.54)

254B: It's a very weird psychological film <,,>
255C: Oh
256A: Yeah
257B: It's It's futuristic
258B: It's set in the future
259C: There was a very weird film about some stranger you know <,>
260C: I never saw it though
261B: Was there
262B: Mm
263B: It's very very good <,>
264B: Uh it's very difficult to understand
265B: You have to see it once and then you have to see it a second time separately to suss it out <,,>
266B: But I like I like uhm <,> I like him jumping and <,>
267A: Oh my God
268C: Oh there are the boys getting out as well
269B: No because they take like it's set in the year two thousand where the State controls
270B: It's rather like ***you know*** Nineteen Eighty-Four that the State controls everybody and they have spies everywhere <,>
271B: uhm and the information technology has grown so much <,,> that they 're ruled by information technology <,,>

272B: And he works in the information technology
(S1A-49-ICE-GB)

In example (4.54), speaker B is explaining to A and C what is happening in a baffling film. By inserting the information of what 'the State controlling' is in line 270, B provides A and C with the information they need to know so as to better understand the background that the film is set in. With *you know*, B can trigger the listener's memory of another film 'Nineteen Eighty-four'. Instead of passively waiting for the speaker to provide further information about 'the State controls', the hearer would search for it in his/her memory at the same time.

4.2.2.1.3 Cause

In this subfunction *you know* is used to link messages which have a cause-effect relationship. The prior message is caused by the coming message marked by *you know*.

(4.55)

181A The other thing is uhm <,,> do you confide in her <,>
182A Does she feel excluded because you don't exactly confide in her
183B Well <,,> I don't
184B But then <,> uh I haven't
185B I mean I never have <,>
186B and I 'm I 'm rather scared that you know that would seem rather artificial to her and as an attempt to win her over
187B and of course <,> ***you know*** she's terribly alive to things like that <,,> uhm
188A But you could start slowly
(S1A-031-ICE-GB)

In example (4.6), A and B are talking about B's teenage daughter from B's previous marriage (who is referred as "her" or "she" here). B's feeling of being scared that her daughter would think her confiding in her artificial because her daughter is very sensitive. *You know* here not only does the job of signaling the cause-effect relationship as *I mean* would do in the same context but also appeals to A to understand why she feels scared and perhaps even can win A's sympathy. The semantic function of *you know* in this example could be performed by 'because'.

4.2.2.1.4 Reason

In this subfunction, *you know* is also used to connect messages which have a cause-effect relationship. What distinguishes this subfunction from the above function of cause is there is human volition involved here because the reason that C made the decision to hoover her bedroom was because she wanted to show Rosie a clean room. Like the above function of cause, this subfunction is also played by *I mean*. The extra work done by *you know* here, again, is that the speaker can appeal to the hearer to understand why she had the bedroom hoovered.

(4.56)

262C Well this was hoovered yesterday
263C I hoovered my bedroom cos Rosie was coming
264C I thought I show her a tidy room ***you know*** <,,>
265B Yeah I
(S1A-048-ICE-GB)

4.2.2.1.5 Result

Like the above three subfunctions, this subfunction is also played by *I mean*. Although *you know* in this subfunction also appears in a cause-effect relationship, in contrast to the above two subfunctions of cause and reason the message marked by *you know* is the consequence of the prior message.

(4.57)

132A My brother-in-law's quite sweet
133A he sends one to a birthday or Christmas or something and once a year I need a great big vase ***you know***
134A I I found this sort of thing I got from Habitat a square thing <,>
135A Like a tadpole tank really
(S1A-019-ICE-GB)

Speaker A's brother-in-law sends her a big bunch of flowers (is referred by 'one' in line 133) for birthday or Christmas, so she needs a big vase every year. With *you know*, A appeals to the hearer to see that her need of a big vase every year is the result of the fact that her brother-in-law sends a big bunch of flowers every year.

The following three subfunctions show how the prior message is modified. *I mean* was also found to play these three roles. But with *you know*, the speaker not only clarifies or amplifies the prior message by giving more details, providing examples or rewording but also appeals to the hearer to accept that those details, examples and rephrasing are relevant.

4.2.2.1.6 Explicitness

(4.58)

145B Yeah
146B I mean for example there was one girl who I met at U who I interviewed at UC L
147B there's the most striking example of it who was
148B she was black <,> really beautiful
149B She had her hair kind of <,> down in a kind of I don't know net
150B ***you know*** hardly any hair and close-cropped to her head <,> absolutely perfect features <,> and very thin and elegant

(S1A-037-ICE-GB)

In example (4.58), B is describing a girl's hair style. In line 149, B said the girl's hair was 'down in a kind of I don't know net', which is rather vague. The hearer may need more details here to understand what B has described. The details given in the next line is certainly very helpful for the hearer to understand what exactly the girl's hair looks like. *You know* in this example can be replaced by 'that is'.

4.2.2.1.7 Reformulation

(4.59)

182B: and they just leave Nell a little bit behind so that she's always got something to aspire to uhm in terms of maturity
183A But she must be clever
184B uhm
185A uhm it's just they're more swotty
186B They're no They're slightly more mature <,>
187B they're a bit more into boys
188B and Nell isn't quite sure what all that's about yet
189B but I think she's getting inklings
190B and they sort of zipped her up on her dress I mean really ***you know*** took her out of herself and her rather old-fashioned ideas

(S1A-054-ICE-GB)

In line 190, B describes the impact that the school had on Nell. Since the impact is put in a rather metaphorical way, to rephrase it in a relatively straightforward way will certainly make it easier for the hearer to understand the message. The use of 'sort of' in line 190 indicates that the speaker is not very certain about the metaphorical expression, which could be another reason for the reformulation. *You know* in this case can be paraphrased as 'in other words'.

4.2.2.1.8 Exemplification

(4.60)

36A Well you don't have a problem with that
37A artistic people don't
38A but people that have been in involved in sciences or <,> or a thing like that they just don't have creative minds
39A My sister for instance
40A She she just doesn't have any kind of a creative mind at all
41A And anything that's just not black and white facts ***you know*** scientific facts or mathematic facts she can't get her head round it at all
42B Yes
(S1A-037-ICE-GB)

To give the hearer a better idea of what black and white facts are, speaker A gives some examples of black and white facts. *You know* here can be replaced by 'for example'.

4.2.2.1.9 Seeking confirmation

This subfunction is also identified by Erman (2001), but termed as 'comprehension-securing'. According to Erman, when *you know* functions in this subcategory, it plays the role of "making sure that the listener has correctly understood specific references made in the text, usually to people but also to objects and other phenomena" (Erman, 2001: 1346). The following are the only two instances of this subfunction found in this study. As can be seen from both examples, the speaker repeats what has just been said to make sure that the hearer has correctly understood what 'cockle' or 'Turkish' is.

(4.61)

305B Could be clams yes

306B And uh
307A Uhm unless they were cockles although they don't <,>
308A ***you know*** cockles eh
(S1A-009-ICE-GB)

(4.62)

160B I mean is it complete nonsense or is it just a language <,>
161A Oh it seems just like Turkish or something like that
162B Uhm
163A Like Turkish or something ***you know***
164B It's like Turkish
(S1A-015-ICE-GB)

However, it has been found that some examples of the subcategory of explicitness (for details see section 4.2.1.6 above) could also be seen as seeking confirmation if the prosodic information of *you know* was available. Example (4.63) can be coded as explicitness because the message marked by *you know* in line 61 gives additional information as to how grapefruit tastes like. But if *you know* here has an interrogative contour, it could be argued that the speaker' motive to ensure that the listener knows what exactly the taste grapefruit has would be more obvious.

(4.63)

60A Well I quite like that sort of quininey taste that grapefruit has
61A Slightly bitter taste ***you know***
62A I find it more refreshing than those sweety things
(S1A-009-ICE-GB)

4.2.2.1.10 Summarization

Like *I mean*, *you know* is also used to summarize the previous talk. In example (4.64), speaker B gives a very long and detailed account of the problems with the course she took. In order to give the listener a clearer idea of those problems, in line 86 B points out her main problem i.e. the course did not help her build up confidence that a counselor needed because it did not give the trainees enough practice.

(4.64)

61B Oh well we didn't have enough practice I don't think on our course <,>
62B Uhm <,> but the reality of it was I could do I mean you know I I <.>c uh <,> I think I <,> I could do the counselling that I had to do there <,>
63B Probably
64B I mean I was I was pretty much advanced on the delegates

65A Right
66A Mm
67A Right
68B I mean one of the one of the the the girls actually said
69B well she's experienced and then she sort of said
70B very experienced
71B I mean I 'm not very experienced but she's you know the one of the trainees <,> said said that <,>
72B So I think it sort of came across
73B I I felt <,> yes I could have done it a lot better but but I certainly could do what they were doing
74B I mean I was sort of going further than the stuff that they were they were doing
75B So that was <,> I wasn't I mean you always learn don't you every time you see something but I but <,> uhm and I did did learn some oh bits on language
76B There were some bits in their uhm <,> uh a little bit on the on N L P
77B uhm but I'd read that and he didn't really use that
78B So I knew the <,> the the content of what he did but it wasn't actually used on the course
79B Nobody really did anything very much with that <,>
80B But uhm <.>th you know a few
81B But I I wouldn't have said that there was very much new <,>
82B Uh uh ninety-eight per cent of material <,> that was covered I knew <,> and <,> had <,> used <,> and could use <,>
83B Uhm but I would <,> uh I I 'm not an experienced counsellor that I could sit there and just counsel anybody through anything and so on
84B So if you had a demonstration if something came up you know there'd be things that I'd sort of go <,>
85B I don't know what to do here
86B ***you know*** *so that 's what I 'm saying that* I feel that I would need much more experience to be able to deliver the course and feel confident that I could demonstrate it or handle anything that came up anywhere
87B And I'd I'd feel you I I'd I need more <,> you know to be able to do that
(SlA-060-ICE-GB)

The co-occurring linguistic expressions ‘so’ and ‘that’s what I’m saying that’ help the identification of this subfunction. In addition to summarizing, *you know* here also appeals to the hearer to agree with the summary that the main problem was that the course did not offer enough practice given the details of the problems mentioned in the prior talk.

This subfunction is also reported in He & Lindsey's (1998) study, which investigates *you know* in one-to-one academic counseling meetings. In their data, as much as 45% of *you know*s function in summarization. They think that the task of counseling is the main factor because with *you know* "counselors make explicit to the students that this is an important part of their advice to which they wish the students to orient" (He & Lindsey, 1998: 144).

4.2.2.1.11 Quotative

You know in this subfunction is used to mark quoted speech. Identification of it is rather straightforward because *you know* almost always co-occurs with reporting verbs such as *say, go, think* and *tell*. Although this subcategory is identified by a few researchers (He and Lindsey, 1998; Erman, 2001; Müller, 2005), different criteria are applied in their studies. He and Lindsey (1998: 143) include cases like example (4.65) and (4.66) where *you know* is followed by direct quotes and example (4.67) where *you know* is followed by indirect reported speech. In Erman's analysis (2001), quotative *you know* can either be cases where *you know* is followed by quotes or cases where *you know* appears after quotes like example (4.68) because she thinks that quotative *you know* is used "to mark transitions between direct and reported speech, close in function to quotation marks in written text" (Erman, 2001:1344). But in Müller's study (2005), quotative *you know* only includes examples where *you know* is followed by quoted speech. In this study, Müller's criterion was applied.

(4.65)

232C I mean she <,> she seems to be quite sort of
233C She said oh well ***you know*** if it doesn't work out it doesn't work out sort of thing
(S1A-071-ICE-GB)

(4.66)

115B I mean I don't force myself to doing anything but I sort of think <,> ***you know*** <.>wh what's going on in your mind
116B What's the matter
117B What are you <,>
118A Mm
119A All this to yourself <,>
120B Yes

(S1A-050-ICE-GB)

(4.67)

163A The Sport
164D The Daily Sport
165C So he was telling me ***you know*** what was in it
166C So I said well bring them home
(S1A-027-ICE-GB)

(4.68)

194B It would need to be definitely something which she'd not
195B She'd agree with somebody if <.>sh if they said that's done this to me
196B She'd go oh yes yes I know
197B I know what he's like
198B ***you know***
199B But she wouldn't ever uh actually say anything bad about him <unclear-w
(S1A-076-ICE-GB)

4.2.2.2 Interactant-relationship-oriented

4.2.2.2.1 Justification

As mentioned earlier, when functioning in justification *you know* also occurs in a cause-effect relationship. As opposed to the subfunctions of cause and reason where *you know* can be interpreted as 'X happens because Y', *you know* in justification would be read as 'I am saying this because...'. Like *I mean*, the three subtypes of justification played by *you know* are also epistemic justification, justification of evaluation and justification of speech act.

(4.69)

56A It might I think it might make an awful lot of sense actually because then ***you know*** the sort of the unity <,> between the two operations would be <,> <unclear>
57A And really I mean you 're welcome to <,,> you 're welcome to sort of adopt the things that we've developed
(S1A-029-ICE-GB)

In example (4.69), *you know* is followed by the reason why A thinks it might make a lot of sense. So *you know* here is used in an epistemic justification.

(4.70)

320A Zara's a Marks and Spencer's person aren't you <,>
321B Well it's easy to buy clothes and rubbish like that <,,>

322C They're all right then
323B I only buy my [B indicates a bra] from there
324C Me too me too
325A Sorry
326A Did you catch that word
327B And the others <laugh>
328B Actually it's strange
329B ***you know*** so many people just don't bother to shop anywhere else for those
330C <laugh>
(S1A-017-ICE-GB)

Example (4.70) is an instance of justification of evaluation because the reason for B's evaluation that it is very strange that so many people only shop at Marks and Spencer is that they don't bother to shop anywhere else.

(4.71)

44B Ah there's an ant
45A So what
46B Well catch it
47C Well put him outside
48C Let him go on to <unclear-word>
49C Look
50C he's on the toaster
51C Now put him outside nicely <,> then brush him out
52C He may be somebody else's ant ***you know*** <,,>
53C Well I remember killing an insect in that train in India
(S1A-032-ICE-GB)

You know in (4.71) is coded as justification of speech act because what is justified here is the command given in line 51.

Although both *you know* and *I mean* can function in justification and both have three subtypes of justification, it has to be borne in mind that due to their semantic meanings, with *I mean*, the speaker can convey a message to the hearer that 'This is just what I think. You do not have to agree with me' while with *you know,* the speaker appeals to the hearer to agree with him/her.

Although justification is not on Erman's function list, she observes that *you know* will typically occur between speaker's position and the backing up of it when it is

used in argumentative discourse (Erman, 2001: 1343). In particular, (4.72), which is cited to support her observation, can be treated as justification in this study because the detailed description that follows *you know* explains why the speaker says that they did it in 'a slapstick farce way'.

(4.72)

/…/they did it in a completely, slapstick farce way, ***you know*** the the men who were dressed up supposed to be women had great big balloons and, had rosy red cheeks and wigs and things/…/

(Erman 2001:1343 cited as example (5))

4.2.2.2.2 Conclusion

Like *I mean*, *you know* also plays the role of signaling conclusion. *You know* in (4.73) is coded as conclusion because in contrast to justification, the previous message in line 38 that B bought a big piece of fresh salmon for just about five pounds explains why B thinks it was very good value. *You know* here not only signals the cause-effect relationship between the two message linked by *you know* but also appeals to the hearer to agree with speaker B's evaluation that buying that much salmon for just about five pounds was a real bargain. Again, *you know* here can be replaced by 'so'.

(4.73)

34B I've bought very good smoked salmon there too
35A Mm
36A Much cheaper than anybody else
37A Mm that's what I 'm talking about
38B Oh well I've also bought fresh salmon you know a piece <,> like that for that for about uhm four pounds <,> five pounds maybe <,,>
39B ***you know*** very very good value
40B But uhm <,,> the trouble is that uh you can't actually rely on what he have always

(S1A-010-ICE-GB)

4.2.2.2.3 Softener of FTA

Like *I mean*, *you know* can also mitigate FTAs. As mentioned in section 3.2.2.4, *I mean* is used almost exclusively to mitigate one FTA, disagreeing while *you know* tends to soften more varieties of FTAs. Giving a suggestion and disagreeing, which are shown by (4.74) and (4.75) respectively, are the main FTAs mitigated by *you*

know. There is one case of breaking bad news shown by (4.76).

(4.74)

195A Are you recording
196A Ah yeah you you're recording in C then
197A D' you normally record in C
198B I 've no idea
199B It 's the first time I 've done it <laugh>
200A No
201A It's just if you're doing phonetics analysis ***you know*** you should never record in Dolby <,,>
(S1A-008-ICE-GB)

In example (4.74), speaker A tells the hearer not to record in Dolby if he is doing phonetic analysis. With *you know*, the speaker not only softens the face threatening act of telling people what they should not do but also appeals to the hearer to take his advice.

(4.75)

221B I had sardines once with a fly on it
222B I was just about to tuck into it and I noticed this great fly soaked in tomato sauce
223B so I took it back to the kitchen <,> and they they tried to convince me that this fish had swallowed the fly before it died
224B I said
225B but they live under the sea ***you know***
226B you can't have flies under water
(S1A-055-ICE-GB)

Speaker B is recounting one incident in her school canteen. B found a fly on her sardines but was told by the kitchen staff that it was because the fish had swallowed the fly, with which B could not agree. So in line 225, B argued that sardines live under the sea where no flies can live. With *you know*, B mitigated the disagreement and at the same time appealed to the hearer to acknowledge that she was right. This is an interesting case, because B is quoting what s/he said to the canteen staff; and s/he doesn't just quote the content of what s/he said but dramatizes the negotiation with 'you know'. This suggests that at some level B is aware that 'you know' mitigates the FTA - or at least is the kind of thing one says when disagreeing.

(4.76)

35C They came round on the third and I got my licence on the fifth
36C But basically they came round
37C and I didn't have one at the time
38C and I got that through this morning <,>
39C The other page is first actually <,,>
40B Right <,>
41B ***you know***
42B the problem with that is they can fine up to twelve hundred pounds
43C Yeah <,>
(S1A-078-ICE-GB)

In example (4.76), speaker B, a university staff member has to tell C, a student who failed to buy a TV license on time, that her problem could cost her up to twelve hundred pounds. *You know* here can help alleviate the damaging impact made on both the speaker and listener's face and at the same time appeal to the hearer to come to terms with the bad news.

4.2.2.2.4 Interactional repair

Interactional repair is another subfunction where you know seems only to play a supporting role because it only occurs twice and in both cases *you know* co-occurs with *I mean*, which is rather frequently used to function in this subfunction. As discussed in section 3.2.2.1, this subfunction is termed 'interactional repair' because what is repaired is the social inappropriateness that the message could convey. The following are the only two examples where *you know* plays the role of interactional repair found in the data of this study.

(4.77)

214B The B B C were like this third party that kind of came in
215B and there were all sorts of dictates from above what they needed that year and uhm and what and what <,> the viewers wanted to see and uhm
216B and so what whatever kind of things have been important to your to your script were were trivialised really in them to get what they wanted to get
217B But uhm *I mean* ***you know*** there might be an element of sour grapes in in what I 'm saying you know but I think uh I think that a lot of writers <,> have had that that experience with television
218B You end up by feeling quite compromised <,>
(S1A-058-ICE-GB)

B, a screen play writer, is complaining about his experience of working with

television. It seems that some repair work is needed after the harsh comments on BBC are made in lines 215 and 216. Firstly, making negative comments is face-threatening. Secondly, the negative comments may sound unfair to BBC because B lately wrote a screen play for BBC but it has not been screened. To repair the inappropriateness, B admits that "there might be an element of sour grapes in what I'm saying" in line 217.

Extract (4.27) was also quoted as an instance where *I mean* plays the role of interactional repair in section 4.1.2.2.1. As discussed in section 4.1.2.2.1, the statement made in line 96 is rather face-threatening because it sounds as if A is questioning what B has just said about dance therapy. In line 97, A turns the statement into a less face-threatening question, which invites B to give his/her opinion.

(4.27)

94B and people now do things like art therapy and dance therapy <,> uh which is great

95B It's very good work but I think <,> that art and dance actually include <,> the idea of therapy within them <,>

96A When you say recovering the whole person it suggests that there is something lost

97A *I mean* ***you know*** is there something incomplete

98A I mean<,> uh what's <.>w what's what's incomplete <unclear-words> <,>

99B Yeah

(S1A-004-ICE-GB)

The above four subfunctions under the interactant-relationship category are all played by both *you know* and *I mean*. But the following subfunctions under this category are only played by *you know.*

4.2.2.2.5 Indicating marked expressions

This subfunction is termed 'indicating marked expressions' because *you know* is used to signal to the hearer that the message marked by *you know* is expressed in a marked way i.e. the linguistic expressions chosen by the speaker are not what the hearer might expect. *You know* in this study is found to indicate mainly three types of

marked expressions - markedly formal or technical, markedly metaphorical and markedly exaggerating.

(4.78)

89A You mean even if you and Bernard had stayed together

90B Even if we had stayed together there might have <,> ***you know*** had endless hormones and glands problems

91A No

92A Yes you might

93A It's quite true <,>

94A She might have been that kind of teenager anyway

(S1A-031-ICE-GB)

In example (4.78), A and B are talking about B's teenage daughter (referred by 'she' in line 94) from B's previous relationship. In line 90, instead of saying 'lots of sexual problems' the speaker chooses 'hormones' and 'glands', which are very technical. As these technical words would not be expected in this kind of daily conversation, with *you know* B can signal to A that the coming expressions are marked and at the same time appeal to A to accept them.

The following two extracts are examples to show that the marked expressions are markedly metaphorical. In (4.79) the message marked by *you know* is expressed by personification while in (4.80) a simile is used.

(4.79)

134A I I found this sort of thing I got from Habitat a square thing <,>

135A Like a tadpole tank really

136A Tall square tall sort of oblong thing <,>

137A And I have to and it sort of comes out about three times a year and the rest of the time it skulks somewhere ***you know***

138A That's the trouble

(S1A-019-ICE-GB)

According to the Longman Online Dictionary, 'skulk' means 'to hide or move about secretly, trying not to be noticed, especially when you are intending to do something bad'. So 'skulk' is a verb which should only apply to human beings or perhaps animals. In line 137, vase is treated as if it were a human by the speaker. *You know* here is used to indicate that 'it skulks somewhere' is marked and at the same time appeal to the hearer to appreciate the creativeness of this personification.

(4.80)

228A Actually he <.>w he he looked like a polished
229B And what kind of body would you like
230B Well about nine inches
231A No no no no no
232A I was just saying that checking who you meant
233A ***you know*** he looked like a polished conker
234A He was really He was shiny and absolutely smooth
235A Not a hair
(S1A-080-ICE-GB)

In example (4.80), *you know* prefaces a simile, which compares a human body to a polished conker. This simile certainly makes the description of a human body more interesting and vivid than does a more usual or unmarked description like 'He was shiny and absolutely smooth' (line 234). Again, *you know* here not only signals the markedness but also appeals to the hearer to appreciate the originality of the simile.

(4.81)

38B Even more frightening than knowing they can understand what you're saying
39C That's what I said
40B Oh I see I thought you said it was very frightening being able to understand what they were saying
41D Yes they they know too much about <unclear-words>
42C Yes
43A Uhm <,>
44B Yes the English are branded on their tongue as they say don't they so uh as soon as you speak ***you know*** they usually know what an idiot you are
45A So this one was <,> lower middle-class in that case
(S1A-020-ICE-GB)

In example (4.81), the message introduced by *you know* can be seen as a dramatic or joking way of saying 'we can easily tell from which social class people come by their English accents'. Obviously, compared to the unmarked version, the exaggerating expression can make the conversation more interesting or entertaining. *You know* here not only indicates that the coming message will be put in a marked way but also urges the hearer to appreciate the joke.

You know used in this subfunction is also observed by other researchers. For instance, Le Lan (2007) observes that *you know* is often used with hyperbole (Le Lan, 2007:110). Extract (4.82) is her example where *you know* is used with hyperbole.

(4.82)

> Um.. the ghastly thing was the …senior common-room conversation that one had to conduct..which was just frightful and the …absolutely grotesque ritual of high table formal dinner in the evening which started with sherry in the senior common room with the …president sitting there like God.. ***you know***, talking absolute insanity on a very profound level (b and c laugh)
> (CEC, 1.3.3.2.5460) (Le Lan, 2007: 109 cited as example (5), prosodic information omitted)

According to Le Lan, the speaker in (4.82) is talking about the formal dinner ritual she had in a women's college. Clearly, exaggerating and formal expressions such as 'absolute insanity' and 'profound level' help to present a rather entertaining description of her experience. With *you know*, the speaker urges the listener to appreciate the creative wording.

Van Bogaert (2007: 16) also reports that *you know* is used to signal that the speaker is using figurative language or an unconventional turn of phrase and appealing to the hearer to accept his/her metaphor, comparison or imaginative use of language. Obviously her observation is very close to my description here. The main reason for this can be due to the fact that she also analyzes the spoken data of ICE-GB although mine only includes face-to-face conversations.

In contrast to the subfunction of indicating marked expressions, in the following two subfunctions *you know* is used to signal to the hearer that there is nothing unexpected or unusual in the coming message.

4.2.2.2.6 Indicating the most likely event

This subfunction is coded as 'indicating the most likely event' because the two messages linked by *you know* are about life routine, which is assumed to be shared by the speaker and hearer. With *you know*, the speaker urges the hearer to predict what the coming message is from the prior message by referring to the shared life

experience.

(4.83)

79A How old is it <,>
80B Who
81A Baby
82B Eight months <,>
83B It's just crawling
84A Is it cute <,>
85B It’s very cute i
86B It's got a grin that's sort of like up to here I mean
87B It says mmm like this <,>
88B And they wanted to photograph it for some baby magazines
89B not my cousin's
90B But uhm they took it to a shop <,> ***you know*** to get some baby clothes and so on
91B and the woman in the shop said
92B oh
93B you must let me photograph your baby for my magazine
(S1A-039-ICE-GB)

In example (4.83), B is talking about her cousin’s baby who was very cute. From line 90, B starts telling the story of how the baby magazine came to know the baby. The message ‘to get some baby clothes’ introduced by *you know* is one of the most likely reasons for taking a baby to a shop. In contrast, the encounter with a woman from a baby magazine who wanted to photograph the baby would be a low probability event. With *you know*, the hearer’s episodic memory of taking babies to shops can be activated. So instead of simply waiting to be told why B’s cousin took the baby to a shop, the hearer would search her memory to guess what could be the most likely reason.

(4.84)

83A Can you describe to me if possible a typical day in your home when you were a boy of less than fourteen
84B Uhm <,,> uh typical a typical day uhm <,,>
85B Right
86B School day'd be a typical day yes
87B Uh we'd get get get up
88B ***you know*** then my sister'd get up and uh <,> uh have breakfast and then <,,> go to school
(S1A-076-ICE-GB)

In example (4.84), B was asked to describe a typical day at home when he was a teenage boy. *You know* here is used to signal what most likely to happen after B got up in the morning. At the same time, *you know* appeals to the hearer's knowledge about what a typical school day would be like.

4.2.2.2.7 Indicating unspoken message to be completed by the hearer

Like 'indicating most likely event', *you know* in this subfunction also indicates that there is nothing unexpected in the coming message. But the speaker chooses to leave the message introduced by *you know* to be completed by the hearer.

(4.85)

100A I mean he be one of my my great mates you know you behind my back he was <,> slagging me
101A You see he told somebody I was weak <,,>
102A You see he was he tried to pull my mate's wife one night in Tramps
103A My my mate Peter is married to Cleo Goldsmith one of the Goldsmith family right and he says to
104A Terry pulled her aside and said why are you with these two
105A And I be one of his mates
106A And why are you with these two
107A they're both weak people <,>
108A And she came back and told her husband you see
109A She said ***you know***
110A So anyway <,,> and then and then he rang me up and I wouldn't answer the phone and I picked it up and said you've been telling people I 'm fucking weak you know

(S1A-052-ICE-GB)

In example (4.85), B is telling a story of Terry, one of his mates, who was slagging him behind his back. In line 109 speaker B does not finish quoting what Cleo told her husband, Paul, another mate of B. Since the conversation between Terry and Cleo has just been mentioned in lines 104 to l07, the speaker does not need to complete the quotation because the listener should be able to work out what the unfinished message was. By doing so, the speaker can not only avoid unnecessary repetition but also make the conversation more engaging by inviting the hearer to mentally participate in the conversation.

4.2.2.2.8 Indicating the coming message is meant to be evaluated

You know in this subfunction signals what the speaker is about to say is meant to be understood as evaluative and at the same time appeals to the hearer to see and agree with the evaluation.

(4.86)

147A Local school but <unclear-words>
148E May even be better than boarding school if you <unclear-words>
149C Well they don't think like that
150C Our parents don't think like that
151C They think that if you pay it must by definition be better
152E Yeah
153B Yeah yeah <,>
154E Right
155C It's funny because much to my amusement some of them send their <.>s send their girls off to ***you know*** local fee-paying schools who are you know not nearly as good as we are

(S1A-012-ICE-GB)

In example (4.86), C is commenting on his parents' idea that if a school needs to be paid, it must be better. With *you know*, C signals that 'local fee-paying schools' is not just a factual description of schools and appeals to his listeners to see and agree with his evaluation that 'how silly it is to go to local fee-paying schools'.

(4.87)

57B And so it's <,> totally accessible <,> for anybody <,>
58B Uhm it's linked to the Spinal Unit so that <,> patients from the Spinal Unit can <,> uhm associate with <,> the public and get back into <,> life all the more quicker because of the contact with <,> with people <,> instead of just ***you know*** the <,> the hospital staff and other patients <,>
59B Think it's a a vital <,> need <,,>

(S1A-003-ICE-GB)

In example (4.87), B is talking about how Mike Heafy centre (referred by "it" in line 57) benefits patients from the Spinal Unit. Similar to example (4.37), the factual description 'the hospital staff and other patients' in line 58 marked by *you know* is meant to be evaluated. *You know* here appeals to the hearer to see and agree with the speaker's opinion that only having contact with the hospital staff and other patients is not good for the recovery of patients from the Spinal Unit.

4.2.3 Speaker-oriented

The first two subfunctions of *you know* in this category are relatively easier to identify because the verbal environments of *you know* in them are very specific and offer very little room for alternative interpretations. Although researchers do agree that *you know* can play these two subfunctions, there is not a consensus as to what terms should be used.

4.2.3.1 Hesitation marker

Both Erman (2001) and Müller (2005) identify this subfunction but label it in different terms. This study takes Erman's term, hesitation marker, because *you know* in this subfunction is used to buy more time to think about what should be said next. Müller terms this subfunction 'marking lexical or content search' by following Östman's description (1981) where the searching function of *you know* is further categorized into lexical search and content search. However, Müller does point out that "it is a bold move to maintain that the two types can (always) be distinguished" (2005: 158) because "the category 'lexical or content search' contains a continuum of instances with clear lexical search at one end, apparent content search at the other, and cases with elements of both in between" (2005: 160).

Erman (2001) gives a very detailed account of the linguistic environments of *you know* as a hesitation marker. For instance, she observes that *you know* usually occurs after function words or after a determiner. Similar patterns are also identified in this study.

(4.88)

155B I think she feels probably that uhm ***you know*** unless I need her for <,,>***you know*** <.>mak making my life happy then I don't really want her

156B And clearly I've got what I need to make my life happy

(S1A-031-ICE-GB)

In example (4.88), both *you know*s are identified as hesitation markers. The first *you know* occurs after the complementizer *that* while the second *you know* appears after the preposition *for*. The two markers of hesitation *uhm* and pauses here clearly help

the speaker win more time. Example (4.89) is another instance of hesitation marker where *you know* occurs after a determiner and is prefaced by a pause. This example supports Östman's observation that "if *you know* was used after an article, that article had to be repeated after *you know*" (Östman, 1981: 31).

(4.89)

127A I did buy a <.>bu ***you know*** a bunch of roses but hardly ever a great big thing like like uh uh a ribbon and <unclear-words> sheet is it really

128E They're quite expensive nowadays

S1A-019-ICE-GB)

You know in hesitation marker can not only help the speaker win more time but also can appeal to the hearer for patience or even help. Example (4.90) is an instance of hesitation marker where the appealing function seems to work for the speaker because it is the listener who gives the answer before the speaker finishes her searching.

(4.90)

94A She might have been that kind of teenager anyway

95B So

96B Mm

97B Quite likely I think

98B I was quite like that and <,> ***you know***

99A Bolshie

100B Yes

101B And and also apt to take you know very completely irrational hates against people for what I think were probably sexual reasons

102B I mean<,> you know at the bottom it was sort <,> uh sexual problems

(S1A-031-ICE-GB)

4.2.2.3.2 Restart

This subfunction is identified by a number of researchers (Goldberg, 1980; Schourup 1985; Holmes, 1986; Erman, 2001; Müller, 2005) but is referred to by different terms. The following are examples of this subfunction cited in previous studies. It is clear that *you know*s in the following examples all share the same linguistic environment: the prior message is not finished and usually there is a syntactic change in the coming message. *You know* in this context is termed repair by most researchers (Goldberg, 1980; Schourup, 1985; Erman, 2001; Müller, 2005) because there is a

syntactic correction involved.

(4.91)

Well, it's just none of their- ***you know***, that's really none of their business.
(White House: 28-2-73:43) (Goldberg, 1980, quoted as example (70) in Schourup, 1985: 123)

(4.92)

… they can get- ***you know*** they can get close (ta you)
(RTS18t, 56) (Schourup: 1985 as example (91))

(4.93)

The question is are you actually interested, attracted to her enough to want to, ***you know***, what are you really interested in doing /…/
(Erman, 2001: 1345 as example (13))

(4.94)

Young man to friends at dinner party
and I've been on this bloody speed reading course which is / ***you know*** so one/ one notices
(Holmes, 1986: 11 as example (22))

However, example (4.91), according to Müller's criteria, will be coded as marking false start rather than repair because no correction but repetition is made because repair in Müller's study only applies to those cases where the coming message starts with a different sentence structure from the prior message.

This study labels this subfunction as restart rather than repair because repair is used as a cover term in this study. In the analysis of *I mean* in section 4.1, four subtypes of repair were identified and restart was one of them. Example (3.39) is an instance of restart of *I mean* (quoted in section 4.1.2.3.2 and repeated here for convenience). As can be seen from example (4.39), there is also a syntactic correction involved.

(4.39)

57A I would because to me <,> it seems
58A ***I mean*** I 'd go to that and I 'd go to the Palmer one if I was you
(S1A-005-ICE-GB)

Like Müller, this study also makes a distinction between whether a correction is made or not. Example (4.95) is treated as a case of restart because a syntactic change is made in the coming message. Example (4.96), where only repetition is made, is put under the above subcategory of hesitation marker because repetition is

considered as a marker of hesitation.

(4.95)

98D Well it be handed in a week on Monday and I've been doing it for about three weeks now

99D And I just ***you know*** when you just can't be bothered it's so boring

100D It's on the National Curriculum

(S1A-040-ICE-GB)

(4.96)

91B Platform in a pond really just to a sort of walkway for kids to look up close because they 're always taking parties of kids around to have a look at the site <,>

92B It's ***you know*** it's a nature reserve

(S1A-081-ICE-GB)

4.2.3.3 Approximator

This subfunction is termed approximator because *you know* is used to signal that the coming message is rather rough or vague. The reason for the imprecision could either be because the speaker could not find the right word at the time of communication or because the accuracy of the message may not be important. This subfunction is identified by both Erman (2001) and Müller (2005). Similar to Erman's findings (2005), as an approximator *you know* is often found to work with other approximators in this study. Among them, *sort of* is the most frequent one, which supports Aijmer's finding that there is a large number of examples of the collocation *sort of you know* (Aijmer, 2002: 189). Extract (4.97) is an instance where *you know* occurs with *sort of* while extract (4.98) is a case where *you know* co-works with another less frequent approximator *whatever*.

(4.97)

32B were you lost in Richard 's seminar the other night

33C Yeah right

34D No I was super <,>

35D I really enjoyed it

36B Did you get the implications of everything then

37D No

38B Uhm cos Caroline was just asking what it was about the other night <,,>

39D Well it was he was ***you know*** *sort of* <,> dodging about <,> about family trees and about and about Deor and about <unclear-word> and about those things and it was interesting <,>

40D I don't I don't suppose there was anything fantastically useful
(S1A-090-ICE-GB)

(4.98)

85B and it's based on your memory <,>
86B So even if you write it down as a script it 's still only going to be what you could in fact actually tell us <,,>
87A Yes
88C There are occasions where
89A Yes
90B So your your range of events or ***you know*** *whatever* it is for emphasis is going to be from you anyway <,>
(S1A-064-ICE-GB)

With *you know*, the speaker not only can warn the hearer that the coming message lacks precision but also appeal to the hearer to cope with the imprecision.

4.2.3.4 Introducing a new topic

You know in this subfunction plays the role of introducing a new topic. It is identified by Müller but is not included on her function list because it occurs less than three times. In this study, this subfunction occurs at moderately low frequency (14 out of 692).

(4.99)

1A Uhm <,> you can break into the pears if you want to or have a piece of choccy
2A You've had plenty of veggies <,,>
3B A piece of choccy
4B What do you mean
5A Well you bought some and I bought some
6B Oh you bought some did you
7B Oh
8A I bought exactly the same thing only I bought the uh the Tobler version
9B Oh
10B I wouldn't mind uhm a pear
11B Just one
12A OK <,,>
13A ***you know*** Gerry actually talks refers to my mother's bedroom as the boudoir
14A She doesn't <laugh>
15B Seriously <,,>
16B When she answers the phone she says's residence
(S1A-023-ICE-GB)

In example (4.99) prior to *you know*, A and B are talking about food. In line 13, with

you know A switches to another topic, which is about Gerry who refers to A's mother's bedroom as the boudoir.

However, Goldberg (1980, 1981) reports that *you know* is frequently used to preface a new topic in her data. She also points out that this topic changing function is face-threatening. Example (4.100) is one of her examples.

(4.100)

President Nixon is talking to John Dean. Dean has just entered the room

P Hi John, how are you?

D Good morning. Good morning.

P Sit down. Sit down. Trying to get my remarks ready to (unintelligible) the building trades.

D So I understand.

P Yes, indeed, yeah. ***You know***, I was thinking we ought to get the odds and ends, uh (unintelligible) we talked, and, uh, it was confirmed that—you remember we talked about resignations and so forth and so on—that I should have in hand, not to be released.

(White House: 16-4-73am:187) (Cit. in Goldberg, 1980:96) (Quoted as example (55) in Schourup, 1985: 107)

As can be seen from (4.100), *you know* is used between phatic talk and the main theme of their talk. The reason that the topic transition in this example is interpreted as face-threatening could be because the new topic *you know* prefaces happens to be the face-threatening act of giving a command because President Nixon is telling John Dean what he should do. The topic change itself may not be inherently face-threatening as Goldberg claims (Goldberg, 1981:3). At least, the reading of the topic change in example (4.99) above does not sound face-threatening.

4.2.3.5 Indicating speaker attitude

Like *I mean*, *you know* is also used to mark speaker attitude. As discussed in section 4.1.2.3.5, *I mean* can indicate three types of speaker attitude - expressing evaluation or judgment, expressing emphasis and expressing sincerity. As shown by following examples, *you know* is only used to indicate the first two speaker attitudes perhaps because the attitude of expressing sincerity is very closely related to the semantic meaning of *I mean*. *I mean* in the context of expressing sincerity can be translated as

‘I’m serious when I say’.

(4.101)

95C Anyway anyway anyway Louis tell us some tell us more about her
96A I don't know much about her
97A I've got her phone number right here
98A The thing is what's the etiquette of this
99A You're then meant to wait a couple of days before you ring them up or else it appears uncool ***you know***

(S1A-021-ICE-GB)

In line 99, A expresses his attitude towards the etiquette that you cannot phone the girl right after you have got her phone number by using the evaluative adjective ‘uncool’.

(4.102)

124B I I I really <,> love it
125B I I look forward to <,> to working each week
126B I mean I'd like to do it every day <,,> but ***you know*** such is life <,,>

(S1A-003-ICE-GB)

In line 126, B uses a statement ‘such is life’ to express his/her attitude that life would not allow you to do what you want.

(4.103)

277E And yet uh I know everybody laughs about Barbara Cartland but they're fabulous to read
278E They're so <,> innocent
279E And it's so <.>honestly <.>the they
280D Yeah uhm
281D Yeah
282D They are
283D They're just so lovely <,>
284D So easy to read ***you know*** I mean and now they hers are all lords and ladies and all

(S1A-016-ICE-GB)

Example (4.103) is an instance of expressing emphasis. The adverb ‘so’ in line 284 helps express the emphasis here. Again, *you know* is not only used to indicate speaker attitude but also appeal to the hearer’s agreement.

You know used in the context of expressing emphasis is also observed in previous studies (Holmes, 1986; Erman, 2001). Extract (4.104) is quoted by Erman as an example of *you know* with emphatic function, “where the speaker urges the listener

to appreciate the force of the utterance as a whole" (Erman, 2001: 1347)

(4.104)

<1> I didn't realize what I was doing. I dunno.
<2> You're so stupid! ***You know***.
<1> Yeah, yeah erm.
Melanie was talk … I was talking to Melanie about <unclear> and Melanie goes to me, <unclear> if you go out with him, and you realize that I'm mad, and you your eyes out of your head!
<2> (laugh)
(B133203, Erman, 2001: 1347 cited as example (18))

The above is the complete description of the pragmatic functions of *you know* identified in the native speaker data of this study. But identification of them is not always straightforward because like *I mean*, *you know* was also found to play more than one function in one context i.e. there were cases where decision had to be made as to which function was more typical. The following two examples show how such decisions were made.

(4.105)

244B Uhm though at the weekends <,> I had because I was on this this silly residential course and things I kept having meals with this are you vegetarian or aren't you
245B and I wasn't allowed to say I was a whimsical vegetarian and fancied a vegetarian option
246B ***you know*** you there are only three vegetarian dinners here
247B Where are the vegetarians to give them the vegetarian dinner
(S1A-011-ICE-GB)

You know in example (4.105) was coded as quotative because it is followed by direct quoted speech, which is the typical context for quotative *you know*. However, it could be argued that the quote introduced by *you know* explains why 'I wasn't allowed to say I was a whimsical vegetarian and fancied a vegetarian option'. Therefore, *you know* could also be coded as reason.

(4.106)

107A Yes well that's fifteenth century isn't it <,>
108B I think so
109A They've got a salt box haven't they over the near the fire ***you know*** *sort of* hollow beam where you keep the salt or used to keep the salt dry
110A Yes
(S1A-009-ICE-GB)

Like example (4.105), *you know* in example (4.106) can also be seen to play two roles. But it is very difficult to tell which role is the primary one. The information introduced *you know* gives additional information about the salt box, so it can be seen as an instance of explicitness. But at the same time, *you know* can also be seen as a signal warning the hearer the upcoming description is not accurate and appealing to the hearer to get on with the imprecision because it co-occurs with an approximator *sort of*. This example was finally coded as explicitness because it seems that *you know* only plays a supporting role when functioning in approximator because it always co-occurs with markers of approximation such as *sort of* or *whatever*. On the contrary, explicitness can be seen as one of the main subfunctions played by *you know* because *you know* was rather frequently (46 out of 590) found to perform this role independently.

4.3 Comparison of *I mean* and *you know* in the native speaker data

Since a detailed description of *I mean* and *you know* in the native speaker data has been completed in the above two main sections, it is the right time to compare their behavior in English conversation because although they are compared in the literature the comparison is either very brief (David & Davy, 1974; Brown & Levinson, 1978; Edmonson, 1981; Östman, 1981; James, 1983; Schourup, 1985) or at a general level (Schiffrin, 1987; Fox Tree & Schrock, 2002). As can be seen from the above discussion, the analysis of *I mean* and *you know* was carried out not only in great depth but also under the same broad framework with the same translation technique. As a result, a more delicate and systematic comparison of *you know* and *I mean* is hoped to be achieved. Since the comparison between *you know* and *I mean* will be conducted from three aspects, this section will be further divided into three subsections. Section 4.3.1, will present the comparison in terms of their overall frequencies, degree of pragmatization, distributions of the specific functions. Section 4.3.2 will then attempt to account for the similarities and differences revealed in section 4.3.3.The final section will compare them in terms of turn positions and propositions.

4.3.1 Quantitative findings of the comparison of the use of *I mean* and *you know* in the native speaker data

Table 4.8 below shows the overall frequencies of *you know* and *I mean* in the face-to-face conversations chosen in this study followed by a chi square test of degree of pragmaticalization of the two markers. As can be seen from Table 4.8, both linguistic expressions are frequently used, with a slightly higher frequency of *I mean*. However, *you know* is less pragmatized than *I mean*, with about 10% more semantic cases. The subsequent chi-square test shows that *I mean* is significantly more pragmatized than *you know*.

Table 4.8 Overall frequencies of *I mean* and *you know* in the native speaker data

	Identified			Unidentified	Total
	Pragmatic	Non-pragmatic	Total		
I mean	686 (95.8%)	30 (4.2%)	716	157	861
you know	587 (84.9%)	104 (15.1%)	691	132	823

	pragmatic	non-pragmatic	Total
I mean	686	30	716
you know	587	104	691
Total	1273	134	1407
X^2=48.136, df=1,p＜0.001			

The significance between *I mean* and *you know* in terms of their degree of being pragmatized appears to echo my observation that in my data for *I mean* the distinction between semantic and pragmatic cases seems clear-cut while *you know* seems to be still in the process of being pragmatized. Example (4.107) is a borderline case of *you know* where *you know* is used as one unit to replace the embarrassing words in line 168.

(4.107)

163A The Sport
164D The Daily Sport
165C So he was telling me you know what was in it
166C So I said well bring them home
167C Let me have a look <,,>

168C you know just for the sheer <,> ***you know*** of it
169B <laugh>
(S1A-027-ICE-GB)

Table 4.9 below shows the pragmatic functions of *you know* and *I mean* identified in the British speakers' data and their distributions, followed by a chi-square test of the three main categories. As shown by Table 4.9, *you know* and *I mean* display similar pragmatic properties because more than half of the pragmatic functions (16 out of 26) are shared by them and the chi-square test shows that there is no significant difference between them in terms of the three main categories.

Table 4.9 Pragmatic functions of *I mean* and *you know* identified in the native speaker data

		I mean (686)	*you know* (587)
Hearer-oriented	Assumption correction	73(10.6%)	1(0.2%)
	Exemplification	48(7.0%)	40(6.8%)
	Explicitness	42(6.1%)	46(7.8%)
	Reformulation	16(2.3%)	13(2.2%)
	Cause	8(1.2%)	7(1.2%)
	Reason	24(3.5%)	16(2.7%)
	Result	2(0.3%)	7(1.2%)
	Summarization	10(1.5%)	11(1.9%)
	Quotative	2(0.3%)	29(4.9%)
	Introducing background information	0(0.0%)	11(1.9%)
	Seeking confirmation	0(0.0%)	2(0.3%)
	Total	225(32.8%)	184(31.2%)
Interactant-relationship-oriented	Softener of FTA	39(5.7%)	13(2.2%)
	Interactional repair	22(3.2%)	2(0.3%)
	Justification	191(27.8%)	41(7.0%)
	Conclusion	13(1.9%)	10(1.7%)
	Indicating marked expression	0(0.0%)	48(8.2%)
	Indicating the most likely event	0(0.0%)	19(3.2%)
	Indicating the unspoken message to be completed by the hearer	0(0.0%)	18(3.1%)
	Indicating the coming message is meant to be evaluated	0(0.0%)	67(11.4%)

	Total	265(38.6%)	218(37.1%)
Speaker-oriented	Indicating speaker attitude	51(7.4%)	76(12.9%)
	Hesitation marker	9(1.3%)	44(7.5%)
	Restart	90(13.1%)	39(6.6%)
	Resumption	14(2.0%)	0(0.0%)
	Transactional repair	32(8.5%)	0(0.0%)
	Approximator	0(0.0%)	13(2.2%)
	Introducing a new topic	0(0.0%)	14(2.4%)
	Total	196(28.5%)	186(31.7%)

	three broad categories			
	Hearer-oriented	Interactant-relationship-oriented	speaker-oriented	Total
I mean	225	265	196	686
you know	183	218	186	587
Total	408	483	382	1273
X^2=1.469, df=2, p>0.47				

Under the hearer-oriented category, the British speakers use both *I mean* and *you know* to clarify the previous message and even use similar proportions of *I mean* and *you know* in the three specific ways of clarification i.e. giving examples, providing more detail and paraphrasing. They also use *I mean* and *you know* to link messages which have cause-effect relationship i.e. X happens because of Y or X happens, so Y happens. In addition, similar percentages of *you know* and *I mean* are used to summarize the previous talk.

Under the interactant-relationship category, *I mean* and *you know* are both used to soften FTAs, which supports Brown & Levinson's (1978) observation that *I mean* and *you know* are found to act as 'hedges' in politeness strategies when the speaker performs face-threatening acts. They are also used to signal that the upcoming message explains why the prior message has been said or that the upcoming message is the inference that can be drawn from the prior talk.

Under the speaker-oriented category, both *I mean* and *you know* are used to indicate the attitude of the speaker and help the speaker to edit his/her real time communicating by either stalling for more time or making syntactic repair.

Under each broad category, there is one or two subfunctions where both *I mean* and *you know* can occur but there is a substantial gap between them in terms of their frequencies. Assumption correction and transactional repair are predominately played by *I mean*. *You know* seems to play just a minor role because it is very rarely used in these contexts and always co-occurs with *I mean*. Similarly, *I mean* seems to only play a supporting role in quotative and hesitation marker because it shows very low frequencies in both subfunctions as compared to that of *you know* and it almost always co-occurs with *you know*.

To sum up, in the face-to-face conversations chosen in this study, both *I mean* and *you know* are frequently used pragmatic markers although *I mean* shows a slight higher frequency. They share over half of the subfunctions identified in this study and there is no significant difference between them in terms of the distributions of the three main broad categories.

However, *you know* is more versatile than *I mean* with regard to the number of subfunctions they can play. The subfunctions played by *you know* is more evenly distributed than those played by *I mean*. For instance justification in *I mean* accounts for as high as 27.8% while the subfunction of ‘indicating the speaker attitude’, the most frequent subfunction of *you know*, only accounts for 12.9%.

The similarities and differences between *I mean* and *you know* in terms of the specific functions they played can be summarized in five categories. The first category includes those subfunctions such as explicitness, reformulation, exemplification and summarization where *I mean* and *you know* show rather similar frequencies. The second category includes subfunctions such as assumption correction and interactional repair, which are predominantly played by *I mean*. The

third category includes subfunctions such as quotative and hesitation marker, which are predominantly played by *you know*. The fourth category is *I-mean*-only category which includes resumption and transactional repair. The final category is *you-know*-only category which includes subfunctions such as 'introducing a new topic', 'indicating the coming message is meant to be evaluated', 'indicating marked expressions' and 'indicating the most likely event'.

4.3.2 Accounting for the features of the use of *I mean* and *you know* in the native speaker data

As shown in section 4.3.1 *you know* and *I mean* can occur in the same contexts because they "may actually accomplish the same interactive task" (Schiffrin, 1987: 309), but it does not mean that they are identical in those overlapping subfunctions. It seems that their general functions or core meanings can help account for the difference between them. Both Edmondson (1981) and Schiffrin (1987) propose that the general functions of *you know* and *I mean* are complementary due to their semantic meanings. *You know* and *I mean* are both considered as fumbles in Edmondson's study but put in two groups. *You know* belongs to a group called "the Cajoler" because it is used by the speaker "as an appeal for understanding i.e. through the Cajoler the speaker seeks to make his illocution more palatable to the hearer" (Edmondson, 1981: 154) while *I mean* is in a group named "the Let-me-explain" because it is used to "communicate the fact that I'm trying to communicate" (Edmondson, 1981: 155). For Schiffrin, *I mean* "focuses on the speaker's own adjustments in the production of his/her own talk, *you know* proposes that a hearer adjust his/her orientation (specially knowledge and attention) toward the reception of another's talk"(Schiffrin, 1987: 309). So for all the contexts where *you know* and *I mean* can both be applied, the complementary role that they play can be summarized as follows: with *I mean* the speaker spells out his/her own message and in the same time can diminish potential imposition on the hearer by suggesting 'this is what I mean, so you do not have to agree with me' while with *you know* the speaker not only spells out his/her message but in the same time appeals to the hearer

to understand, appreciate or agree to the spelt out message.

Since in the shared contexts both *you know* and *I mean* can occur, it is not surprising to find that *you know* and *I mean* co-occur in nearly all shared subfunctions except cause, result and conclusion, which are all rather infrequent. The switch from *you know* to *I mean* or vice versa seems to suggest that the speaker keeps adjusting how much he/she wants the hearer to get involved in the conversation or assuming how much the hearer knows. It seems that such a switch is not random because the occurrence of *I mean you know* (24) is about two times that of *you know I mean* (10). However, it is not clear why the British speakers switched from *I mean* to *you know* more often than vice versa.

It can be seen from the above discussion that in the subfunctions where both *you know* and *I mean* can occur, it is their general functions that tell one from the other i.e. apart from the same subfunctions they both play, with *you know* and *I mean*, the speaker can choose whether to invite the listener to participate or just focus on his/her own talk. As to the subfunctions where only one of them can fit, it is, again, their general functions that can help answer the question why only one of them works because as can be seen from the following discussion, in the *you-know*-only subfunctions the listener's participation is essential while in the *I-mean*-only subfunctions the listener's participation is least needed or even impossible.

In the hearer-oriented category, the two subfunctions of 'introducing background information' and 'seeking confirmation' are only played by *you know*. As discussed in sections 4.2.2.1.2 and 4.2.2.1.9, when functioning in 'indicating background information', *you know* introduces background information that the speaker assumes the hearer needs to know so as to follow the talk while in the subfunction of 'seeking confirmation', with *you know* the speaker can make sure that the listener knows what exactly he/she refers to. So in both subfunctions the hearer's participation is crucial. *I mean* certainly would not be a preferred candidate here because it would discourage

the hearer's involvement. Although quotative is a subfunction where both *I mean* and *you know* can work, it is predominately played by *you know* because *I mean* rarely occurs in this context and if it does, it co-occurs with *you know*. The reason that *you know* is preferred to *I mean* in this subfunction could be because when we quote something, there must be something worth reporting in that quote. We do not really just quote something for the sake of quoting it. So with *you know*, the speaker can indicate the coming message is a quote and in the same time appeal to the hearer to see something is being quoted. The reason why assumption correction and interactional repair in the interactant-relationship category are predominately played by *I mean* will be discussed later when it comes to the comparison of how *you know* and *I mean* do their repair in conversation.

The appealing function, the general function of *you know*, again, is indispensable to *you-know*-only subfunctions in the interact-relationship category. In 'indicating marked expression', the speaker needs the hearer to appreciate his/her linguistic efforts of making the conversation more interesting. In 'indicating the most likely event' the speaker tries to claim a common ground between him/her and the hearer by appealing to the shared knowledge or experience. In 'indicating the unspoken message to be completed by the hearer' and 'indicating the coming message is meant to be evaluated', the speaker urges the hearer to complete the unspoken message and see the implication. It is self-evident that the general function of *I mean* would not fit in any of the above subfunctions.

In the speaker-oriented category, there are two *you-know*-only subfunctions and two *I-mean*-only subfunctions. Both resumption and 'introducing a new topic' play the role of topic management. 'Introducing a new topic' is only played by *you know* because the speaker would certainly hope that the hearer would be interested in the new topic and be willing to participate. On the contrary, resumption is only played by *I mean* because what the speaker needs is to resume his/her topic, which is interrupted in the prior talk. Instead of inviting the hearer's participation, the speaker

needs to focus on his/her own talk.

As discussed in section 4.2.2.3.3, when *you know* functioning as approximator, it tends to co-occur with other approximators such as *sort of* and *whatever*, which means *you know* might only play a supporting role here. Only signaling to the hearer that the coming message is not put in an accurate way seems not enough. There is a need to appeal to the hearer to cope with the imprecision. Obviously, a self-involvement marker like *I mean* does not really suit this occasion.

The fact that *you know* cannot be used in transactional repair is reported by Schourup (1985). According to Schourup, the reason that *you know* cannot appropriately used in transactional repair is because the repair is too radical. Example (4.108) is cited by Schourup to explain why *you know* cannot work in this context. He thinks that in example (4.108) "it is unlikely that the addressee could grasp the speaker's intention (cat) from that speaker's having said dog, without further context or foreknowledge" although *I mean* goes well in (4.108)(Schourup, 1985: 122).

(4.108)

? I got a dog YK cat for my birthday.
(Schourup, 1985: 122 cited as example (66))

As discussed in section 4.1, due to the semantic meaning of *I mean*, repair is one of the main jobs done by *I mean*. The five subtypes of repair identified in this study are hesitation marker, restart assumption correction, interactional repair, and transactional repair. As shown in Table 4.1 above, although *you know* was found to function in most of the repairing jobs, it was very infrequent and showed much lower frequencies than *I mean* except in hesitation marker. It can be argued that *you know* does not really contribute much in assumption correction or interactional repair because it occurred only one or two times and all co-occurred with *I mean*. The reason *you know* could occasionally occur in these two subtypes of repair is because there can be a potential need to appeal to the hearer to understand why an unwanted assumption needs to be corrected or why an inappropriate message needs to be

modified. But in transactional repair, what the speaker needs to correct is a piece of wrong information. Since the change involved in transactional repair is much more radical than that in either assumption correction or interactional repair, appealing to the hearer would not help because the hearer would not know the correct information. Only the speaker knows exactly what the information he/she has intended to provide. Therefore, it should be solely the speaker's job to make the correction and *I mean* is certainly a perfect candidate for doing this radical repair.

Apart from those subfunctions where both *you know* and *I mean* can play, there are a number of subfunctions where only one of them can work. The different general functions of *you know* and *I mean* seem to be able to account for why some subfunctions allow both pragmatic markers while others can only have one. In shared subfunctions such as clarifying previous message by being more specific, exemplifying and paraphrasing, the speaker can choose either to get the hearer involved by appealing to the relevance of the upcoming message or just keep the job to him/herself. Subfunctions such as 'indicating marked expression' and 'indicating the coming message is meant to be evaluated' are only played by *you know* because in those contexts the hearer's participation is essential. In contrast, in those *I-mean*-only subfunctions such as resumption and transactional repair the hearer's participation is least needed or even impossible.

To sum up, the above comparison of the pragmatic functions of *you know* and *I mean* gives a clear view of how *you know* and *I mean* behave in face-to-face conversation by highlighting the similarities and differences between them. The general functions of *I mean* and *you know* adopted in this study have successfully accounted for contexts where both pragmatic markers occur with similar frequencies; contexts where one shows a much higher frequency than the other; and contexts where only one of the marker occurs. Therefore the comparison of *I mean* and *you know* made in this study does seem to support the claim (Schourup, 1985: Holmes, 1986) that the approach of general function plus specific function will not only result in a unified

account of a pragmatic marker but also help account for why it is A marker rather than B marker that works in a certain context.

4.3.3 Comparison of turn positions and proposition positions between *I mean* and *you know* in the native speaker data

Table 4.10 below shows the distribution of turn positions of *I mean* and *you know* in the chosen British data in this study followed by a chi-square test. As shown by Table 4.10, for both pragmatic markers turn medial position is the preferred turn position for about 90% *you know* and *I mean* appear in this turn position. However, *I mean* very rarely occurs in turn final position while turn initial position is the least preferred one for *you know*. The chi-square test further shows that the difference between the two markers in terms of turn positions is significant. In other words, *you know* enjoys significantly more freedom than does *I mean* with regard to which turn position they can appear.

Table 4.10 Distributions of turn positions of *I mean* and *you know* in the native speaker data

	turn position			
	turn initial	turn medial	turn final	Total
I mean	60(8.7%)	623(90.8%)	3(0.4%)	686
you know	20(3.4%)	525(89.4%)	42(7.2%)	587
Total	80	1148	45	1273
X^2=54.798, df=2, p<0.001				

Table 4.11 below presents the distribution of proposition positions of *I mean* and *you know* followed by a chi-square test. As can be seen from Table 4.11, although proposition initial position is preferred by both pragmatic markers, the difference between them is significant according to the chi-square test.

Table 4.11 Distributions of proposition positions of *I mean* and *you know* in the native speaker data

	Proposition position			
	proposition initial	proposition medial	proposition final	Total
I mean	674(98.3%)	9(1.3%)	3(0.4%)	686
You know	319(54.3%)	114(19.3%)	154(26.2%)	587
Total	993	123	157	1273
X^2=356.232, df=2, p<0.001				

In contrast to *I mean*, which is predominantly used in proposition initial position and very restricted in the other two proposition positions, *you know* is more evenly distributed over the three positions, which can explain the general impression that *you know* seems to turn up anywhere in a conversation.

As explained in Chapter 3, when *I mean* and *you know* appear in proposition medial position they can occur in the positions of either 'within a constituent' or 'between constituents'. In this study, *I mean* very rarely occurs within a constituent and if it does, it co-occurs with *you know*. Among the nine cases of proposition medial *I mean*s, only two occur within a constituent while the other seven occur between constituents. Example (4.109) is an instance where *I mean* occurs within the prepositional phrase 'for three months'. In example (4.110), *I mean* occurs between subject 'the boat' and its predicate 'sleeps'.

(4.109)

269B And Nell Nell's breast-feeding only lasted for ***you know I mean*** three months drying up kind of thing <,>

(S1A-056-ICE-GB)

(4.110)

64C The boat seats uh ***I mean*** sleeps fifteen hundred you see

(S1A-021-ICE-GB)

In contrast, *you know* enjoys much more freedom in proposition medial position. About 40% (43 out of 114) proposition medial *you know*s appear within a constituent although more than half of them occur between constituents. Example (4.111) is an

instance where *you know* appears between verb ‘take’ and its complement ‘hates’. Example (4.112) is a case where *you know* occurs within a constituent. But note in this example that *you know* actually occurs between auxiliary verb ‘have’ and finite verb ‘done’, which further supports the impression that *you know* seems to be virtually everywhere.

(4.111)

101B And and also apt to take ***you know*** very completely irrational hates against people for what I think were probably sexual reasons

(SlA-031-ICE-GB)

(4.112)

152B I think we should have ***you know*** done a bit of digging at that point sort of you know make her feel particularly special and find out if she was feeling at all bothered about it or whether she thought she had to be nice to this child because again she was earning her place

(S1A-031-GB)

The proposition final position is where *you know* and *I mean* show the biggest gap. Only three (out of 686) *I mean*s are proposition final while about one in four *you know*s appears in this position. But it has to point out here that the proposition final *you know* include cases like example (4.113) where the position of *you know* can be more accurately coded as incomplete proposition final. In example (4.113), *you know* is used to introduce a message which is left unspoken and expected to be completed by the hearer. B leaves out what Alice says because A can easily figure it out given what has just been said in line 117.

(4.113)

112A we do all different kinds of ones this winter
113Z It's basically Surrey
114B Mmm
115A Yeah <,> up to Dorking
116 A From Leatherhead up to Dorking that area
117B Yeah we go back <,> you know when we're needed you know
118B My friend Alice who lives opposite <,> she organises the task
119B She rings up the wardens <,> and says ***you know***
120A Yeah

(S1A-081-ICE-GB)

To sum up, for both pragmatic markers turn medial and proposition initial are the most preferred positions. But overall *you know* enjoys more freedom to appear in any of the positions while *I mean* is particularly restricted in turn final, proposition medial and proposition final positions.

Chapter Five

Uses of *I mean* and *you know* in the non-native speaker data

This chapter consists of two main sections, which will display how *I mean* and *you know* are used by the Chinese EFL learners as compared to the native British speakers respectively. The comparison of each marker between the native speaker data and non-native speaker data will be carried out in terms of overall frequencies, distributions of specific functions, distributions of turn positions and distributions of proposition positions. So under each main section, there are four subsections.

Before moving on to the description of *I mean* and *you know* in the learners' data, it is important to note that this study holds a neutral view towards the differences between the learners and the native speakers: i.e. the differences do not necessarily indicate the learners' deviant or wrong uses of the two markers. Therefore, neutral terms such as over-representation or under-representation rather than terms like overuse or underuse are used.

5.1 Uses of *I mean* in the non-native speaker data

5.1.1 Overall frequencies of *I mean* in the non-native speaker data and the native speaker data

Table 5.1 shows the overall frequencies of *I mean* in the British speakers' data and the Chinese EFL learners' data followed by the chi-square result of the distributions of pragmatic and non-pragmatic cases of the identified cases.

Table 5.1 Overall frequencies of *I mean* in the British speakers' data and the Chinese EFL learners' data

	Identified			Unidentified	Total
	Pragmatic	Non-pragmatic	Total		
British data (ICE-GB)	686	30	716	157	861
Chinese EFL data (SECCL 1.0& 2.0)	82	13	95	4	99

	Identified cases		
	Non-pragmatic	Pragmatic	Total
British data (ICE-GB)	686(95.8%)	30(4.2%)	716
Chinese EFL data (SECCL 1.0 & 2.0)	82(86.3%)	13(13.7%)	95
Total	768	43	811
X^2=15.057, df=1, p<0.001			

As can be seen from Table 5.1 (above), the overall frequency of *I mean* in the speakers' data is nearly 9 times that of *I mean* in the Chinese EFL learners' data. As to the identified cases, not only do the Chinese students use *I mean* much less but the proportion of non-pragmatic uses is statistically much higher. Romero Trillo's (2002) study, where Spanish EFL learners and British speakers are compared, also shows that *I mean* is more frequent in native speaker data than in learners' data, but the frequency of *I mean* in his adult native data (London-Lund Corpus, 50,000 words) is only about 2 times that of *I mean* in his adult learners' data (spoken production of 3rd and 4th year English Philology students at Universidad Autónoma de Madrid Corpus, abbreviated as UAM-Corpus, 19,614 words). In contrast to the Chinese EFL learners' data where *I mean* is underpragmatized with only 86% pragmatic cases, *I mean* in UAM-Corpus is overpragmatized with 100% pragmatization given that the pragmatizated rate of *I mean* in the baseline data of both this study and Romeo Trillo's is about 95%.

The reason that *I mean* in the Chinese EFL learners' data is less pragmatized might be that *I mean* tends to be represented as two separate lexical items in the learners' mind while *I mean* in the native speakers' mind is more treated as a whole unit. For instance, the Chinese learners tend to preface their answers to questions like 'what do you mean' or 'you mean X?' with *I mean* while the native speakers tend to give their answers to either question without *I mean*. Examples (5.1) and (5.2), (5.3) and (5.4) show how the Chinese learners and native speakers answer the question 'you mean X?' and 'what do you mean?' respectively.

(5.1)

A But you know, he asked me for my advice, I think he should go. It's a good chance ... good opportunity. He can learn ... he can learn everything ... eh ... not everything ... but he can learn a lot of things ... which can not be learnt <learn> in here ...

B In China ... ***you mean***?

A ***I mean*** in China.

(01-099-16-SECCL)

(5.2)

188C Oh with Ramsey ***you mean***

189A Yeah

(SLA-048-ICE-GB)

(5.3)

A Some, some students still, still feel very puzzled about how to, about social contact.'

B Social contact?

A Yes.

B What did ***you mean***?

A Eh ...***I mean*** some students when they are about study they are interested by others. Because others want to.

(00-074-26-SECCL)

(5.4)

1A Uhm <,> you can break into the pears if you want to or have a piece of choccy

2A You 've had plenty of veggies <,,>

3B A piece of choccy

4B What do ***you mean***

5A Well you bought some and I bought some

(SLA-023-ICE-GB)

Even when *I mean* is used in non-pragmatic cases in the native speaker data, it tends to occur in fixed expressions such as *you see what I mean*, *you know what I mean* and *that's what I mean*. There are only a very small number of cases like "I think *I mean* that everybody did" (ICE-GB-014-126) where *I mean* is used semantically. In contrast, there are very few cases where *I mean* is part of fixed expressions in the learners' data.

For the learners who are of intermediate level in this study, it seems that their language proficiency has just reached a developmental stage where the pragmatization process of *I mean* has not been completed because borderline cases

like example (5.5) occur in the learners' data while *I mean* in the native speaker data is clearly used either pragmatically or semantically. In example (5.5), *I mean* could be argued to play the role of reformulating but can also be seen as a case of semantic use of *I mean*.

(5.5)

A Well, in my opinion, I think uh the more uh you know the more available uh method for us to to gain our experience I mean for the preparation for our future careers is to touch more new things ***I mean*** to get in touch with more new things. I think the security social work is a good opportunity for us to to be a member of it and to get in touch with a lot of people and learn from them I mean do some meaningful and significant things for the society.

(06-001-23-SECCL)

The pragmatic cases of *I mean* in the learners' data of this study include not only borderline cases like example (5.5), but also cases of a deviant use in which *I mean* plays a pragmatic role but is followed by the complementizer *that*. The reason for including the deviant cases is because the complementizer seems to be optional in the learners' mind because it does not have any semantic meaning and there is no grammatical equivalent of it in Chinese EFL learners' L1. Therefore, the learners' production of *I mean* as a pragmatic marker with or without the complementizer *that* might just be a random choice. Presumably, adding *that* to *I mean* may buy the speaker a bit more time to organize his/her online communication. Extracts (5.6) and (5.7) are two examples of disagreeing, one without *that* and the other one with *that*. Speakers in both example(5.6) and (5.7) use the same strategy to disagree with the hearer by admitting that what has been said by the hearer is right before giving his/her own opinion prefaced by *but* and *I mean*.

(5.6)

B Yeah, but, er, to gain some money to support your tuition. I think we can, maybe you can study hard and to gain the scholarship. I think that is also a good way.

A Er. What you said I think is right, but ***I mean*** just, er, er, for parents they just have responsibility to pay tuition for us, as our Chinese, China is a very traditional country and everyone just says that parents have responsibility.

(04-199-08-SECCL)

(5.7)

A But the main task for student is study, right?

B Yes, but I... ***I mean*** **that** in my... in our spare time and they also will practise our work ability.

(04-199-18-SECCL)

5.1.2 Distributions of the pragmatic functions played by *I mean* in the non-native speaker data and the native speaker data

Table 5.2 shows the distributions of the pragmatic functions of *I mean* in the British speakers' data and the Chinese EFL learners' data followed by the chi-square result of the distributions of the three main categories of *I mean*.

Table 5.2 Distributions of the pragmatic functions of *I mean* in the British speakers' data and the Chinese EFL learners' data

	Pragmatic functions of *I mean*	The British speakers (686)		The Chinese EFL learners (82)	
Hearer-oriented	Assumption-correction	73	10.6%	3	3.7%
	Cause	8	1.2%	1	1.2%
	Exemplification	48	7.0%	1	1.2%
	Explicitness	42	6.1%	13	15.9%
	Reason	24	3.5%	1	1.2%
	Reformulation	16	2.3%	9	11%
	Result	2	0.3%	0	0.0%
	Quotative	2	0.3%	0	0.0%
	Summarization	10	1.5%	2	2.4%
	Total	225	32.8%	30	36.6%
Interactant-relationship-oriented	Softener of FTA	39	5.7%	12	14.6%
	Justification	191	27.8%	14	17.1%
	Interactional repair	22	3.2%	3	3.7%
	Conclusion	13	1.9%	4	4.9%
	Total	265	38.6%	33	40.2%
Speaker-oriented	Hesitation marker	9	1.3%	5	6.1%
	Restart	90	13.1%	7	8.5%
	Resumption	14	2.0%	0	0.0%
	Indicating speaker attitude	51	7.4%	0	0.0%
	Transactional repair	32	4.7%	7	8.5%
	Total	196	28.5%	19	23.2%

	The three main categories			
	hearer-oriented	interactant-relationship-oriented	speaker-oriented	Total
The British speakers' data	225(32.8%)	265(38.6%)	196(28.5%)	686
The Chinese EFL learners' data	30(36.6%)	33(40.2%)	19(23.2%)	82
Total	255	298	209	768
X^2=1.129, df=2, p>0.56				

As shown by Table 5.2, both groups display a roughly even distribution in terms of the three main categories and the chi square test indicates that the difference between them is not significant. In both corpora, the category of interactant-relationship-oriented has the highest frequency and the category of speaker-oriented has the lowest frequency. They also show similar frequencies in subfunctions such as cause and interactional repair.

However, there are more differences between the native speakers and the learners regarding the frequencies of individual subfunctions. In the hearer-oriented category, the native speakers show much higher frequencies in correcting potential unintended interpretations and giving examples while the learners use much more of their *I mean*s to give more details and reformulate messages. Low frequency subfunctions, result and quotative, in the native speaker data are absent from the learners' function list.

In the interactant-relationship-oriented category, the British speakers use 10% more of their *I mean*s to justify what has just been said than do the Chinese EFL learners although justification shows the highest frequency in both corpora. But the learners use more *I mean*s in softener of FTA and conclusion.

In the hearer-oriented category, the biggest gap between the two groups is found in indicating speaker attitude, which is rather frequently used in the native speaker data but absent from the learners' data. The other subfunction that the native speakers use

more often is restart while hesitation marker and transactional repair are preferred by the learners. The low frequency subfunction of resumption in the native speaker data is absent from the non-native speaker data.

Although there are obvious differences between these two groups in terms of distributions of the individual subfunctions, the question that whether the difference between them is significant or not cannot be answered here because chi-square test cannot be used due to the low frequencies of subfunctions such as result, quotative, and indicating speaker attitude.

To sum up, as compared to *I mean* in the native speaker data, *I mean* is markedly under-represented in the Chinese EFL learners' data. In addition, *I mean* in the non-native speaker data is significantly less pragmatized. As to the specific functions of *I mean*, there is no significant difference between the two groups in terms of the distribution of the three main categories. *I mean* in the non-native speaker data occurs in slightly more restricted contexts because only 14 out of 18 subfunctions of *I mean* identified in the native speaker data are used by the Chinese EFL learners. However, there are quite a few subfunctions of *I mean* which are either over-represented or under-represented in the non-native speaker data. The subfunctions that are absent in the non-native speaker data are mainly the low frequency subfunctions in the native speaker data except 'indicating speaker attitude', which is frequently used.

5.1.3 Distributions of turn positions of *I mean* in the non-native speaker data and the native speaker data

As explained in section 3.5, *I mean* occurs in three turn positions i.e. turn initial, turn medial and turn final. Table 5.3 shows the distribution of these three categories.

Table 5.3.Distributions of turn positions of *I mean* in the British speakers' data and the Chinese EFL learners' data

	turn position			
	turn initial	turn medial	turn final	Total
The British speakers' data	60(8.8%)	623(90.8%)	3(0.4%)	686
the Chinese EFL learners' data	8(9.8%)	71(86.6%)	3(3.7%)	82

As shown by Table 5.3, more similarities than differences are shown between the two groups. For instance, the preferred turn position for both groups is turn medial position. Similar proportion of *I mean*s are used in turn initial position and turn final position is least used in both corpora. Since the frequencies of turn final position are too low, turn initial position and turn final position are conflated as non-medial position so that a chi-square test of medial position and non medial position can be run instead. As can be seen from Table 5.4 below, the difference between the native speakers and the learners in terms of medial and non-medial positions is not significant.

Table 5.4 Distributions of medial and non-medial positions of *I mean* in the British speakers' data and the Chinese EFL learners' data

	turn position		
	Medial	non-medial	Total
The British speakers' data	623(90.8%)	63(9.2%)	686
The Chinese EFL learners' data	71(86.6%)	11(13.5%)	82
Total	686	82	768
X^2=1.506, df=1, p>0.21			

However, at a more delicate level of analysis, differences emerge. 42% (25 out of 60) turn initial *I mean*s in the British speakers' data co-occur with other linkers and they are predominately used as post-linkers as in example (5.8).

(5.8)

148B I 'm not

149B I 'm I 'm

150B I 'm I 'm not very happy with that because uh that is uh that is how the Russians powered uh the

151A *Well **I mean*** they're putting chunks of plutonium up there which in fact they are

152A then that plutonium is going to shatter and disintegrate over quite a large area of the country

(S1A-088-ICE-GB)

There are only four cases (out of 25) where *I mean* is used either as a pre-linker as shown in example (5.9) or as a medial-linker in example (5.10).

(5.9)

90B when you're working with someone whether they're disabled or able-bodied you're still <,> it's still new to work with them and to learn to trust them and learn to <,> sort of have contact with them <,>

91C Uhm

92B ***I mean*** I *I suppose* that is awkward but then that that is awkward anyway <,>

(S1A-002-ICE-GB)

(5.10)

157C What do they do

158C I don't know I mean

159B *Well* ***I mean*** *say* eighteenth century uhm relationships between men and women were very different because if it was a low class woman

(S1A-020-ICE-GB)

Among the co-occurring linkers, pragmatic marker *well* and marker of hesitation *uhm* are the most frequent ones. Others include *but*, *and*, *because* and *yeah*.

In contrast, nearly all *I mean*s appearing in turn initial position in the learners' data are on their own as in example (5.11).

(5.11)

A but you are college student, right? And You know a lot of things, a lot of knowledge, and so if you are a bellboy, and you can do a well a good bellboy. and other people maybe he just graduate from the primary school, he also can can do, a good job as a bellboy, so what is the difference between a college graduate and er, another kinds of people ?

B ***I mean*** college doesn't mean much in a job if you are a capable person, and

(05-023-11B-SECCL)

There is only one case (out of 8) where *I mean* is used as a post-linker as shown in example (5.12).

(5.12)

A ... eh ..., I see you mean we can leave something from them and learn to cooperate with others, right?

B Um, yeah.

A *But ... eh ...* ***I mean*** I still have another question, I am afraid if I cannot do as good as I was in senior school, I mean the score on others things ... eh ... I will

feel ... I will feel very ... unconfident. How can I adjust my attitude?

(00-058-07-SECCL)

In example (5.12), *I mean* occurs in the turn initial position but prefaced by "but" and "eh".

Although *I mean* very rarely appears in turn final position in both corpora, the Chinese learners have less restriction in this position because the frequency of *I mean* occurring in this position in the learners' data is about 9 times of that of *I mean* in the native speaker data (although still very low in raw frequency). Also, the turn final *I mean* plays different roles in the native speaker data and the learners' data. Two out of the three instances of turn final *I mean*s in the native speaker data are used to giving details as shown by example (5.13) where speaker A specifies what food she wants to ask for. One of them is used in transactional repair as shown by example (5.14) where speaker C tells the hearer what he/she meant to say is adult story-tale tapes rather than good story-tale tapes.

(5.13)

119A And uh <,> <.>sh she showed me some
120A Oh quite nice that <,>
121A Good mm
122A What was the grub like <,,>
123A In France ***I mean***
124B Well we had uhm <,>
125B The people we were staying with they <,,> uh cooked us a traditional Normandy dinner <,,>

(SLA-009-ICE-GB)

(5.14)

115C Thank God thank God must do the same must do the same this year and get you some good story-tale tapes and maybe some more adult story-tale tapes ***I mean***
116B <laugh>

(SLA-011-ICE-GB)

On the other hand, in the learners' data, two of the turn final *I mean*s are functioning as conclusion as shown by example (5.15) where the speakers are arguing if university students should take the job of being a bellboy with a good salary.

(5.15)

B Yeah, I know what you mean, you mean er the position maybe er er a not en a high one. You mean serve <So> as a bell boy (A: Yeah) does not may not take er take full use of your... knowledge <knowledeg> in the university. But in my opinion, you should en work <word> from the bottoms of you job then you may... en know all the setters <settlers> in the hotel. And in that case can you... just en you you can work as a higher position, ***I mean.***

A Do you know what how <do> they <you> work?

(05-010-14B-SECCL)

It seems that B thinks that with experience of working from the bottom as a bellboy, which is an opportunity to know how a hotel works, a university graduate can be promoted to a higher position although it is not clear what 'know all the setters <settlers> in the hotel' means here. So the message that 'you can work as a higher position' followed by *I mean* is the inference that the speaker makes from the condition that if you know very well about how a hotel works by being a bellboy.

One of them is used in reformulation as shown by example (5.16) where "train station" in the previous message is reformulated by its synonym "railway station".

(5.16)

B Yes, but if the terrorism can attack the plane, why can't they attack the train next time? And compare with a compare with the er the train, yeah, the plane can er has still have much er advantages. Like example is er a softer, you know, and when you get into the em station, it it it won't be as crowded as in a train station. Have you ever been to a train station? Railway station, ***I mean.***

A Yeah yeah I've been there. I've been enter it. But em you know that em in my point of view for my point of view, I think the terrorist to terrorist the railway station is less attractive than an airplane, you know that.

(03-160-25B-SECCL)

To sum up, *I mean* in both data sets shows rather similar features in terms of turn positions. For both groups turn medial position is the dominant position and turn final position is the least preferred although *I mean* in the non-native speaker data has less restriction in turn final position.

5.1.4 Distributions of proposition positions of *I mean* in the non-native speaker data and the native speaker data

As discussed in section 3.5,the three proposition positions that *I mean* occurs are proposition initial, proposition medial and proposition final. The investigation of proposition position is particularly helpful for depicting a more detailed picture of the turn medial positioning of *I mean* because it can appear in any of the three proposition positions while turn initial *I mean* would also be proposition initial and turn final *I mean* would also be proposition final. All cases of proposition final *I mean*s in the native speaker data turn out to be turn final but there is one case in the Chinese learners' data where the proposition final *I mean* appears in turn medial position, which is shown in example (5.17).

(5.17)

B Um ... yeah, you're right. Um and what what do you think about the high school graduates um-um go overseas for their college education?

A I know the-the education system abroad is better than in China. You know they're related to the British culture, Australia, ***I mean***. And the teating the teachers their teating method are more suitable I think for everyone to develop their own characters.

(01-100-31- SECCL)

Table 5.5 below shows the distribution of proposition positions of *I mean* in the British speakers' data and the Chinese learners' data. As can be seen from Table 5.5, proposition initial is the predominant position for both groups. But compared to the British speakers, the Chinese EFL learners show less restriction in the other two positions. However, whether the difference between them is significant or not can not be tested here because of the low frequencies in proposition medial and proposition final positions. The frequencies of *I mean* in proposition medial position and proposition final position are so low that the difference between the two groups in terms of initial position and non-initial position (i.e. the conflation of medial and final positions) cannot be tested either.

Table 5.5 Distributions of proposition positions of *I mean* in the British speakers' data and the Chinese EFL learners' data

	Proposition position			
	proposition initial	proposition medial	proposition final	Total
The British speakers' data	673(98.1%)	10(1.5%)	3(0.4%)	686
The Chinese EFL learners' data	74(90.2%)	4(4.9%)	4(4.9%)	82

Despite the fact that both groups show very high frequency in proposition initial position, one third of *I mean*s in the British speakers' data co-occur with other linkers while only one fifth of *I mean*s in the learners data co-occur with other linkers. In both data *I mean* in proposition initial position is mainly used as a post-linker and least as a medial-linker. But the co-occurring linkers in the British speakers' data cover a wider range of expressions than those in the Chinese learners' data. The linkers that co-occur with *I mean* in the British speakers' data can be classified into the following groups : pragmatic markers such as *well, yeah, like, you know* and *you see*; stance adverbs (Powell: 1992) such as *actually, basically, honestly, obviously, probably, really* and *certainly*; epistemic expressions like *I suppose, I think, I guess* and *I presume*; conjunctive items such as *cos, because, but, and, so* and *on the other hand*; hedges such as *in a sense, to some extent* and *more or less;* and marks of hesitation like *uhm*. Of these, the most frequent co-occurring linkers are *but, uhm, well, you know, and, I think, so, because, yeah, like, so* and *obviously*. The picture of the co-occurring linkers in the Chinese learners' data is much simpler. Unlike the British speakers, the Chinese learners use a very limited number of linkers. There are only 8 linkers found in their conversations. They are *but, er, you know, because, so, and, yeah* and *actually*. Among them *but, er,* and *you know* are the most frequent ones.

When *I mean* occurs in proposition medial position, it will either appear between constituents or within a constituent. In both corpora, *I mean* is found both between constituents and within a constituent although cases where *I mean* is inside a constituent are very rare. Example (5.18) is an instance where *I mean* occurs between

constituents while example (5.19) is a case of within a constituent.

(5.18)

278B Well <,> it was sent to her home and then it was sent ***I mean*** to her own home uh and she 'd moved from there of course and was in the nursing home and they sent it there

(S1A-007-ICE-GB)

(5.19)

A But it's really a chance for for me to have a ... a ... a ... ***I mean*** explanation.

(01-008-25-SECCL)

To sum up, like *I mean* in the native speaker data, *I mean* in the Chinese EFL learners' data predominantly occurs in proposition initial position. However, it enjoys more freedom in proposition medial and proposition final positions. In terms of co-occurring linkers, overall *I mean* in the non-native speaker data tends to be a lot more independent and co-occurs with a much smaller number of other linkers.

5.2 Uses of *you know* in the non-native speaker data

5.2.1 Overall frequencies of *you know* in the non-native speaker data and the native speaker data

Table 5.6 shows the overall frequencies of *you know* in the British speakers' data and the Chinese EFL learners' data, followed by the results of the chi-square test of the distribution of identified cases.

Table 5.6 Overall frequencies of *you know* in the British speakers' data and the Chinese EFL learners' data

	Identified			Unidentified	Total
	Pragmatic	Non-pragmatic	Total		
British data (ICE-GB)	587 (84.9%)	104 (15.1%)	691	132	823
Chinese EFL data (SECCL 1.0& 2.0)	1080 (91%)	110 (9%)	1190	60	1250

	Identified		
	Pragmatic	Non-pragmatic	Total
The British speakers' data	587	104	691
The Chinese EFL learners' data	1080	110	1190
Total	1667	214	1881
X^2=14.620, df=1, p<0.001			

As can be seen from Table 5.6, overall the Chinese EFL learners use more *you know* than do the British speakers. In addition, the chi-square test shows that the proportion of pragmatic uses in the Chinese EFL data is significantly higher. Romero Trillo (2002) reports a different statistic finding in comparing adult British and Spanish EFL data. In his data, not only do the Spanish learners use markedly less *you know* but their *you know*s are about 10% less pragmatized (the British data 90.9% and the Spanish data 81.8%, respectively). Müller's (2005) investigation of *you know* used by German advanced EFL learners shows a similar pattern to Romero Trillo's (2002) study. The German learners in her data use 75% less *you know*s than do the American native speakers.

However, it should be pointed out here that like *I mean* the pragmatic uses of *you know* in the learner's data also include cases of a deviant use where *you know* is used pragmatically but followed by the complementizer *that*. As discussed in section 5.5.1, the compelmentizer *that* seems to be just a random choice to the Chinese speakers as examples of the same pragmatic function either with or without *that* are present. In the following two examples, A is asking B's advice as to which of the two job offers should be taken. Both of them were coded as 'introducing a new topic' although in example (5.21) *you know* is followed by the complementizer *that* because as can be seen from (5.20) and (5.21), the two *you know*s are used in exactly the same context i.e. with *you know* the speaker moves from the phatic talk to the main theme of the talk.

(5.20)

A Hello, Mary.

B Hello.

A ***You know*** we have, I have the chance to work in the government or um ... in the joint venture. But I was confused to choose what one to choose. Can you give me some suggestions?

(99-035-19-SECCL 1.0&2.0)

(5.21)

A Eh ... Hi, Umily.

B Hi, Alice.

A Eh ... ***You know*** *that* it has already take me a long time to look for a job and I

have just been offer two, just at the same time.
B Oh, is that <that's> real?
(99-035-07-SECCL 1.0&2.0)

The deviant use of *I mean* and *you know* – i.e. both expressions are used pragmatically but followed by the complementizer *that* in this study – seems to be the first report of deviant or wrong use of pragmatic markers in L2 English because there have been no such cases of learners' errors reported in most of the previous studies (Hays, 1992; Nikula, 1996; Demirci & Kleiner, 1997; Romero Trillo, 2002; Fuller, 2003; Fung & Carter, 2007; Müller, 2005) on learners' use of pragmatic markers in L2 English. Fuller (2003: 203) explicitly points out that in her data "none of the individual instances of *y'know* could be labeled as 'incorrect'".

5.2.2 Distributions of the pragmatic functions played by *you know* in the non-native speaker data and the native speaker data

Table 5.7 shows the distributions of the pragmatic functions of *you know* in the British speakers' data and the Chinese EFL learners' data followed by the chi-square results of the distributions of the three main categories.

Table 5.7 Distributions of the pragmatic functions of *you know* in the British speakers' data and the Chinese EFL learners' data

	Pragmatic functions of *you know*	The British speakers (587)	The Chinese EFL learners (1080)
Hearer-oriented	Assumption correction	1(0.2%)	0(0%)
	Exemplification	40(6.8%)	47(4.4%)
	Explicitness	46(7.8.%)	15(1.4%)
	Reformulation	13(2.2%)	3(0.3%)
	Cause	7(1.2%)	28(2.6%)
	Reason	16(2.7%)	24(2.2%)
	Result	7(1.2%)	3(0.3%)
	Summarization	11(1.9%)	1(0.1%)
	Quotative	29(4.9%)	0(0.0%)
	Introducing background information	11(1.9%)	6(0.6%)
	Seeking confirmation	2(0.3%)	8(0.7%)
	Total	183(31.2%)	135 (12.5%)
Interactant-relationship-oriented	Softener of FTA	13(2.2%)	368(34.1%)
	Interactional repair	2(0.3%)	0(0.0%)
	Justification	41(7.0%)	295(27.3%)
	Conclusion	10(1.7%)	14(1.3%)
	Indicating marked expression	48(8.2%)	2(0.2%)
	Indicating the most likely event	19(3.2%)	2(0.2%)
	Indicating the unspoken message to be completed by the hearer	18(3.1%)	0(0.0%)
	Indicating the coming message is meant to be evaluated	67(11.4%)	44(4.1%)
	Total	218(37.1%)	725(67.1%)
Speaker-oriented	Indicating speaker attitude	76(12.9%)	82(7.6%)
	Hesitation marker	44(7.5%)	33(3.1%)
	Restart	39(6.6%)	32(3.0%)
	Approximator	13(2.2%)	0(0.0%)
	Introducing a new topic	14(2.4%)	73(6.8%)
	Total	186(31.7%)	220(20.4%)

	The three main categories			
	hearer-oriented	interactant-relationship-oriented	speaker-oriented	Total
The British speakers' data	183	218	186	587
The Chinese EFL learners' data	136	752	220	1108
Total	319	970	406	1695
X^2= 158.588, df=2, p<0.00l				

As shown by Table 5.7, there are obvious differences between the native speakers and the learners in terms of the distributions of individual functions and the results of

the chi-square test show that there is a significant difference between them in terms of the distribution of the three main categories.

First of all, *you know* in the learners' data plays a narrower range of pragmatic functions. Instances of quotative, 'indicating the unspoken message to be completed by the hearer', approximator, assumption correction and interactional repair have not been found in the learners' data.

Secondly, the Chinese learners use the majority of their *you know*s in a very limited number of subfunctions while the subfunctions in the native speaker data are much more evenly distributed. Strikingly high proportions of *you know*s are used to soften face-threatening acts (34.1%) and explain why something has been said (27.3%) by the Chinese students. There are only two subfunctions 'indicating speaker attitude' (7.6%) and 'introducing a new topic' (6.8%), which are moderately used. The remaining subfunctions all show very low frequencies. In contrast, the most frequent subfunctions 'indicating speaker attitude' and 'indicating the coming message is meant to be evaluated' in the native speaker data only account for 12.9% and 11.4% respectively. There are a number of subfunctions such as 'indicating marked expression' (8.1%), explicitness (7.8%), hesitation marker (7.5%), justification (7.0%) and exemplification (6.8%), which show moderate frequencies.

To sum up, *you know* in the Chinese EFL learners' data shows a very different pattern from *you know* in the British speakers' data. Firstly, *you know* is markedly over-represented and is significantly more pragmatized in the native speaker data. Secondly, there is a significant difference between the two groups in terms of the distribution of the specific functions. Thirdly, the Chinese EFL learners use *you know* in more restricted contexts. 5 (out of 26) subfunctions are absent from the learners' function list. Finally, the Chinese EFL learners tend to use majority of their *you know*s in a very limited number of individual functions while the subfunctions of *you know* in the native speaker data are more evenly distributed. Softener of FTA and

justification account for as high as 34.1% and 27.3% respectively in the non-native speaker data while the most frequent subfunction in the native speaker data, 'indicating speaker attitude', only accounts for 12.9%.

5.2.3 Distributions of turn positions of *you know* in non-native speaker data and the native speaker data

Like *I mean*, *you know* also occurs in turn initial, turn medial and turn final positions. Table 5.8 below shows the distribution of these three positions in the two data sets, followed by the chi-square results of the distributions of the three positions.

Table 5.8 Distributions of turn positions of *you know* in the British speakers' data and the Chinese EFL learners' data

	turn positions			
	Turn initial	turn medial	turn final	Total
The British speakers' data	20(3.4%)	525(89.4%)	42(7.2%)	587
The Chinese EFL learners' data	266(24.6%)	753(69.7%)	61(5.6%)	1080
Total	286	1278	103	1667
X^2=120.516, df=2, p<0.00l				

As can be seen from Table 5.8 there are some similarities between the native speaker data and non-native speaker data. Firstly, both groups show their preference in turn medial position. Secondly, both groups use similar proportion of *you know*s in turn final position. In addition, the comparison of co-occurring linkers of turn final *you know*s in the two data sets shows that turn final *you know*s in both groups tend not to work with other linkers. The only exceptions are the 4 cases in the native speaker data where *you know* is used to indicate the unspoken message to be completed by the hearer. All these 4 turn final *you know*s co-occur with *so*. Example (5.22) is such an example.

(5.22)

288B well look the thing is that it's I can tell you now it's highly unlikely to take more than fifty pounds

289B We're having a few bottles but that's all

290B But they're not an alcoholic lot *so* ***you know***

291A More than fifty quid yeah

(S1A-008-ICE-GB)

However, the results of the test show that there is a significant difference between the two groups. According to Table 5.8 there is a substantial gap between the two data sets in turn initial position. Nearly one in four *you know*s in the learners' data is turn initial while in the native speaker data turn initial *you know* only accounts for 3.4%. The higher proportion of turn initial *you know* in the learners' data is mainly caused by the high frequency of the subfunction of softener of FTA where 161 out of 368 *you know*s are turn initial.

In addition, as high as 70% (188 out of 266) of turn initial *you know* in the non-native speaker data co-occurs with other linkers compared to 30% (6 out of 20) in the native speaker data, which is, again, contributed by the learners' frequent use of softener of FTA where *you know* is predominantly used as a post-linker co-occurring with *but*. Extract (5.23) is an example of this typical use. In example (5.23), A, a first year university student, is asking B's advice about how to lead a better university life. With *but* and *you know*, A softens his/her disagreement with what B has said about university life and appeals to B to understand why he/she disagrees.

(5.23)

A Oh, then you are a sophomore of our department, but I am a freshman, and I feel afraid <fraid> in this university, you know, I feel very lonely and I left all my friends behind.

B Oh, eh ... , don't ... don't be so discouraged. And you know the university life is very colorful and eh ... , wonderful life. And I hope that in the time to come, you can enjoy <enjoin> your college life.

A *But **you know*** in my senior middle school, all my teachers and classmates like me very much and they care for me. But here all the people around me are strangers. I feel much afraid and when my ... when I lay on my bed, I just think of my parents, think of my warm home.

(00-065-14 SECCL 1.0& 2.0)

The turn initial *you know* in the learners' data also co-occurs with a wider range of linkers than does that in the native speaker data. Apart from the most frequent linker of *but*, markers of hesitation such as *um, eh, er* are also frequently used. Other moderately used linkers include *well, and, oh, because, I think* and *yeah*.

Co-occurring items such as *in my opinion*, *maybe*, *personally*, *on the other hand*, *firstly* and *first* are very infrequent. In contrast, there are only 6 (out of 20) turn initial *you know*s that have co-occurring linkers in the native speaker data, which include *well*, *right*, *oh*, *ok*, *I mean*, *so*, and *uh*.

The same gap between the two groups is found in turn medial position. The native speakers use 20% more of their y*ou know*s in turn medial position than do the Chinese learners despite the fact that this position is preferred by both of them.

To sum up, there is a significant difference between the two groups in terms of the three turn positions. Although both groups show the highest frequency in turn medial position and a similar frequency in turn final position, the Chinese EFL learners use 20% more of their *you know*s in turn initial position and 20% less of them in turn medial position.

5.2.4 Distributions of proposition positions of *you know* in the non-native speaker data and the native speaker data

Like *I mean*, *you know* also occurs in the three proposition positions of propositional initial, proposition medial and propositional final. Table 5.9 shows the distributions of proposition positions of *you know* in the two data sets, followed by the chi-square results of the distributions of the three positions. For both groups, proposition initial position shows the highest frequency and proposition medial position shows the lowest frequency.

However, the results of the chi-square test show that the two groups show significant difference in terms of the distribution of the three proposition positions. *You know* in the native speaker data is more evenly distributed over the three categories than that in the learners' data. So *you know* in the native speaker data enjoys more freedom as to which proposition position it can appear in, while *you know* in the learners' data is predominantly proposition initial and is particularly restricted in proposition medial

position.

Table 5.9 Distributions of proposition positions of *you know* in the British speakers' data and the Chinese EFL learners' data

	Proposition positions			
	proposition initial	Proposition medial	proposition final	Total
The British speakers' data	319(54.3%)	114(19.4%)	154(26.2%)	587
The Chinese EFL learners' data	913(84.5%)	38(4.5%)	129(11.9%)	1080
Total	1232	152	283	1667
X^2=198.130, df=2, p<0.001				

As mentioned earlier, proposition medial position can be further classified into two positions, namely 'between constituents' and 'within a constituent'. A comparison of these two positions will give a finer description of how the two groups differ in proposition medial position. Despite the fact that *you know* in both corpora prefers to appear between constituents, *you know* is much more evenly distributed over the two positions in the native speaker data than in the non-native speaker data. About 40% (43 out of 114) proposition medial *you know*s occur within a constituent in the native speakers' corpus while only 10% (4 out of 38) proposition medial *you know*s occur within a constituent in the learners' corpus. The following four extracts are examples of the two proposition medial positions from the Chinese learners' data and the British speakers' data respectively. Extract (5.24) is an example where *you know* appears inside the noun phrase 'English skills' while extract (5.25) is an example where *you know* occurs inside the noun phrase 'this image'.

(5.24)

B Oh, I don't think so, using, en, u, u, en, the, the famous <famars> hotel must have some forein, foreigners. You see, my, our major were English. There must many foreigners <foreigner>, you can practice your English with them. You can gain, en, en, gain many en, speaking or en, or other English ***you know*** skill, skills.

(05-015-08B-SECLL 1.0& 2.0)

(5.25)

248B And uhm I realise how heavy going they were <,> and partic it wasn't the right thing to be reading at all to her

249B And I had sort of had this ***you know*** image of how I remembered

(SLA-013-ICE-GB)

In both examples (5.26) and (5.27), *you know* appears between a copula and its complement.

(5.26)

A Oh, you know nowadays many er foreign companies they still pay attention... pay very... er... pay a kind of... er special attention on the students' ... er voluntary work experience. They will think that this kind of students they will willing to help others without payment and this quality is... ***you know***... kind of virtue, so this kind of foreign's <foreign> corporation still pay attention to the students' this kind of qualities.

(06-001-19A-SECLL 1.0 & 2.0)

(5.27)

259A Where whereabouts was it

260B It was just by the market <,> just down some steps

261B It was just ***you know*** the market in Cambridge it was yeah

(SLA-015-ICE-GB)

These two groups also show great difference in terms of the frequency of co-occurring linkers of *you know*. 44% (476 out of 1080) *you know*s are used with other linkers by the learners while only 22% (131 out of 590) *you know*s co-occur with other linkers in the native speaker data. *You know* as a medial-linker is least used by the native speakers while *you know* as a pre-linker is the least frequent in the learners' data. However, for both groups *you know* is most frequently used as a post-linker. The following three extracts from the Chinese learners' data are examples of *you know* as pre-linker, medial-linker and post-linker respectively.

(5.28)

B Yes. The study is the main task in college, but not the whole things. ***You know*** *because en* became en tutor or the other kinds of part-time job can accu accurate the experience. En and an helpful for them to en enter the society.

A I know em you mean you mean that we can go en non a lot from taking a part-time job.

(04-128-24B-SECLL 1.0& 2.0)

(5.29)

A yes, but the life is the most important element, don't you think so? I I I still is a little bit afraid about the possibility to meet across the air crash you know.

B oh yes, *but* ***you know*** *I think um* people to people choose to take plane um is polite than those who choose to train.

(03-168-02A-SECLL 1.0& 2.0)

(5.30)

B Perhaps you are right. *But after all*, ***you know*** it too really depends on the decision of their parents. But asked for me, I will try my best to convince Jim the fact that he <she> should finish his college education and wait.

(01-067-30-SECLL 1.0& 2.0)

The two groups also share most of the co-occurring linkers. Among them, *but, uhm, and, because, well, I think, so* and *oh* are the most frequent ones that co-occur with *you know* in both corpora. Less frequent items for both groups include *actually, maybe, ok, yeah,* and *just*. However, co-occurring linkers which play textual function such as *first* and *firstly* only appear in the learners' data. *Sort of*, a moderately frequent co-occurring linker in the native speaker data, is absent in the learners' data. *I mean*, the second frequent linker, next to the marker of hesitation *uhm* in the native speaker data rarely occurs in the learners' data.

To sum up, *you know* in the native speaker data enjoys much more freedom in proposition positions while *you know* in the Chinese learners' data is predominately proposition initial and very restricted in proposition medial position and very rarely occurs within a constituent. Although *you know*s in both data sets share most of the linkers, about 20% more *you know*s in the learners' data co-occur with other linkers than those in the native speaker data.

In this chapter, I have described how *I mean* and *you know* are used by the Chinese EFL learners as compared to the British speakers in terms of specific functions they play and their positioning. The features of the uses of *I mean* and *you know* in the learners' data as compared to the native speaker data are summarized as follows:

1) In terms of overall frequency, *I mean* is markedly under-represented and significantly less pragmatized in the learners' data. In contrast, *you know* is markedly over-represented in the Chinese EFL learners' data and is significantly more pragmatized.
2) In terms of number of subfunctions, both *I mean* and *you know* are used in more

restricted contexts, which echoes Fung & Carter's (2007) observation that native speakers tend to use pragmatic markers in a wider variety of contexts.

3) In terms of the distribution of the three main categories, hearer-oriented, interactant-relationship-oriented and speaker-oriented, *you know* in the Chinese learners' data shows significant difference from *you know* in the native speaker data while *I mean* in the non-native speaker data does not.
4) With regard to the distribution of individual functions, the learners' *you know* heavily depends on a very small number of subfunctions while the subfunctions of *I mean* in the learners' data are more evenly distributed.
5) As to turn positions, similar to the native speakers, the Chinese learners tend to use *I mean* in turn medial position although they show a less restriction on turn final position. Although the Chinese learners also prefer to use *you know* in turn medial position like the British speakers, they show higher frequency in turn initial position.
6) As to proposition positions, like *I mean* in the native speaker data, *I mean* in the Chinese EFL learners' data predominantly occurs in proposition initial position, but enjoys more freedom in proposition medial and proposition final positions. In contrast, *you know* in the Chinese learners' data is predominately proposition initial and very restricted in proposition medial position and very rarely occurs within a constituent.
7) As to the co-occurring linkers, overall *I mean* in the non-native speaker data tends to be a lot more independent and co-occurs with a much smaller number of other linkers than *I mean* in the native speaker data. In contrast, 20% more *you know*s in the learners' data co-occur with other linkers than those in the native speaker data.

Chapter Six

Accounting for the uses of *I mean* and *you know* in the non-native speaker data

This chapter will attempt to account for the characteristics of how *I mean* and *you know* are used by the Chinese EFL learners as compared to the British speakers. As pragmatic markers *I mean* and *you know* are very unlikely to be taught in the classroom, the accounting for their features in the Chinese learners' data will be approached from the perspective of second language acquisition (SLA). Since L2 acquisition is highly complex, it is not surprising that previous research in SLA shows that there are many variables that can affect learners' L2 performance. Among them, L1 influence and universal constraints and processes are two most widely studied variables. So section 6.1 will review major findings revealed by previous studies on how these two factors may influence L2 production. Section 6.2 will, then, discuss how to establish the role of L1 influences on the Chinese EFL learners' use of *I mean* and *you know*. Sections 6.3 and 6.4 will discuss the features of *I mean* and *you know* in the learners' data respectively. Finally, section 6.5 will summarize the major findings of this chapter.

6.1 Major findings of previous studies on how L1 influence and universal constraints and processes may affect L2 acquisition

L1 influence is one of the most widely studied variables in SLA. There is a good reason for assuming that L1 can play a role in L2 acquisition because as Ringbom (2007: 1) points out any learning activity, including language learning, is built upon prior knowledge. So when people learn a second language, their L1 knowledge naturally becomes the prior knowledge they can draw upon. Assumption of how L1 affects L2 acquisition in 1950s and 1960 is rather straightforward. Researchers believed that differences between L1 and L2 should be responsible for L2 learning difficulties and a systematic comparison between L1 and L2 would be able to predict what errors will be made by L2 learners with a particular L1 background. However, subsequent research shows a much more complicated picture of how L1 can affect

L2 acquisition.

Firstly, it turns out that differences between L1 and L2 may not lead to L1-induced errors. Hyltenstam's (1977) investigation of negation of 160 beginners who learn Swedish as a second language shows that they show the same patterns of negation regardless of their L1 backgrounds. In the same time, similarities do not seem to be helpful. Some of the Turkish-speaking learners in his study show an initial preference for pre-verbal negation although both Turkish and Swedish are post-verbal negation languages. Instead, researchers have found that it is those L1-L2 similarities or differences perceived by L2 learners rather than those revealed by systematic L1-L2 comparisons done in Contrastive Analysis that have an impact on L1 transfer i.e. for L1 transfer to occur learners need to make "the judgment that something in the native language and something in the target language are similar" (Odlin, 2003: 454). Singleton's (1987) case study of a beginner learner of L2 French provides such evidence. His data suggest that his subject might be aware that the L2 solutions he came up with might not be right because one-third of his transfer errors co-occurred with indications of uncertainty such as hesitation and interrogative intonation. When he was asked to listen to his own errors, he commented "I knew that it probably wasn't right, but it was the nearest I could get to something that might be right" (Singleton, 1987: 335).

Secondly, comparison between L1 acquisition and L2 acquisition shows that there are universal developmental forces which constrain both L1 acquisition and L2 acquisition i.e.. L1 acquisition and L2 acquisition actually follow the same natural developmental paths. For instance, as Ortega (2009: 34) points out L1 English children will make errors such as *runned* or *the car was crashed also appear* before they arrive at the final competence. The same errors are also found in the speech and writing of adult learners of English with different L1 backgrounds.

Another universal force which can control both L1 and L2 is markedness, which refers to " a closed set of possibilities within a linguistic system, where the given possibilities rank from simplest and most frequent across languages of the world, or unmarked, to most complex and most rare, or marked" (Ortega, 2009: 37). For example, voiced stops are considered as marked as compared to voiceless stops. /b/, /d/ and /g/ in *tab*, *seed* and *bag* are voiced stops while /p/, /t/ and /k/ in *tap*, *seat*, and *back* are voiceless stops. According to Ortega (2009:37) voiced finals are considered marked because all languages have some voiceless stops while only some languages have both voiced and voiceless stops, and there is no language which has voiced stops without also having voiceless ones. She goes on to point out that children whose L1 has both voiced and voiceless stops will acquire the unmarked form, voiceless stops, first. The same acquisition order also applies to L2 learners.

Although L1 influence cannot override the natural developmental path of L2 acquisition, there is plenty of evidence suggesting that L1 influence can affect the rate of learners' progress of acquisition i.e. L1 influence can either delay or facilitate the progress of L1 acquisition. Spada and Lightbown's (1999) study provides evidence for the negative impact that L1 influence can impose on L2 acquisition. Their study shows that L1 French learners' acquisition of formation of English questions is delayed by the way that French questions are formed.

However, compared to the negative effects that L1 influence may have, the facilitative effects of L1 knowledge can easily be ignored because L1-induced errors against the norm of target language (TL) are easier to spot while L1 facilitating learning is much more difficult to notice (Ringbom, 2007: 2). Ringbom (2007) studies how English as a L2 is acquired in Finland, which he believes is an ideal place to investigate how language proximity i.e. similarities between L1 and L2 can facilitate L2 acquisition because there are two co-existing groups of learners i.e. L1 Finnish and L1 Swedish EFL learners in Finland. Swedish is very close to English because they both belong to Germanic languages while Finnish can be seen as a

language unrelated to English because Finnish is an agglutinative language. Ringbom's (2007) investigation shows that Swedes overall show great advantage over Finns in learning English. Wayland and Guion's (2004) experiment reports similar findings. In their experiment, both native speakers of English and native speakers of Chinese were taught Thai high and low tones. After training they were asked to identify these two tones. The result shows that the native speakers of Chinese outperformed the native speakers of English. The better performance of the Chinese speakers can be due to the fact that the Chinese language, like Thai, is also a tonal language while English is more distant to Thai.

To sum up, previous research in SLA shows that L1 influence can be a factor that affects L2 acquisition, but it cannot override universal developmental processes in acquisition of all human languages such as developmental sequences and markednesss. However, as Ortega (2009: 51) points out "to ascribe all interlanguage forms attested in a given data set to either the L1 or universal influence may be a futile enterprise. There are too many variables that affect the amount of L1 transfer that materializes for a given learner". Following this idea, this study assumes that L2 production can be seen as the result of the interaction of all potential factors and the importance of a certain factor varies from one L2 phenomenon to another. As will be seen in the following discussion, in addition to universal developmental processes and L1 influence, other potential variables that have been identified in this study include discourse types, learners' L2 proficiency and even the co-occurring linguistic elements of pragmatic markers.

6.2 Method of establishing the role of L1 influence on the Chinese EFL learners' uses of *I mean* and *you know*

To establish the role of L1 influences in L2 acquisition, one method is to compare the particular structure in the native language, the target language and the interlanguage (Odlin, 2003: 445). Selinker (1969) has successfully used this approach. By comparing a number of structures in Hebrew children's L2 English, their native

speech and native English data, Selinker finds that the interlanguage data resemble the word order of Hebrew significantly more than the word order of the target language. But for this study, such a comparison cannot be made because there is no parallel L1 data available. However, to have a rough idea of how the learners' L1 might affect their uses of *I mean* and *you know*, a general picture of how the Chinese equivalents of these two markers, *wode yisi shi* (my meaning is) and *ni zhidao* (you know), are used in the Chinese language is needed. Unfortunately, there are very few studies of *wode yisi shi* and *ni zhidao* available in Chinese studies. As a result, to obtain a crude picture of them I decided to carry out a small-scale study of them. Since I did not have access to face to face conversations in Chinese, a corpus of Chinese scripts of films and TV dramas from a script trading website, http://www.juben.cn, was used instead. Therefore, any findings of potential L1 influences here are only suggestive.

The Chinese corpus of scripts consists of about one million characters. It has to be pointed out here that the size of the Chinese corpus is counted on the basis of number of characters rather than words because unlike English there is no space between words in the Chinese language and a word can have one character or more. *Wode yisi shi* consists of five characters or three words while *ni zhidao* consists of three characters or two words. Therefore, instead of a direct comparison between the frequency of *I mean* in English and that of *wode yisi shi* in Chinese, frequencies of the other pragmatic marker in this study *you know* and its Chinese equivalent *ni zhidao* need to be used as references and vice versa.

The presentation of Chinese examples in this study follows the conventions set up by Li & Thompson (1981). Each Chinese sentence has two lines of English below. For the sake of easier reading, the Chinese sentence is transcribed in Pinyin (used in Mainland China) without tone markers rather than in Chinese characters. The first line of English translation consists of the most literal English equivalents of each Chinese element while the second line offers a proper English translation which

attempts to retain the meaning of the Chinese sentence as much as it can. Since this study only focuses on pragmatic functions of *wode yisi shi* and *ni zhidao,* only a couple of abbreviated terms are used in the transcription to represent those Chinese monosyllabic words which do not have direct English translations; these do not have semantic meanings; but carry grammatical meanings or pragmatic meanings. The abbreviated terms are NEG, CL and PRT, which stand for negation, classifier and particle respectively.

6.3 Accounting for the uses of *I mean* in the non-native speaker data

For the learners in this study, one developmental universal process that could influence the learners' acquisition of *I mean* is the congruence between the semantic meaning and the pragmatic function of *I mean* i.e. those pragmatic functions which can be easily derived from the semantic meaning of *I mean* should be easier for the learners to acquire because the more transparent the pragmatic meaning is the less difficult the learning task is.

The notion 'congruence' here is borrowed from studies done on the development of L2 morphology. Ortega (2009) cites studies done on L2 English to support the Aspect Hypothesis formulated by Andersen and Shirai (1996), which predicts that "the developmental pathway of emergence of tense and aspect will reflect prototypical pairings, that is, combinations where the semantics of the verb morphology is congruent with the semantics of the meaning of the verb to which the morphology is attached" (Ortega, 2009: 127). For instance, it is found that in L2 English, the combination of imperfective marking *-ing*, which has a prototypical durative meaning, with verbs depicting events that imply duration, such as *run*, *walk*, *sing* or *watch*, emerges first in interlanguage.

In order to show how congruence affects the learners' production of *I mean*, all the subfunctions of *I mean* are regrouped on the basis of the congruence between the pragmatic meaning and the semantic meaning of *I mean*. Table 6.1 below is the

reclassification, which is made over a continuum of congruence rather than a dichotomy because as Schiffrin (1987) points out the literal meaning of *I mean* not only influences its function of marking a speaker's upcoming modification of the meaning of his/her own prior talk but also is responsible for its use with self repair (Schiffrin, 1987: 296, 300).

Table 6.1 Reclassification of the pragmatic functions of *I mean* on the basis of congruence between pragmatic meaning and semantic meaning of *I mean*

			The British data (686)		The Chinese EFL data (82)	
More congruent: overall total			371	54%	59	72%
Repair	Message correction	Assumption correction	73	10.6%	3	3.7%
		Interactional repair	22	3.2%	3	3.7%
		Transactional repair	32	4.7%	7	8.5%
		Total Correction	127	18.5%	13	15.9%
	Discourse Editing	Hesitation marker	9	1.3%	5	6.1%
		Restart	90	13.1%	7	8.5%
		Total Editing	99	14.4%	12	14.6%
	Total Repair		226	32.9%	24	30.5%
Expansion of ideas		Explicitness	42	6.1%	13	15.9%
		Exemplification	48	7.0%	1	1.2%
		Reformulation	16	2.3%	9	11%
		Total Expansion	106	15.4%	23	28.1%
Softener of FTA			39	5.7%	12	14.6%
Less congruent : overall total			315	46%	23	28%
Because		Reason	24	3.5%	1	1.2%
		Cause	8	1.2%	1	1.2%
		Justification	191	27.8%	14	17.1%
			223	32.5%	16	19.5%
So/thus		Result	2	0.3%	0	0.0%
		Conclusion	13	1.9%	4	4.9%
		Summarization	10	1.5%	2	2.4%
			25	3.7%	6	7.3%
Indicating speaker attitude			51	7.4%	0	0.0
Quotative			2	0.3%	0	0.0%
Resumption			14	2.0%	0	0.0%

The reason that repair and expansion of ideas, which correspond to Brinton's category of appositional meanings are considered as more congruent is because they are the closest to the literal meaning of *I mean* (Brinton, 2007: 45). Softener of FTA is put under the same category because it can be seen as a kind of reformulation at the level of speech act rather than the content. The last three subfunctions, 'indicating speaker attitude', quotative and resumption, in the less congruent

category can be seen as the least congruent because they seem to have the least link with either modification of upcoming talk or repair.

The general picture of the acquisition of *I mean* seems to support the prediction that the more congruent pragmatic functions of *I mean* should be easier for the learners to acquire. As shown by Table 6.1 the learners use about 20% more of their *I mean*s in the more congruent category and 20% less in the less congruent category. The only four subfunctions that did not appear in the learners' data are all under the less congruent category.

The above discussion seems to suggest that the universal constraint i.e. the congruence between the pragmatic meaning and semantic meaning of *I mean* does have an impact on the Chinese EFL learners' acquisition of the specific functions of *I mean*. But it appears that the congruence cannot account for why *I mean* is markedly under-represented and significantly less pragmatized in the learners' data. The following comparison of *I mean* in the L1 English data, *I mean* in the L2 English data and *wode yisi shi* in the L1 Chinese data seems to indicate that the learners' L1 influence can be responsible for the Chinese EFL learners' avoidance of *I mean*.

Table 6.2 below shows frequencies of *I mean* and *you know* in the learners' conversations and frequencies of *wode yisi shi* and *ni zhidao* in the Chinese data. As can be seen from table 6.2, like *I mean* in the L2 English data, *wode yisi shi* shows a very low frequency in the L1 Chinese data as compared to *ni zhidao*.

Table 6.2 Frequencies of *I mean* and *you know* in the Chinese EFL learner's data and frequencies of *wode yisi shi* and *ni zhidao* in the native Chinese data

Pragmatic markers	The Chinese EFL learners' data	Chinese equivalents	The Chinese data
I mean	99	*wode yisi shi*	26
You know	1247	*Ni zhidao*	413

A closer look at how *wode yisi shi* is used in the L1 Chinese data seems to suggest that *wode yisi shi* is not typically used as a pragmatic marker in Chinese, which, again, could inhibit the learners from using *I mean* in their L2 English and also explain why *I mean* is less pragmatized in the learners' data. In order to get a better picture of how *wode yisi shi* is used, Webcorp was used to search the same script trading site, http//www.juben.cn, so as to generate more examples of *wode yisi shi*. With the help of Webcorp, another 27 instances of *wode yisi shi* were extracted. Among the 53 cases of *wode yisi shi*, 38 (72%) were found to be used semantically i.e. *wode yisi shi* is literally used to specify the speaker's meaning. Example (6.1) is a typical case where *wode yisi shi* comes in the response to an unintended interpretation of what the other person has just said.

(6.1)

A konglong! ni shi zhege sheshui de konglong!
Dinosaur! you be this society PRT dinosaur!
Dinosaur! You are the dinosaur in this society!

B ni zhe shi shenme yisi ?
You this be what meaning
What do you mean by this?

A Oh! ***Wode yisi shi***…xianzai xiang ni zhiyang de
Oh! my meaning be now like you such PRT
nühaizi zhengshi tai shangyou le
girl really too rare PRT
Oh! ***I mean*** people can hardly see girls like you nowadays!
(http//www.juben.cn)

Example (6.1) is a conversation between A, a young man and B, a young woman. By saying that A is a dinosaur, B means that A is an exceptionally good girl. But apparently B does not take it as a compliment but asks A to clarify why she is being called a dinosaur. So here *wode yisi shi* is used to specify the speaker's meaning.

Example (6.2) is the only example of *wode yisi shi* cited in Wang and Ge's study (2004) on Chinese discourse markers, the only one that I have found to discuss *wode yisi shi* in Chinese studies. As can be seen from example (6.2), *wode yisi shi*, again, is used to specify the speaker's meaning. Instead of directly answering whether he/she will go to the library, M3 uses a marker of hesitation to indicate that he/she does not

understand why M1 asks the question. With *wode yisi shi*, M1 explains why the question is asked.

(6.2)

M1 ni xiaowu qu tushuguan ma?
you afternoon go library PRT
Are you going to the library this afternoon?

M3 En…
Er…
Er…

M1 ***wode yisi shi***, ni ruguo qu tushuguan, jiu shunbian
My meaning be you if go library just by the way
bang wo xuejie yi xia ka shang de shu
help me renew one CL card on PRT book
I mean if you go to the library, I'd like to ask you to renew the books on the card for me
(Wang and Ge, 2004: 120)

However, example (6.2) is treated as a case of pragmatic use of *wode yisi shi* in their study where *wode yisi shi* seems to be regarded as a discourse marker in Chinese simply because its English equivalent is a discourse marker. There is no discussion as to how to distinguish the semantic use of *wode yisi shi* from its pragmatic use in their study.

The typical context of *wode yisi shi* shown by the above examples (6.1) and (6.2) actually reveals the key difference between *wode yisi shi* in L1 Chinese and *I mean* in L1 English i.e. *I mean* tends to occurs in self-initiated self-repair while *wode yisi shi* is more likely to occur in other-initiated self-repair. According to Schegloff et al. (1977) self-initiated self-repair is the most preferred and most common repair in conversation, which means in L1 English the speaker can use *I mean* whenever he/she needs to make a correction. In contrast, the speaker in L1 Chinese will not use *wode yisi shi* until he/she is told/hinted by the hearer that some clarification is needed. So this could be the reason for the low frequency of *wode yisi shi* in L1 Chinese, which in turn, can be seen as a factor that prevents *wode yisi shi* from further being pragmatized.

When *wode yisi shi* behaves more like a pragmatic marker, it is not used to respond to what the other speaker has said and it typically occurs in turn medial position and between two propositions. Example (6.3) is such a case where *wode yisi shi* is very similar to its English equivalent.

(6.3)

A zhe zhong shiqing zenme bu fen bici?
This kind thing how NEG distinguish each other
How can we share this?
Xiaoshang ke shi ge ren a !
Name PRT be CL man PRT
Xiaoshang is a man (as opposed to an object)!

B ni xianzai hai mei dui ta biaoshi, Xiaoshang ye mei biaoxian chu
You now still NEG to him express, Name either NEG show PRT
You haven't expressed to him yet, Xiaoshang hasn't shown who he prefers
dui shui you haogan.
to who have preference
either.
Wode yisi shi, women gongping jingzheng ba ?
My meaning be we fair compete PRT?
I mean let's have a fair competition, shall we?
(http//www.juben.cn)

Both A and B are female and they are friends. They fall in love with the same man Xiaoshang. B thinks it would be fair to have a competition between them because A has not told Xiaoshang that she likes him and Xiaoshang has not explicitly said who he prefers either. B's suggestion of having a competition is a conclusion drawn from her previous message. So this example can be coded as a case of conclusion.

Apart from the developmental force i.e. the congruence between semantic meaning and pragmatic meaning of *I mean* and the L1 influences, there are other potential variables which could contribute to the learners' pattern of *I mean*. As can be seen from Table 6.1 above, the higher use of the more congruent category is mainly contributed by the higher use of expansion of ideas and softener of FTA. The learners use a higher proportion of their *I means* in expanding their ideas than do the British speakers. But the subfunction of expanding ideas by giving further examples is substantially underproduced. One possible reason for the avoidance of this

subfunction could be because of the co-occurring linkers of *I mean*. When functioning in exemplification in the British data, *I mean* is found to work with other linkers in 44% of cases (21 out of 48) and the most frequent linkers are conjunctives such as *for example*, *say* and *like* whose main function is to introduce examples. Compared to *I mean*, these conjunctives can be seen as more transparent markers. As the previous studies (Hays, 1992; Nikula, 1996; Fung & Carter, 2007) of learners' use of pragmatic markers show that learners tend to use more transparent markers, the Chinese learners would be more likely to use *for example* rather than *I mean*, either on its own or in combination when they need to introduce examples. In contrast, when functioning in explicitness and reformulation, *I mean* in the native speaker data is more independent. There are only 21% (9 out of 42) cases of explicitness and 19% (3 out of 16) cases of reformulation where *I mean* co-occurs with other linkers. Therefore, the factor of co-occurring linkers would be less likely to discourage the learners from using *I mean* in the more congruent subfunctions of explicitness and reformulation.

The learners show a proportionally higher use of the subfunction of softener of FTA than the British speakers. Apart from the fact that this subfunction is rather close to the semantic meaning of *I mean*, its over-representation could also be due to the tasks that the Chinese EFL learners were asked to do in the oral exams. As can be seen from Table 6.3 below, which shows the list of tasks in the learners' data, the students were asked to perform three types of tasks, asking for advice, discussing or arguing with their partners. In most of the topics the students were made to take opposite views. For instance, in 2001 task if one student said he/she believed that high school graduates should study abroad then his/her partner would have to argue against it. As a result, the learners had to perform more FTAs such as disagreeing or turning down suggestions.

Although the learners and the British speakers use a similar percentage of *I mean*s in editing their on-line production, the learners' editing is rather evenly divided between

hesitation marker and restart while the native speakers edit their talk mainly through restart. One possible reason for the over-representation of hesitation marker could be relevant to the learners' proficiency level. As intermediate learners, apparently they need to hesitate more when speaking in a foreign language during an exam.

Table 6.3 Topics of role-plays in the Chinese EFL learners' data

Year	Topic of the role play
1996	Discuss with your partner if second-year students should apply take part-time jobs.
1997	Ask your friend for advice on whether to attend an English speech contest
1999	Ask your friend for advice on how to make decision between two job offers.
2000	Being a new student, you ask a second-year student for advice on how to live your university life
2001	You and your friend are discussing if high school graduates should study abroad.
2002	You argue with your friend on the issue that if the geology department should admit 30 male and 5 female while the overall scores of female students are higher.
2003	You ask your friend for advice on whether you should travel by air or by train.
2004	Discuss with your partner who should pay university tuition fee, students or their parents.
2005	Argue with your partner if university student should apply for a bellboy of a famous hotel, which offers 3000RMB salary.
2006	Discuss with your friend whether to work for money or do some voluntary work in the coming summer vacation.

Another reason for the over-representation *I mean* as a hesitation marker might be that *I mean* in the learners' data plays the role on its own because it rarely co-occurs with other linkers. In contrast, *I mean* in the native speaker data seems only playing a supporting role, because in over half of the cases it co-occurs with other linkers. Among them the most frequent one is *you know*, one of whose main functions is being a hesitation marker (for details see section 4.2.2.3.1).The following two examples are typical ones from the native speaker data and the learners' data which were quoted above as examples (4.40) and (5.19) (repeated here for convenience) respectively.

(4.40)

127A And the physical contact out there is is the only<unclear-words>problem of being isolated in your <,> wheelchair

128B Uhm

129B Uh uh in a in a chair s

130A Uhm

131B I think so

132B Yes

133B I mean I think uhm space is you know ***I mean*** <,> you know just the obstacles that you have in a room

(S1A-003-ICE-GB)

(5.19)

A But it's really a chance for for me to have a….a….a… ***I mean*** explanation.

(01-008-25-SECLL)

As mentioned earlier, all the four subfunctions that have not been transferred are under the less congruent category. But unlike result, quotative, and resumption, which all show very low frequency in the native speaker data, indicating speaker attitude is rather frequently used by the native speakers. One reason that learners' data fails to display an instance of this may be because the learners would not see any need for this subfunction due to the fact that *I mean* is simply used as an indicator and the speaker attitude is actually expressed by the upcoming message itself. As discussed in section 4.1.2.3.5, there are three types of speaker attitudes prefaced by *I mean*. The following are examples of the three speaker attitudes identified in the native speaker data, which were quoted above as example (4.43), (4.45), (4.46) respectively (repeated here for convenience).

(4.43)

105A And I've got masses of vases I could have lent her
106E Yeah we used to buy Mum a vase every year for her birthday
107A Yes
108E So we got uh inundated with them
109E We never use any of them hardly
110E ***I mean*** it's just amazing

(S1A-019-ICE-GB)

(4.45)

290D This looks delicious
291A Don't think makes that much difference <unclear-words>
292D Very on time <,,>
293B Well <,> one of us was
294C Mm
295D ***I mean*** <,> you were very on time
296D You were <,>
297D Absolutely on time

(S1A-022-ICE-GB)

(4.46)

136B I shall have the dry first liquid second
137B I I I shall try that
138B Have a go at this <,,> whi which has a kind of colour that I was born to appreciate

139B ***I mean*** I love it
(S1A-056-ICE-GB)

Example (4.43) is an instance of signaling speaker's evaluation or judgment, the most common speaker attitude identified in this study. Speaker E's attitude is clearly expressed by the adjective 'amazing' and the adverb 'just', which is "never semantically neutral but has an evaluative overlay" (Aijmer, 2002: 158). In example (4.45), the speaker attitude indicated by *I mean* is expressing emphasis. The emphasis is explicitly expressed by repetition of 'very on time' and adverb 'absolutely' in lines 295 and 296. So in both examples, the learners could perceive *I mean* as a superfluous expression.

Example (4.46) is an instance where *I mean* prefaces the attitude of sincerity, which is very rarely found in the native speaker data. *I mean* here can be translated as 'I'm serious when I say it'. So this is a subfunction where the pragmatic meaning of *I mean* derives from the sense of 'mean', 'to be serious about what the speaker is saying or writing'. But *yisi* in *wode yisi shi*, the Chinese equivalent of 'mean' does not cover this sense. In other words, to express the same meaning of 'I mean I love it' in Chinese, different lexical entities rather than *wode yishi* need to be used instead. So this might prevent the Chinese learners from using *I mean* to indicate sincerity in L2 English.

In sum, the above discussion of the uses of *I mean* in the non-native speaker data suggests that the pattern of the Chinese learners' *I mean* is shaped by the interaction of various factors. Firstly, it seems that the congruence between pragmatic meaning and semantic meaning of *I mean* plays the key role. Since most subfunctions of *I mean* are influenced by its semantic meaning, 14 out of 18 subfunctions have been successfully transferred, although at a much lower level of overall frequency than in the British speakers' data. The more congruent category is proportionally over-represented while the less congruent category is under-represented. The four subfunctions that have not emerged in the learners' conversations are all under the

less congruent category. Three of them can be seen as the least congruent ones because the connection between the pragmatic function and semantic meaning is almost opaque.

Secondly, the influence of the learners' L1 seems to be responsible for the marked overall under-representation of *I mean* in the learners' data because the small scale study of *wode yisi shi*, the Chinese equivalent of *I mean* shows that *wode yisi shi* is also very infrequent in the native Chinese data. In addition, the analysis of how it is used seems to suggest that it may not (yet) qualify as a pragmatic marker in the Chinese language because it shows very low frequency and it is used semantically in over 70% of cases.

Thirdly, there are also a few other variables which could account for under-representation or over-representation of certain subfunctions. For instance, the learners' L2 proficiency level could potentially lead to the over-representation of *I mean* as a hesitation marker. The fact that the learners' data is predominately argumentative could cause the over-representation of softener of FTA. The co-occurring linkers of *I mean* could be the reason why the learners avoid using *I mean* in exemplification.

6.3 Accounting for the uses of *you know* in the non-native speaker data

As argued in section 6.2 above the congruence between pragmatic meaning and semantic meaning of *I mean* can be seen as one major factor that shapes the learners' uses of *I mean*, but the accounting for the uses of *you know* will consider L1 influence as a major factor instead. This is because firstly it seems that this universal constraint may not apply to *you know* because as mentioned earlier the appealing function of *you know*, which can be easily derived from the semantic meaning of *you know*, can be extracted from all the specific pragmatic functions of *you know*. In other words, all the subfunctions of *you know* can be seen as equally congruent to the semantic meaning of *you know*. Secondly, it seems that L1 influence may play a

more important role in shaping the learners' uses of *you know* than those of *I mean* because in contrast to *I mean*, which cannot be literally translated into Chinese and whose Chinese equivalent may not even be considered as a pragmatic marker in the Chinese language, *nizhidao,* the Chinese equivalent and also the literal translation of *you know,* is a frequently used pragmatic marker in Chinese (Tao, 2003; Liu, 2005; Yang, 2007).

As will be seen in the following comparison between *you know* in the learners' L2 English and *ni zhidao* in the native Chinese data, striking similarities between them emerge, which may indicate that L1 influence does play a very important role in shaping the Chinese learners' uses of *you know*.

The first striking feature of *you know* in the learners' data revealed in section 5.2 is that it is markedly over-represented. Similarly, according to Table 6.2 (repeated here for convenience) below, *ni zhidao* is also very frequently used as compared to *wode yisi shi* in the native Chinese data.

Table 6.2 Frequencies of *I mean* and *you know* in the learner's data and frequencies of *wode yisi shi* and *ni zhidao* in the Chinese data

Pragmatic markers	The Chinese EFL learners' data	Chinese equivalents	The Chinese data
I mean	99	*wode yisi shi*	26
You know	1247	*Ni zhidao*	413

Secondly, *you know* is significantly more pragmatized in the Chinese EFL learners' data. Although as shown by Table 6.4, which shows the distribution of pragmatic cases and non-pragmatic cases of *ni zhidao* in the native Chinese data, *ni zhidao* is much less pragmatized than *you know* in the learners' L2 English or *you know* in the L1 English data (*nizhidao* 24%, *you know* in L2 English, 91% and *you know* in L1 English, 85%, respectively), a further analysis of non-pragmatic cases of *nizhidao* seems to suggest that the lower percentage of pragmatic use of *ni zhidao* could well be caused by discourse type.

Table 6.4 Distribution of pragmatic cases and non-pragmatic cases of *ni zhidao* in the native Chinese data

Pragmatic cases	Non-pragmatic cases			Total
99 (24%)	314(76%)			413
	Rhetorical question	Yes-no question and wh-question	Statements	
	75 (24%)	158(50%)	81(26%)	314

Rhetorical questions, like example (6.4) accounting for 24% in the semantic use of *nizhidao*, would be very unlikely to occur in the learners' conversation where the Chinese students are asked to argue with each other most of the time.

(6.4)

Ni zhidao wo duome xiang ni ma? ***Ni zhidao*** wo duome nanguo ma?
You know I how much miss you PAR? ***You know*** I how much upset PAR?
Do you know how much I miss you? Do you know how upset I am?
(http://www. juben.cn)

Again, half of the semantic cases of *nizhidao* are either yes-no questions like example (6.5) or (6.6) or wh-questions like example (6.7), which would be more frequent in more narrative discourse than the learners' conversation.

(6.5)

Ni zhidao tade guoqu ma?
You know his past PRT?
Do you know his past?
(http://www. juben.cn)

(6.6)

Zhe ***ni zhidao*** ma?
This ***you know*** PRT?
Do you know this?
(http://www. juben.cn)

(6.7)

Ni zhidao ta zhu naer ma?
You know he live where PRT?
Do you know where he lives?
(http://www. juben.cn)

Although Tao (2003: 298) claims that *nizhidao*, the Chinese equivalent of *you know*, is used as a pragmatic marker like *you know* in English, but is less pragmatized than *you know*, he does not provide the distribution of pragmatic cases and non-pragmatic cases in his study. Nor do other studies on *nizhidao* (Liu, 2005; Yang, 2007).

Another distinctive feature of how *you know* is used by the Chinese learners is that

vast majority of their *you know*s are used for just a small number of subfunctions, which actually roughly matches the picture of *nizhidao* described by Liu(2005) or the analysis of this study. As shown by Table 5.7 in section 5.2.2, *you know* in the learners' data is mainly used in softener of FTA, justification, indicating speaker attitude and introducing a new topic among the 24 identified subfunctions. Liu (2005) gives a rather detailed corpus-based account of *nizhidao* in her PhD thesis. Her data consists of recorded face-to-face conversations, telephone conversations, TV interviews, TV dramas and dialogues from novels. But her corpus is counted by size of the audio files rather than the number of characters. There are three main pragmatic functions of *nizhidao* identified in her study. The first function is to get the hearer's attention, which is similar to the subfunction of introducing a new topic in this study given her cited examples. Example (6.8) is one of them.

(6.8)

(Liu, Gao and Huang used to be postgraduate students of master in the same university. Although they studied different subjects, they were acquaintances because they had attended the same English class.)

Gao1 zenmeyang, du bo lei bu lei?
How read PhD tired NEG tired?
How about you? Is it tiring to do a PhD?

Liu1 lei, tebie lei, erqie yali tebie da
Tired, very tired, and pressure very big
Yes, I'm very tired and also under great pressure.

Gao2 ei ***ni zhidao*** ma, Huang si le
Oh, ***you know*** PRT, Huang die PRT
Oh, *you know* Huang has died.

Liu1 a? zenme keneng, ta bu shi kaoshang boshi le ma?
What? How possible he NEG be qualify PhD PRT PRT?
What? How come? He has just been admitted into the PhD program, hasn't he?

(Liu, 2005: 109, cited as example (135))

In example (6.8), Gao moves from asking about Liu's PhD study to a new topic about Huang who has just died.

In both her second and third functions, the message introduced by *nizhidao* serves the purpose of giving further explanation so that it would be easier for the hearer to understand the prior message. But under her second category, a distinction is made as

to whether the information marked by *nizhidao* is likely to be known by the hearer or not while in her third category the information marked by *nizhidao* is shared by the speaker and hearer. As mentioned earlier, the analysis in this study is more concerned about the potential logical relationship between the linked messages than whether the upcoming message is known by the hearer or shared by interactants or not. Given the examples cited under her last two categories, most *nizhidao*s are used in cause-effect relationship and can be classified under the subfunction of justification in this study. Example (6.9) and (6.10) are quoted in her second category. They would be coded as reason and justification respectively.

(6.9)

(Li is telling his wife the story of a group of people who he met on a plane.)

Li tamen haoxiang shi diyi ci zuo feiji , mei ge ren dou zi bei
They seem be first CL sit plane each CL person all self prepare
It seemed that it's the first time for them to take a flight. Each of them bought his own
de kuangquan shui, xiangchang shenme de.
PRT mineral water sausage so on PRT
mineral water, sausage and so on
….. zhe shi fuwuyuan songcan lai le,
….. This time steward serve meal PRT PRT
Then the stewards started to serve meal
tamen ke langbei le, yi ren nan yi ge kuanquanshui pingzi,
they very embarrassed PRT each person hold one CL mineral water bottle
they were very embarrassed, each of them held a bottle of mineral water
mei difang fang ***nizhidao*** ba.
no place put ***you know*** PRT
There was no place for them to put their bottles ***you know***.

In example (6.9), the message marked by *nizhiao* explains why they had to hold their mineral water bottles. So *nizhidao* in this example would be coded as reason.

(6.10)

(Wang is on a diet)

Child baba, wo ma zuo de shaoniurou ke haochi le, ni chang yi kou ba.
Dad, my mum cook PRT stewed beef very delicious PRT you try one CL PRT
Dad, the stewed beef cooked by mum is so delicious. Why not have a taste of it?

Wang qu qu qu bie duan zhe wan zai wo yan qian huang.
Go go go NEG hold PRT bowl in my eyes front move
Go away. Don't let me see the bowl.
wo kan zhe fang ***ni zhidao*** ba.
I see PRT annoyed ***you know*** PRT
It is very annoying to see it ***you know***.

ganjin qu chi qu, chiwang le jiu qu zuo gongke qu, kuai qu.
Hurry go eat PRT eat up PRT then go do homework go quick go
Go and eat it quickly. Do your homework after you finish eating. Hurry up.

Nizhidao in this example is used to justify why Wang told his child to leave him alone.

Extract (6.11) is an example from her third category and can be recoded as justification.

(6.11)
(At a teacher's office)

Teacher wo renwei shouxian yao zuo de shi, xiang ba zhe liang haizi chaisan,
I think first should do PRT be first PRT this two kid separate
I think what we should do first is to separate these two kids
bu xu tamen gao dao yiqi
NEG allow them play PRT together
and stop them from playing together.

Parent wo hen ganxie xuexiao de peihe, dashi yao wo xiaminglin
I very thank school PRT cooperation, but ask me give an order
I am very grateful for the cooperation of the school, but if you ask me to stop
bu xu tamen jiezhu, zuihao hai you yixie shishizaizaide zhengming
Neg allow them contact best PRT have some real proof
them from seeing each other, you'd better have some convincing proof that
ta gen Tiejun hun zaiyiqi queshi mei haochu de zhengju
he with Tiejun play together indeed NEG advantage PRT proof
can prove playing with Tiejun is, indeed, bad.
nizhidao xianzai de haizi, ni nao bu chu youshuofuli de dongxi
you know now PRT kid, you produce NEG PRT convincing PRT stuff
You know nowadays if you can't give convincing evidence,
ta jiu bu ting nide
he PRT NEG listen your
kids wouldn't listen to you at all.

The message followed by *nizhidao* in this example explains why the parent does not think that he/she can easily stop his/her son from playing with Tiejun unless there is some convincing proof to prove that playing with Tiejun is bad for his/her son. So *nizhidao* here would be coded as justification.

To sum up, given the examples cited in her three categories, it seems that *nizhidao* in her data typically occurs in contexts where it is used to justify what has just been said

or introduce a new topic.

As can be seen from Table 6.5 below, which shows the distribution of pragmatic functions of *nizhidao* in the corpus of scripts in this study, the two main functions of *nizhidao* identified in Liu's study (2005) are also among the frequently used subfunctions identified in this study. The most frequent subfunctions of *nizhidao*, justification, 'indicating speaker attitude', 'introducing a new topic' and softener of FTA, are also the most frequently used subfunctions of *you know* in the learners' data (see Table 5.7 in section 5.2.2 above). The conflation of these four subfunctions accounts for a strikingly high frequency in both the L1 Chinese data and the L2 English data (*nizhidao*, 70%, and *you know*, 75%, respectively).

Table 6.5 Distribution of subfunctions of *nizhidao* identified in the native Chinese data

Subfunctions (99)	Frequency
Justification	32 (32%)
Indicating speaker attitude	15 (15%)
Introducing a new topic	12 (12%)
Softener of FTA	10 (10%)
Indicating the coming message is meant to be evaluated	8 (8%)
Indicating marked expression	6 (6%)
Reason	5 (5%)
Exemplification	5 (5%)
Cause	4 (4%)
Result	2 (2%)

Finally, *you know* in the L2 English data and *ni zhidao* in the native Chinese data even show some similarities in terms of their turn positions and proposition positions. Table 6.6 shows the distributions of turn positions of *you know* in the learners' data and *nizhidao* in the native Chinese data. As shown in Table 6.6, both *you know* and *nizhidao* show the highest frequency in turn medial position and lowest frequency in turn final position.

Table 6.6 Distribution of turn positions of *you know* in the Chinese EFL learners' data and *nizhidao* in the native Chinese data

	Turn initial	Turn medial	Turn final	Total
The Chinese EFL data	266 (24.6%)	753(69.7%)	61(5.6%)	1080
The native Chinese data	16 (16.2%)	70 (70.7%)	13(13.1%)	99

With regard to proposition position *you know* and *nizhidao* also show a similar pattern because as can be seen from Table 6.7, which shows the distribution of proposition of *you know* in the learners' data and *ni zhidao* in the native Chinese data both of them show the highest frequency in proposition initial position and the lowest frequency in proposition medial position.

Table 6.7 Distribution of proposition of *you know* in the Chinese EFL learners' data and *nizhidao* in the native Chinese data

	Proposition initial	Proposition medial	Proposition final	Total
The Chinese EFL data	913(84.5%)	38 (4.5%)	129 (11.9%)	1080
The native Chinese data	72(72.7%)	3(3.0%)	24 (24.2%)	99

As discussed in section 3.5, proposition medial position can be further distinguished into two positions 'between constituents' and 'within a constituent'. *You know* in the learners' data not only rarely occurs in proposition medial but rarely occurs between a constituent, which matches the picture of proposition medial *nizhidao*. But it should be pointed out here that in the Chinese data when *nizhidao* occurs in proposition medial position, it only occurs between topic and comment, which corresponds to the position of 'between constituents' (Chinese is a topic-comment language as opposed to English, which is a subject-predicate language). So *nizhidao* would not occur within a constituent like *you know* does in English. This could discourage the learners from using *you know* in this position. Extract (6.12) is an example where *nizhidao* occurs between topic and comment i.e. between constituents.

(6.12)

Lingshan xiaozhang, xuexiao bu neng jiejue yixie ma?
headmaster school NEG can solve some PRT
Headmaster, can't our school give a bit financial help?

Jingzhao xuexiao de jingfei ***nizhidao*** jiu shi na dian bangong jinfei
school PRT funding ***you know*** just be that bit stationery funding
The funding of our school ***you know*** is just for stationery, cannot even
liang zhe qiche de you qian dou jiejue bu liao, shizai shi
even this car PRT petrol money PRT solve NEG PRT indeed be
cover the petrol of this car. Honestly…
(http://www. juben.cn)

nizhidao in this example occurs between the topic 'the funding of our school' and the

comment ‘is just for stationery and cannot even cover the petrol of this car’. The message marked by *nizhidao* is used by the headmaster to turn down Lingshan’s application for some financial support.

In addition to L 1 influence, other potential variables which might contribute to the over-representation or under-representation of certain individual functions of *you know* in the learners’ data also include discourse type and proficiency. The subfunction of softener of FTA is substantially over-represented in the L2 English. It accounts for 34% as compared to 10% in the native Chinese data or 2.2% in the British speakers’ data. Apart from the fact that *nizhidao* is frequently used in this subfunction in learners’ L1, another important contributing factor for the overproduction seems to be the tasks that the Chinese students were asked to perform. As mentioned earlier, one of the main tasks performed by the Chinese students was to argue with their partners about certain issues. Like *I mean* in the British speakers’ data, *you know* in the learners’ L2 is almost exclusively used to soften one fact FTA disagreeing. There are only a few cases where *you know* is used in speech acts of giving suggestion and refusing. The following three examples are instances of disagreeing, giving suggesting and refusing respectively.

(6.13)

A Really? I think it. It is not worthy to do, to do that. We can to do many other things contribute control our society. Not only to gain the money, you know.

B And, *but* ***you know*** every position servicing need people to the poor bill and that's.

A Yes, *but* ***you know***, I, I think this kind of job is for the lower educated people. you know. For example, the senior highschool <should> graduate or something likes <like> them.

(05-030-15A-SECLL 1.0 & 2.0)

In example (6.13) A and B are arguing if university students should apply for the job of being a bellboy in a five-star hotel. Following *but* and *you know*, B states his/her opinion that every job has its own value no matter how trivial it is, which disagrees with A’s idea that being a bellboy is worthless. In return, with *but* and *you know*, A argues that bellboy is only a job for people with less education than university students such as high school graduates. *You know* here not only makes disagreeing

less face-threatening but also appeals to the hearer to agree with the speaker.

(6.14)

B You should take it easy.

A Take it easy. But I ... Someone gives ... give me a lot of advice, but I can't be calm because you know it's a first time for me to attend and English speech contest.

B ***You know*** the first before it you should prepared <prepare> for it very well. Then you can have confidence.

A Yes. But I don't know how to prepare. I, I want to have a good speech. But I can't find some ... um ... satisfactory materials.

B But ***you know*** you should read some short stories, and read some materials about oral English.

(97-011-08- SECLL 1.0 & 2.0)

In example (6.14), A is asking for B's advice as to how to prepare an English speech contest. The two pieces of advice given by B are all prefaced by *you know*. *You know* here again play a dual role, softening the FTA of telling what people should do and appealing to the hearer to take the speaker's advice.

(6.15)

A While why not why not go with me. I think we can cooperate with with each other very well.

B Well, thank you for your invitation. But my parents also hope me to do some voluntary <volunteery> work, ***you know***. Er, and thank you. Thank you all the same. Er, well, I must be leaving off now or I'll be late for the class.

(06-004-03A-SECLL 1.0 & 2.0)

In example (6.15), the message marked by *you know* is used to turn down A's invitation indirectly. With *you know*, the refusal would sound less unpleasant to A and in the same time can appeal to A to understand why speaker B has to turn down the invitation.

Like softener of FTA, justification is another subfunction which is markedly over-represented in the learners' data. As mentioned above, the fact that *nizhidao* is frequently used in justification could encourage the overproduction of justification in the learners' L2 English. Another possible factor that could contribute to the over-representation might be again to do with the tasks that performed by the learners because apart from the task of arguing, where the speaker needs to justify

his/her opinion, another main task that the learners were asked to perform was to ask for advice, which is a speech act that needs to be justified. Extract (6.16) is such an example.

(6.16)

A Hi, Stone. I haven't see <seen> you for ages. How thing's going?

B Me too. I'm fine. How about you?

A Oh, it's really bad.

B What's the matter with you?

A Um ... because ***you know***, I ... I'm a freshman and ... eh ... it's the first time I step into university ... Eh ... <a click of tongue>. Everything is so different from my ... life in my middle school. So, ... mm ... I just feel in lost and can you give some ... suggestion to me?

B Um ... as a matter of fact ... maybe ... I've walked a long ... longer journey than you. So maybe I'd like to give you some advice <adwice> <advice>.

A <simultaneous with B> Oh, thank you very much.

(00-011-05-SECCL 1.0 & 2.0)

In example (6.16), A, a first-year-student asks B's advice as to how to handle university life. This example shows a rather frequent pattern of discourse which only occurs in the learners' data. Instead of answering B's question 'What's matter with you?' directly and then explaining why, A begins with giving the background information why he/she needs to ask for help before he/she actually performs the action of asking for suggestion.

As a frequently used subfunction in the learners' L1, 'introducing a new topic' is over-represented in the learners' L2 English. A more delicate analysis of 'introducing a new topic' in the learners' data reveals a very distinctive pattern. In 80% (60 out 73) cases, *you know* occurs in the beginning of the conversation and is used to direct the conversation from greetings to the main talk i.e. the task that the learners were asked to perform. Extract (6.17) is such an example.

(6.17)

A Hi.

B Hi.

A I wonder you are sophomore here.

B Yes.

A Ah, I'm Any. How do you do?

B How do you do? Nice to meet you.

A ***You know*** I'm a freshman in this university. I was very happy to enter the university

B Congratulations.

A Thank you. I heard that university life is very wonderful.

B Yeah.

A But to be honest, I'm a little bit lost and afraid of the coming life. Can you help me?

(00-082-19-SECCL 1.0 & 2.0)

As can be seen from example (6.17), both speakers are very nervous, which is understandable because as strangers they have to converse in a foreign language in an exam. After greeting each other, A attempts to direct the conversation to the topic they are supposed to talk about by asking B if he/she is a sophomore and expects B to ask back about him/her so that their conversation can move forward. But unfortunately, B gives up his/her floor immediately by saying 'yes'. Then A restarts the conversation by greeting and reattempts to lead the conversation onto the right track.

With *you know*, B has successfully made the transition from the phatic talk to the main talk. The reason for *you know* being used repeatedly in this context could be that *you know* in the learners' data not only can function in introducing a new topic but also play a role of helping the speaker get out of an awkward point of a conversation and then move on. This extra job done by *you know* in the learners' data seems to echo the way that *nizhiao* functions in introducing a new topic. Like *you know* in the L2 English, *nizhidao* is not only used in context where the topic is changed because the prior topic naturally comes to its end, but also used to switch to a new topic because of the awkwardness of the previous topic. Extract (6.18) is an example where *nizhidao* is used to introduce a new topic because the prior topic is too awkward to be continued.

(6.18)

Caomei Shitouge, ni zenme lian zijide shengri ye bu jide.
Stone brother you how including own birthday too NEG remember
Stone, how come you can't even remember your own birthday?

Stone ji ta gan shenme ? wode fumu zao bu zai le.
remember it do what ? my parents early NEG alive PRT

What's the point of remembering it? My parents died long time ago.
wo xianzai weiyi de qingren, zhi you yi ge meimei
I now only PRT relative only have one CL younger sister
Now the only relative I have is my younger sister.
ke ta ye chujia le
but she also marry PRT
but she has already got married.

(Caomei does not know what to say)

Caomei Stonege. ***Nizhidao*** ma qishi jintian dui wo lai shuo,
Stone brother. ***You know*** PRT actually today for me come say
Stone. ***You know*** for me today is actually also a very important day.
ye shi ge zhongyaode rizi
too be CL important day

In example (6.18), Caomei is a girl, who fancies Stone. When Stone tells why he can't even remember his own birthday, as the script suggests Caomei is shocked by the sad story and does not know what to say. Obviously their conversation comes to a point where it cannot be carried on. With *you know*, Caomei stops the topic about Stone and starts to talk about herself.

One important group of subfunctions of *you know* are either under-represented or absent in the learners' data. Their overall function is to make the conversation more interesting and engaging. They include indicating marked expression, indicating the most likely event, indicating the unspoken message to be completed by the hearer and indicating the coming message is meant to be evaluated. One reason for the under-representation or absence of them could be due to the fact that the learners' data is less narrative than the native speaker data because most of the time the learners have to think about how to win their argument rather than make their conversation more entertaining and engaging.

The under-representation of 'indicating marked language' can also be relevant to the learners' L2 proficiency. Since the learners in this study are only at intermediate level, they have not reached a point where they can appreciate stylistic differences in English. It would be very difficult for them produce examples like (4.78), which was quoted in section 4.2.2.2.5 (repeated here for convenience)

(4.78)

89A You mean even if you and Bernard had stayed together

90B Even if we had stayed together there might have <,> ***you know*** had endless hormones and glands problems

91A No

92A Yes you might

93A It's quite true <,>

94A She might have been that kind of teenager anyway

(S1A-031-ICE-GB)

In example (4.78), B uses two very technical words, hormones and glands. In the same context, the learners would be more likely to go for a much simpler expression 'sexual problems'.

However, the absence of the subfunction quotative which is moderately frequent in the British data seems very unlikely to be caused by the learners' proficiency level because the advanced German EFL learners in Müller's (2005:170) study did not use *you know* in quotative at all. Müller (2005) explains that the German learners apparently did not know about this function. For the Chinese learners in this study the absence of quotative of *you know* could be due to the fact that *ni zhidao* in their L1 does not have this subfunction.

To sum up, L1 influence seems to play a major role in the Chinese learners' uses of *you know* because there are striking similarities shown between the pattern of *you know* in the learners' data and *ni zhidao* in the native Chinese data. For instance, previous studies on *ni zhidao* and the small-scale study of this expression conducted in this study all show that *ni zhidao* is also a frequently used pragmatic marker in Chinese. This prior knowledge of L1 could prime the learners to overproduce *you know* in their L2 English. In addition, the analysis of *ni zhidao* shows that it is only used in very restricted contexts in Chinese, which echoes another striking feature of *you know* in the learners' data i.e. the majority of *you know*s are only used in a very small number of subfunctions among 24 identified subfunctions of *you know*. Even the positioning of *ni zhidao* seems to influence that of *you know* in the non-native speaker data because they show similar patterns in both turn positions and

proposition positions. In addition to L1 influence, other variables such as discourse type and proficiency may also play a role in the over-representation or under-representation of certain subfunctions in the learners' data. For instance, the overproduction of the subfunction softener of FTA could also be caused by the tasks given in the exams, many of which involved performance of the speech act of disagreeing. The under-representation of indicating marked expression could be due to the learners' proficiency because as intermediate learners they may not be able to appreciate the stylistic differences in English.

6.5 Summary

This chapter addressed the question why the Chinese EFL learners' uses of *I mean* and *you know* show those features revealed by chapter 5. To account for those features, the approach that L2 production should be viewed as the result of the interaction of all potential factors, among which universal developmental processes and L1 influence are the most widely considered, was adopted. In accounting for the features of the learners' uses of *I mean*, it was hypothesized that the learners' acquisition of *I mean* might be influenced by the congruence between the pragmatic meaning and semantic meaning of *I mean* because the more congruent the pragmatic meaning is the easier to be acquired. The result of the analysis seemed to support this assumption because the more congruent category was over-represented while the less congruent category was under-represented and all the subfunctions that the learners failed to acquire were under the less congruent category.

In addition to this universal constraint, the comparison between *wode yisi shi* in the native Chinese data and *I mean* in the non-native speaker data seemed to suggest that L1 influence could be responsible for the marked under-representation of *I mean* because *wode yisis shi* showed a very low frequency and was not typically used as a pragmatic marker in the Chinese data. Other variables such as discourse type, proficiency and co-occurring linkers of *I mean* also appeared to contribute to the over-representation or under-represenation of certain functions of *I mean* in the

learners' data. For instance the over-representation of softener of FTA could be because of the tasks that the Chinese students were asked to do in the exams while the over-representation of hesitation markers could by caused by the learner' proficiency level.

Since the congruence between the pragmatic meaning and semantic meaning of pragmatic markers does not seem to apply to *you know*, the accounting for the features of *you know* in the learners' data started with L1 influence. The comparison between *ni zhidao* in the Chinese data and *you know* in the L2 English data seemed to suggest that L1 influence played a very important role in the way that the Chinese EFL learners used *you know*. The analysis of *ni zhidao* showed that it was very frequently used as a pragmatic marker in the native Chinese data. So the learners' overproduction of *you know* in the learners' data could be encouraged by the frequent use of *ni zhidao* in their L1. In addition, *ni zhidao* in the native Chinese data occurred in much more restricted contexts than did *you know* in the British data, which might be the reason why most of *you know*s in the learners' data were only confined to just a few subfunctions.

In addition to L1 influence, other factors such as discourse type and proficiency, again, seemed to play a role in over-representation or under-representation of certain functions of *you know* in the learners' data. For instance like *I mean* the tasks that the students did in the exams could also be responsible for the over-representation of softener of FTA while proficiency might be the reason for the under-representation of 'indicating the marked expression'.

Chapter Seven

Conclusion

This chapter aims to summarize the major findings of this thesis and then make recommendations for further research. The summarization will be organized by reviewing how each of the four research questions (repeated here for convenience) proposed in Chapter 1 was addressed.

1) What pragmatic functions do *I mean* and *you know* play in L1 English conversation?
2) What are the similarities and differences between *I mean* and *you know* with regard to their pragmatic functions and why? What are the similarities and differences between them in terms of distributions and positioning?
3) What are the characteristics of the Chinese EFL learners' use of *I mean* as compared to the native speakers of English? What are the potential factors that could contribute to the similarities and differences between them and why?
4) What are the characteristics of the Chinese EFL learners' use of *you know* as compared to the native speakers of English? What are the potential factors that could contribute to the similarities and differences between them and why?

This chapter will be subdivided into 4 sections. Section 7.1 will outline the patterns of *I mean* and *you know* in the native speaker data chosen by this study. Section 7.2 will highlight the similarities and differences between the patterns of *I mean* and *you know* mentioned in section 7.1 and then the proposal as to what could account for those similarities and differences between them. Section 7.3 will present both the patterns of *I mean* and *you know* in the Chinese EFL learners' data as compared to their uses in the native speaker data and discuss possible factors which might play a role in shaping the features of *I mean* and *you know* in the non-native speaker data.

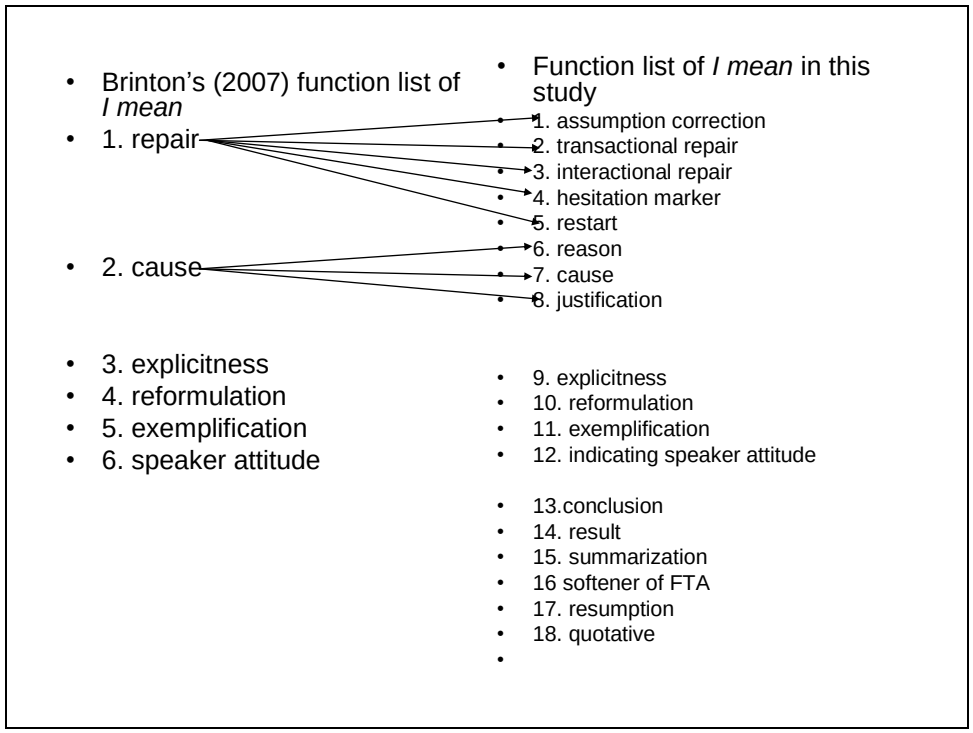
Finally, in section 7.4 suggestions for further research will be made.

7.1 Patterns of *I mean* and *you know* in the L1 English data

Given the fact the chosen sample of spoken data produced by British speakers in this study is not only bigger but also involves a wider range of topics, age groups and speakers than the data used in the previous studies (e.g. Goldberg 1980; Östman 1981; Schourup 1985; Holmes 1986; Schiffrin 1987; Müller 2005), a finer picture of two of the most evasive pragmatic markers, *I mean* and *you know*, than descriptions of them offered in previous research was hoped for. To achieve the best possible description of *I mean* and *you know* the following measures were taken in this study. Firstly, instead of being confined to one single model or theory, the identification of the pragmatic functions of the two markers in question was open to all plausible interpretations. Secondly, to ensure the consistency of data coding, a guiding principle of giving priority to coherence relations was followed in the process of setting up my own categories. Thirdly, a polysemous approach was adopted so that the analysis in this study not only zoomed in on the specific contexts where the markers occurred but also zoomed out to the general function or core meaning of each marker.

The above measures proved to be effective because a finer picture of the two markers than previous descriptions of them did emerge out of my analysis. For instance, the analysis of *I mean* in this study came up with a much longer function list than Brinton's (2007) corpus-based description of *I mean*, which was the most detailed account available when this study began. As shown by Table 7.1, which compares Brinton's (2007) function list and my function list of *I mean*, one group of the newly identified functions of *I mean* in my analysis were the result of further categorization of Brinton's (2007) categories.

Table 7.1 Function lists of *I mean* in Brinton's (2007) study and this study



Brinton's category of repair was further classified into five subtypes of repair – assumption correction, interactional repair, interactional repair, hesitation marker and restart – in my analysis. Another of her categories, cause, which is paraphrased as either 'because or 'I'm saying it because' in her analysis, was also further categorized in this study. Those *I mean*s which can be paraphrased as 'I'm saying it because' were coded as justification while those paraphrased as 'because' were labeled either cause or reason because of the different coherence relations shown by the two messages linked by *I mean*.

As can be seen from Table 7.1, the other group of newly identified functions of *I mean* on my list are not reported in Brinton's analysis. One moderately frequent subfunction in my data, softener of FTA, which accounted for 5.7% of the identified cases, does not appear on Brinton's function list. Another category which is absent in Brinton's analysis but showed rather moderate frequency (3.7%) was those cases where *I mean* was translated as 'so'. They were further grouped into three subgroups – result, summarization and conclusion. Other subfunctions such as quotative and

resumption which do not appear on Brinton's list showed very low frequencies in my data.

By following the same approach, my analysis of *you know* also came up with a much longer list than do the most recent studies (Erman, 2001; Müller, 2005) of *you know* available when this study started. Like *I mean*, some of the newly identified subfunctions of *you know* in this study are also the result of further categorization of functions of *you know* observed in previous research while other new subfunctions were identified only in my data because they were rather infrequent and therefore they may not have been sampled in the data used in previous studies. For example, subfunctions such as exemplification, explicitness and reformulation specify ways how *you know* introduces qualifying information (Holmes, 1986) or modification of previous discourse (Erman, 2001). The two subfunctions 'indicating the unspoken message to be completed by the hearer' and 'indicating the most likely event' revealed how the speaker claims common ground, which is claimed to be a role of *you know* by a few researchers (e.g. Östman, 1981; Schourup, 1985; Jucker & Smith, 1998). Holmes (1986) observes that the speaker uses *you know* to make his/her narration more interesting and entertaining. The newly identified subfunctions – 'indicating the marked expression' and 'indicating the coming message meant to be evaluated' – in this study showed how the speaker achieves his/her goal of retaining the hearer' interest in conversation. The subfunctions which are not reported in previous research but showed very low frequencies in my data included cause, reason, result, conclusion, assumption correction and interactional repair.

However, there are two subfunctions mentioned in Erman's (2001) and Müller's (2005) lists which are not included on my list. One is the role of turn management and the other is the subfunction of appealing. The exclusion of the former is because it is difficult to rule out other possible reasons which might lead to turn exchange (Schourup,1985). In addition, identification of this role involves prosodic

information Müller (2005) which was unavailable in my data. With regard to the latter, it is treated as the core meaning or general function rather than one of the subfunctions of *you know* because it could be extracted from all the subfunctions of *you know* identified in this study.

7.2 Comparison of *I mean* and *you know* in the L1 English data

Following the completion of the detailed account of *I mean* and *you know* a comparison between them was conducted. Unlike the previous comparisons (Östman, 1981; Schourup, 1985; Schiffrin, 1987; Fox Tree and Schrock, 2002), which are either too brief or too general, the comparison in this study was made both at the level of their general functions and at the level of their specific functions.

At the general level, the comparison shows that *I mean* and *you know* can be seen as pragmatically complementary as Schiffrin (1987:309) points out that *I mean* focuses on the speaker's own talk while *you know* is about the hearer's involvement. In addition, I proposed that under Brown and Levinson's (1987) Politeness Theory, the core functions of *I mean* and *you know* can also be seen as complementary. *I mean* can be used to serve the hearer's negative face because with *I mean* the speaker can diminish potential imposition on the hearer by suggesting 'this is what I mean, so feel free to disagree with me' while *you know* can be used to serve the hearer's positive face because with *you know* the speaker can appeal to the hearer's participation in his/her talk i.e. with *you know* the speaker can show his/her interest in the hearer.

The findings of the comparison at the level of specific functions supported the general complementary picture of *I mean* and *you know* presented above. For instance, on the one hand there was no significant difference between *I mean* and *you know* in terms of the distributions of the three broad categories, hearer-oriented, interactant-relationship-oriented and speaker-oriented, which suggests that *I mean*

and *you know* play rather similar functions in conversation. On the other hand, they shared about half of their subfunctions (16 out of 26) and the complementary general functions of *I mean* and *you know* provides a plausible explanation as to why some contexts allow both markers while other contexts only allow one of them.

To give a better picture of how *I mean* and *you know* relate to each other, their pragmatic functions are regrouped as shown by Table 7.2 below on the basis of how many contexts they share. As can be seen from Table 7.2, two main categories are set up on the basis of whether the contexts are shared by the two markers or not. Subfunctions in the category of shared contexts are further grouped into three subcategories – contexts where both markers show similar frequencies, predominantly-*I-mean* subfunctions and predominantly-*you-know* subfunctions. Under the category where only one marker occurs, two subgroups – *I-mean*-only subfunctions and *you-know*-only subfunctions – are set up.

From the perspective of the general or core meanings of *you know* and *I mean*, the shared contexts are actually those subfunctions where the speaker is allowed to choose whether to focus on his/her own talk or to invite the hearer to participate i.e. whether to serve the hearer's negative face or positive face. For instance, subfunctions such as exemplification, explicitness, reformulation, cause, reason and summarization where *I mean* and *you know* showed very similar frequencies can be seen as the contexts where the speaker achieves a balance between the hearer's need of negative face and positive face by equally sharing the job in conversation with the hearer. The occurrence of either *I mean you know* or *you know I mean* in all shared subfunctions seems to suggest that the speaker may be aware that such a balance is important because considering too much of the hearer's positive face could damage the hearer's negative face or vice versa.

Table 7.2 Classification of the pragmatic functions of *I mean* and *you know* on the basis of how much context they share in the British speakers' data

		Pragmatic functions
Shared contexts	Similarly frequent	Exemplification
		Explicitness
		Reformulation
		Cause
		Reason
		Summarization
		Conclusion
		Result
	Predominantly-*I- mean*	Assumption correction
		Softener of FTA
		Justification
		Interactional repair
		Restart
	Predominantly-*you-know*	Quotative
		Indicating speaker attitude
		Hesitation marker
One marker only	*I-mean*-only	Resumption
		Transactional repair
	You-know-only	Introducing background information
		Seeking confirmation
		Indicating marked expressions
		Indicating the most likely event
		Indicating unspoken message to be completed by the hearer
		Indicating the coming message to be evaluated
		Approximator
		Introducing a new topic

As to those shared contexts where one marker showed much higher frequency than the other, the reason could be that one of the needs related to face becomes the speaker's priority at the time of communication. For example, the frequency of *I mean* was nearly quadruple that of *you know* in the subfunction of justification. One reason for the preference for *I mean* to *you know* may be because the British speakers feel that the act of appealing to the hearer to agree has a very high risk of damaging the hearer's negative face. In contrast, in the subfunction of hesitation marker the frequency of *you know* was nearly five times that of *I mean*. Apparently, this is a

subfunction where the need for asking the hearer to share the work of searching right words or content is more urgent than the hearer' negative face.

As to the contexts where only one marker occurred, the general functions of *I mean* and *you know* also work very well. The *you-know*-only subfunctions are those where the hearer's involvement is essential while in the *I-mean*-only ones the hearer's participation is least needed or even impossible. For instance, the subfunction 'indicating the marked expression' is a *you-know*-only subfunction because the hearer's appreciation is the reason why the speaker produces the utterance. A typical example of *I-mean*-only subfunction is interactional repair because this subtype of repair involves correcting the wrong information provided by the speaker him/herself who knows what the correct information is.

7.3 Patterns of *I mean* and *you know* in the Chinese learners' L2 English

Unlike most of the previous studies (Hays, 1992; Nikula, 1996; Romero Trillo, 2002, He, 2002; Fung and Carter, 2007) on L2 learners' use of pragmatic markers, which aim to draw a general picture by reporting what pragmatic markers are used by learners and their frequencies, this study followed Müller's (2005) study, which aims to describe the use of learners' pragmatic markers at the level of specific functions by investigating what pragmatic functions are played by *I mean* and *you know* in the Chinese EFL learners' data as compared to the British speakers' data. But this study continued to look further into potential variables which might be responsible for the characteristics of the Chinese learners' use of *I mean* and *you know*.

This section will be further divided into two subsections, section 7.3.1 and section 7.3.2 , which will describe the Chinese learners' *I mean* and *you know* respectively.

7.3.1 Pattern of *I mean* in the Chinese learners' data

Compared to *I mean* in the British speakers' data, *I mean* was markedly

under-represented in the Chinese EFL learners' data, while 14 out of 18 subfunctions of *I mean* which appeared in the native speaker data were also used by the Chinese learners although most of them showed different frequencies i.e. they were either under-represented or over-represented in the learners' data.

The underproduction of *I mean* in the learners' data could be a L1-induced phenomenon because a small-scale study of the Chinese equivalent of *I mean*, *wo de yisi shi*, showed it was very infrequent in the native Chinese spoken data and its position was very restricted, which suggests it may not be qualified as a pragmatic marker in the Chinese language.

But the under-representation or over-representation of the specific functions of *I mean* seems more relevant to the degree of congruence between the pragmatic functions and semantic meaning of *I mean* because the more congruent category was over-represented while the less congruent was under-represented and the subfunctions which were absent in the learners' data were all under the less congruent category. However, in addition to the L1 influence and the developmental force i.e. congruence, there are other contributing factors such as discourse type, proficiency level and co-occurring linkers. For instance, the over-representation of hesitation marker under the more congruent category may also be caused by the learners' L2 proficiency because the intermediate learners in this study might need to hesitate more while conversing in a foreign language. Although exemplification is under the more congruent category, it was greatly under-represented in the Chinese learners' data. One reason for the learners' avoidance of this subfunction could be caused by the co-occurring linker of *I mean.* When *I mean* functions in exemplification, it tends to co-occur with more transparent markers such as *for example* and *things like* in the L1 English data. According to previous research on learners' use of pragmatic markers learners tend to use more transparent markers, which means the Chinese learners would more likely to use *for example* to *I mean*

when they need to introduce examples.

7.3.2. Pattern of *you know* in the Chinese learners' data

In contrast to *I mean*, *you know* was markedly over-represented in the Chinese EFL learners' data. Although *you know* in the learners' L2 English served most subfunctions (19 out of 24) identified in the L1 English data, it only focused on a very small number of subfunctions. The comparison between *you know* in the Chinese learners' data and the Chinese equivalent of *you know*, *ni zhidao*, in the Chinese spoken data seems to suggest that L1 influence could be the major contributor to the pattern of the uses of *you know* in the learners' data. Unlike *wo de yisi shi*, *ni zhidao* was much more frequent in the native Chinese spoken data chosen by this study. Previous studies (Tao, 2003; Liu 2005) of *ni zhidao* and my analysis of *ni zhidao* all show that *ni zhidao* plays a very limited number of subfunctions and the four most frequent subfunctions of *ni zhidao* – justification, indicating speaker attitude, introducing a new topic and softener of FTA – were also among the most frequently used subfunctions of *you know* in the learners' data. In addition, even the pattern of the positioning of *you know* in the learners' L2 English was similar that of *ni zhidao* in the L1 data.

Other variables such as language proficiency and discourse type also seem to play a role in the shaping of learners' use of *you know*. For instance, the subfunction of indicating a marked expression was markedly under-represented in the learners' L2 data but showed a rather moderate frequency in the L1 Chinese data. The reason for the underproduction of this subfunction in the learners' data could be because the intermediate learners have not reached a point where they could appreciate stylistic differences in English. The overproduction of softener of FTA also seems to be relevant to the tasks that the learners did in the exams.

7.4 Recommendations for further research

Although this study did achieve a finer description of how *I mean* and *you know* were used by the native speakers and a detailed account of the usage of these two markers in the non-native speaker data as it hoped, the main problem it had was the comparability of its chosen corpora. Therefore, to achieve a more accurate picture of Chinese EFL learners' use of these two markers, future study should look for a collection of data which resemble more closely the native English data chosen for this study. To better understand how the L1 would affect the acquisition of these two markers, a study of their equivalents by analyzing comparable L1 Chinese spoken data will also be needed. In addition, to know more about how universal developmental force can affect the path of the acquisition, an investigation of how they are acquired by children will be very helpful. Finally, a comparison between intermediate learners and advanced learners will help to answer the question whether the features of *I mean* and *you know* found in this study are features of intermediate learners or rather features of speakers whose L1 is Chinese.

References

Abraham, Werner. 1991. Discourse particles in German: how does their illocutive force come about?. In Werner Abraham (ed.), *Discourse particles. Descriptive and Theoretical Investigations on the logical, syntactic and pragmatic properties of discourse particles in German*, 203-252. Amsterdam: John Benjamins.

Aijmer, Karin. 2002. *English discourse particles. Evidence from a corpus.* Amsterdam: John Benjamins.

Aijmer, Karin & Anne-Marie Simon-Vandenbergen. 2004. A model and a methodology for the study of pragmatic markers: semantic field of expectation. *Journal of Pragmatics* 36. 1781-1805.

Andersen, Elaine S., Maquela, Brizuela, Beatrice, DuPuy & Laura, Gonnerman. 1999. Cross-linguistic evidence for the early acquisition of discourse markers as register variables. *Journal of Pragmatics* 31. 1339-1351.

Andersen, Gisle. 1997. *They gave us these yeah, and they like wanna see like how we talk and all that*. The use of *like* and other pragmatic markers in London teenage speech. In Kotsinas Ulla-Britt, Anna-Malin Karlsson & Anna-Brita StenstrÖm (eds.), *Ungdomsspårk i Norden*, 82-95. Stockholm: MINS.

Andersen, Gisle. 1998. The pragmatic marker *like* from a relevance-theoretical perspective. In Andreas H. Jucker & Yael Ziv (eds.), 147-170.

Andersen, Gisle. 2000. The role of the pragmatic marker *like* in utterance interpretation. In Gisle Andersen & Thorstein Fretheim (eds.), *Pragmatic markers and propositional attitude*, 17-38. Amsterdam: John Benjamins.

Andersen, Gisle. 2001. *Pragmatic markers and sociolinguistic variation: a relevance-theoretic approach to the language of adolescents.* Amsterdam: John Benjamins.

Andersen, Roger W. & Yasuhiro Shirai. 1996. Primacy of aspect in first and second language acquisition: the pidgin- creole connection. In William C. Ritchie & Tej K. Bhatia (eds.), *Handbook of second language acquisition*, 527-570. San Diego, CA: Academic Press.

Ariel, Mira. 1994. Pragmatic operators. In Ronald E. Asher (ed.), The encyclopedia of language and linguistics. Oxford: Pergamon Press.

Ariel, Mira. 1998. Pragmatic operators. In R.E. Asher (ed.), *The encyclopedia of language and linguistics,* 3250-3253.Oxford: Pergamon Press.

Bazzanella, Carla. 1990. Phatic connectives as interactional cues in contemporary spoken Italian. *Journal of Pragmatics* 14. 629-647.

Bazzanella, Carla & Lucia Morra. 2002. Discourse markers and the indeterminacy of translation. In Iørn Korzen & Caria Marello (eds.), *Argomenti per una linguistica della traduzione, On linguistic aspects of translation, Notes pour une linguistique de la traduction*. Alessandria: Edizioni dell' Orso. 149-157.

Blackwell, Susan. 2000. Looking up *look*: discourse markers in the Bank of English. In John M. Kirk (ed.), Corpora galore. Analyses and techniques in describing English. Papers from the Nineteenth International Conference on English Language Research on Computerised Corpora (ICAME 1998), 3-16. Amsterdam: Rodopi.

Blakemore, Diane. 1987. *Semantic constraints on relevance*. Oxford: Blackwell.

Blakemore, Diane. 1998. 'So' as a constraint on relevance. In Ruth M., Kempson (ed.), *Mental representations. The interface between language and reality,* 183-195. Cambridge: Cambridge University Press

Blakemore, Diane. 1992. *Understanding utterances. An introduction to pragmatics*. Oxford: Blackwell.

Blakemore, Diane. 2002. *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Blass, Regina. 1990. Constraints on relevance. A key to particle typology. *Notes on Linguistics* 48. 8-21.

Bloomer, Aileen. 2005. *Introducing language in use: a coursebook*. London: Routledge.

Blum-Kulka, Shoshana, Juliane House, & Gabriele Kasper. (eds.), 1989. *Cross-cultural pragmatics: requests and apologies*. Norwood, NJ: Ablex

Bolinger, Dwight. 1989. *Intonation and its uses: melody and grammar in discourse*.

Stanford, CA: Stanford University Press.

Brinton, Laurel J. 1990. The development of discourse markers in English. In Jacek Fisiak (ed.), *Historical linguistics and philology*, 45-71. Berlin: Mouton de Gruyter.

Brinton, Laurel J. 1998. 'The flowers are lovely; only, they have no scent.': the evolution of a pragmatic marker in English. In Raimund Borgmeier, Herbert Grabes & Andreas H. Jucker (eds.), *Anglistentag 1997 Giessen. Proceedings*, 9-33. Tier: Wissenschaftlicher Verlag.

Brinton, Laurel J. 2007. Development of *I mean*: implications for the study of historical pragmatics. In Susan Fitzmaurice & Irma Taavitsainen (eds) *Methods in historical pragmatics*, 37-80.Berlin: Mouton de Gruyter.

Brown, Penelope & Stephen C. Levinson. 1978. Universals in language usage: politeness phenomena. In Esther N. Goody (ed.), *Questions and politeness*, 56-289. Cambridge: Cambridge University Press.

Brown, Penelope & Stephen C. Levinson. 1987. *Politeness. Some universals in language usage.* Cambridge: Cambridge University Press.

Carlson, Lauri. 1984. *"well" in dialogue games: a discourse analysis of the interjection 'well' in idealized conversation.* Amsterdam, etc.: John Benjamins.

Chappell, Hilary. 1991. Strategies for the assertion of obviousness and disagreement in Mandarin: a semantic study of the modal particle *me*. *Australian Journal of Linguistics* 11. 39-65.

Chen, Shaojie. 2011. *Yinghan tuokouxiou biaojiyu shiyong de duibi yanjiu [A comparative study of discourse markers in English and Chinese TV talk shows]*. Jinan: Shandong University MA thesis.

Chen, Xing, Ye, Lei, & Zhang, Yanyin. 1995. Refusing in Chinese. In Gabriele Kasper (ed.), 119-163.

Chen ,Yiya &Agnes Weiyun He. 2001. *Dui bu dui* as a pragmatic marker: evidence from Chinese classroom discourse. *Journal of Pragmatics* 33. 1441-1465.

Coates, Jennifer. 1998. Gossip revisited: language in all-female groups. In Jenny Cheshire & Peter Trudgill (eds.), *The sociolinguistic reader* 2 (Gender and

Discourse), 127-152. London: Edward Arnold. Reprinted from Jennifer Coates & Deborah Cameron (eds.), 1989. *Women in their speech communities: new perspectives on language and sex*, 94-122. London: Longman.

Crystal, David. 1988. Another look at, *well, you know*…. *English Today* 13. 47-49.

Crystal, David & Derek Davy. 1975. *Advanced conversational English*. London: Longman.

De Fina, Anna. 1997. An analysis of Spanish *bien* as a marker of classroom management in teacher-student interaction. *Journal of Pragmatics* 28. 337-354.

Demirci, Mahide & Brian Kleiner. 1997. Discourse markers in Second Language Research. *Journal of Intensive English Studies* 11. 131-42.

Du, Jinwen Steinberg. 1995. Performance of face-threatening acts in Chinese: complaining, giving bad news, and disagreeing. In Gabriele Kasper (ed.), 165-206.

Edmonson, Willis. 1981. *Spoken discourse: a model for analysis*. New York: Longman.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Erman, Britt. 1986. Some pragmatic expressions in English conversation. In Gunnel Tottie & Ingegerd Bäcklund (eds.) *English in speech and writing. A symposium*, 131-147. Stockholm: Almqvist & Wiksell.

Erman, Britt. 1987. *Pragmatic expressions in English*. Stockholm: Almqvist & Wiksell.

Erman, Britt. 1992. Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change* 4. 217-234.

Erman, Britt. 2001. Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics* 33. 1337-1359.

Feng, Guangwu. 2008. Pragmatic markers in Chinese. *Journal of pragmatics* 40. 1687-1718.

Fischer, Kerstin & Martina Drescher. 1996. Methods for the description of discourse particles: contrastive analysis. *Language Sciences* 18. 853-861.

Flowerdew, John & Steve Tauroza. 1995. The effects of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition* 17. 435-458.

Foolen, Ad. 1997. Pragmatic particles. In Jef Verschueren, Jan-Ola Östman, Jan Blommaert & Chris Bulcaen (eds.), *Handbook of pragmatics 1996 installment*. Amsterdam, etc.: John Benjamins.

Fox Tree, Jean E. & Josef C. Schrock. 1999. Discourse markers in spontaneous speech: oh what a difference an oh makes. *Journal of Memory and Language* 40. 280-295.

Fox Tree, Jean E. & Josef C. Schrock. 2002. Basic meanings of *you know* and *I mean. Journal of Pragmatics* 34. 727-747.

Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguistica Hungarica* 38. 19-33.

Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14. 383-395.

Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6. 167-190.

Fraser, Bruce. 1998. Contrastive discourse markers. In Andreas H. Jucker & Yael Ziv (eds.), 301-326.

Fraser, Bruce. 1999. What are discourse markers in English?. *Journal of Pragmatics* 31. 931-952.

Fuller, Janet M. 2001. The principle of pragmatic detachability in borrowing: English-origin discourse markers in Pennsylvania German. *Linguistics* 39. 351-369.

Fuller, Janet M. 2003. Discourse marker use across speech contexts: a comparison of native and non-native speaker performance. *Multilingua* 22. 185-208.

Fung, Loretta & Ronald Carter. 2007. Discourse markers and spoken English: native and learner use in pedagogic settings. *Applied Linguistics* 28 (3). 410-439.

Garner, Richard. 1969. Caton on epistemic qualifiers. In William Todd (ed.), *Studies in philosophical linguistics,* 55-76. Evanston: Great Expectations Press.

Gerhardt, Julia & Charles Stinson. 1994. The nature of therapeutic discourse:

accounts of the self. *Journal of Narrative and Life History* 4(3). 151-191.

Goldberg, Julia A. 1976. *The syntax, semantics, pragmatics and sociolinguistics of some conventionalized parenthetical clauses in English: 'you know' and 'I mean"*. Cambridge: Cambridge University Diploma Dissertation.

Goldberg, Julia A. 1980. *Discourse particles: an analysis of the role of 'you know', 'I mean', 'well' and 'actually' in conversation*. Cambridge: Cambridge University PhD thesis.

Goldberg, Julia A. 1981. *Hey, y'know, have I got a topic for you*. Unpublished paper.

Goss, Emily L. & Joseph C. Salmons. 2000. The evolution of a bilingual discourse marking system: modal particles and English markers in German-American dialects. *International Journal of Bilingualism* 4 (4). 469-484.

Grice, Paul H.1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics* 3 (Speech Acts) 41-58. New York: Academic.

Grice, Paul H. 1989. *Studies in the ways of words*. Cambridge, Mass.: Harvard University Press.

Halliday, M.A.K & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Halliday, M.A.K 1994. *An introduction to functional grammar,* 2nd edn. London: Edward Arnold.

Hansen Mosegaard, Maj-Britt. 1998. *The function of discourse particles. A study with special reference to spoken standard French.* Amsterdam, etc.: John Benjamins.

Hasund, Kristine. 2002. *Congratulations, like! – Gratulerer, liksom!* Pragmatic particles in English and Norwegian. In Leiv Egil Breivik & Angela Hasselgren (eds.), *From the COLT's mouth… and others. Language corpora studies in honour of Anna-Brita Stenström.* 125-139. Amsterdam: Rodopi.

Hays, Paul R. 1992. Discourse markers and L2 acquisition. In Don Staub & Cheryl L. Delk (eds.), *The proceedings of the twelfth second language research forum*, 24-34. Michigan: Papers in Applied Linguistics – Michigan.

He, Agnes Weiyun & Brian Lindsey. 1998. "You know" as an information status enhancing device: arguments from grammar and interaction. *Functions of Language*

5. 133-155.

He, Anping. 2002. On the discourse marker *so*. In Pam Peters, Peter Collins & Adam Smith (eds.), *New frontiers of corpus research*, 41-52. Amsterdam: Rodopi.

Hellermann, John & Andrea Vergun. 2007. Language which is not taught: The discourse markers use of beginning adult learners of English. *Journal of Pragmatics* 39. 157-179.

Holmes, Janet. 1986. Functions of *you know* in women's and men's speech. *Language in Society* 15 (1). 1-22.

Holmes, Janet. 1995. *Women, men and politeness*. Harlow: Longman.

Huspek, Michael. 1989. Linguistic variability and power: an analysis of *you know/I think* variation in working-class speech. *Journal of Pragmatics* 13. 661-683.

Hyltenstam, Kenneth. 1977. Implicational patterns in interlangauge syntax variation. *Language Learning* 27. 383-411.

James, Allan R. 1983. Compromisers in English: a cross-disciplinary approach to their interpersonal significance. *Journal of Pragmatics* 7. 191-206.

Jucker, Andreas H. 1988. Review of Deborah Schiffrin, discourse markers. (Studies in International Sociolinguistics 5). *Multilingua* 7. 219-224. Cambridge: CUP

Jucker, Andreas H. 1993. The discourse marker *well*: a relevance theoretical account. *Journal of Pragmatics* 19. 435-452.

Jucker, Andreas H. & Yael Ziv (eds.). 1998. *Discourse markers: description and theory.* Amsterdam, etc.: John Benjamins.

Jucker, Andreas H. & Sara W. Smith. 1998. And people just you know like 'wow': discourse markers as negotiating strategies. In Andreas H. Jucker & Yael Ziv (eds.), 171-202.

Kroon, Caroline. 1995. *Discourse particles in Latin. A study of nam, enim, autem, vero and at*. Amsterdam; J.C. Gieben Publisher.

Kasper, Gabriele (ed.).1995. *Pragmatics of Chinese as a native and target Language.* Honolulu, HI: University of Hawaii Press.

Kasper, Gabriele. 2000. Data collection in pragmatic research. In Helen

Spencer-Oatey (ed.), *Culturally speaking: managing rapport through talk across cultures*, 316-341. London: Continuum.

Keller, Eric. 1979. Gambits: Conversational strategy signals. *Journal of Pragmatics* 3. 219-238.

Kryk, Barbara. 1992. A cross-linguistic look at discourse particles. In Christian Mair & Manfred Markus (eds.), *New departures in contrastive linguistics. Neue Ansätze in der Kontrastiven Linguistik. Proceedings of the conference held at the Leopold-Franzens-University of Innsbruck, Austria, 10-12 May 1991*. 43-50. Innsbruck: Verlag des Institus für Sprachwissenschaft.

Kyratizis, Amy & Susan Ervin-Tripp. 1999. The development of discourse markers in peer interaction. *Journal of Pragmatics* 31. 1321-1338.

Labov, William. 1970. The study of language in its social context. *Studium Generale* 23. 30-87.

Lakoff, Robin. 1973. Questionable answers and answerable questions. In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli & Sol Saporta (eds.), *Issues in linguistics. Papers in honor of Henry and Renée Kahane*, 453-467.Urbana: University of Illinois Press.

Lamiroy, Beatrice. 1994. Pragmatic connectives and L2 acquisition: the case of French and Dutch. *Pragmatics* 4. 183-201.

Lee-Wong, Song Mei. 2001. Coherence, focus and structure: The role of discourse marker *ne*. *Pragmatics* 11. 139-154.

Le Lan, Barbara. 2007. Orchestrating conversation. The multifunctionality of *well* and *you know* in the joint construction of a verbal interaction. In Angès Celle & Ruth Huart (eds.), *Connectives as discourse landmarks*, 103-116. Amsterdam: John Benjamins.

Lenk, Uta. 1998. *Marking discourse coherence. Function of discourse markers in spoken English*. Tübingen: Gunter Narr.

Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: a functional reference grammar*. Berkeley, CA: University of California Press.

Liao, Chao-chih & Mary I. Bresnahan. 1996. A contrastive study on American English and Mandarin refusal strategies. *Language Sciences* 18(3-4). 703-727.

Liao, Silvie. 2009. Variation in the use of discourse markers by Chinese teaching assistants in the US. *Journal of Pragmatics* 41(7). 1313-1328.

Liu, Binmei. 2009. Chinese discourse markers in oral speech of mainland Mandarin speakers. *Proceedings of the 12th North American Conference on Chinese Linguistics (NACCL-21)*2. 358-374.

Liu, Yanli. 2005. *Kouyu jiaoji zhong de huayu biaoji [Discourse markers in spoken interaction]*. Hangzhou: Zhejiang University PhD thesis.

Maschler, Yael. 1994. Metalanguaging and discourse markers in bilingual conversation. *Language in Society* 23. 235-366.

Meng, Katharina & Susanne Schrabback. 1999. Interjections in adult-child discourse: the cases of German *hm* and *na*. *Journal of pragmatics* 31. 1263-1287.

Montes, Rosa Graciela. 1999. The development of discourse markers in Spanish: interjections. *Journal of Pragmatics* 31. 1289-1319.

Müller, Simone. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam, etc.: John Benjamins.

Murray, Dinah. 1979. "Well". *Linguistic Inquiry* 10 (4). 727-732.

Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring natural language: working with the British component of the International Corpus of English*. Amsterdam: John Benjamins.

Nikula, Tarja. 1996. *Pragmatic force modifiers. A study in interlanguage pragmatics.* Jyväskylä: University of Jyväskylä

Norrick, Neal R. 2009. Interjections as pragmatic markers. *Journal of Pragmatics* 41. 866-891.

Odlin, Terence. 2003. Cross-linguistic influence. In Catherine J. Doughty & Michael H. Long (eds.), *Handbook of second language acquisition*, 436-486. Malden, MA: Blackwell.

Olynak, Marian, Alison d'Angeljan & David Sankoff. 1990. A quantitative and

qualitative analysis in the native and second language speech of bilinguals. In Robin Scarcella, Elaine S. Andersen & Stephen D. Krashen.(eds.), *Developing communicative competence in a second language*, 139-155. Boston: Heinle & Heinle.

Ortega, Lourdes. 2009. *Understanding second language acquisition.* London: Hodder Education.

Östman, Jan-Ola. 1981. *You know. A discourse-functional approach.* Amsterdam: John Benjamins.

Östman, Jan-Ola. 1982. The symbiotic relationship between pragmatic particles and impromptu speech. In Nils Erik Enkvist (ed.), *Impromptu speech: A symposium*. Åbo: The research Institute of the Åbo Akademi Foundation. 147-177.

Östman, Jan-Ola. 1995. Pragmatic particles twenty years after. In Brita Wårvik, Sanna-Kaisa Tanskanen & Risto Hiltunen (eds), *Organization in discourse. Proceedings from the Turku Conference.* University of Turku, Finland, 95-108.

Park, Yong-Yae. 1998. A discourse analysis of contrastive connectives in English, Korean, and Japanese conversation: with special reference of dispreferred responses. In Andreas H. Jucker & Yael Ziv (eds.), 227-300.

Powell, Mava Jo. 1992. The systematic development of correlated interpersonal and metalinguistic uses in stance adverbs. *Cognitive Linguistics* 3. 75-110.

Quirk, Randolph , Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1972. *A grammar of contemporary English*. London: Longman.

Redeker, Gisela. 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14. 367-381.

Redeker, Gisela. 1991. Review article: Linguistic markers of discourse structure. Review of Deborah Schiffrin, 1987, *Discourse markers*, Cambridge: CUP. *Linguistics* 29. 1139-1172.

Ringbom, Håkan. 2007. *Cross-linguistic similarities in foreign language learning.* Bristol: Multilingual Matters Ltd.

Romero Trillo, Jesús. 2002. The pragmatic fossilization of discourse markers in

non-native speakers of English. *Journal of Pragmatics* 34. 769-784.

Salmons, Joseph C. 1990. Bilingual discourse marking: code switching borrowing, and convergence in some German-American dialects. *Linguistics* 28. 453-480.

Sankoff, Gillian, Pierrette Thibault, Noami Nay, Hélène Blondeau, Marie-Odile & Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9. 191-217.

Selinker, Larry. 1969. Language transfer. *General Linguistics* 9. 67-92.

Schegloff, Emanuel, Gail Jefferson & Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53(2). 361-382.

Schiffrin, Deborah. 1995. Conversational coherence: the role of *well*. *Language* 61. 640-667.

Schiffrin, Deborah. 1987. *Discourse markers.* Cambridge: Cambridge University Press.

Schourup, Lawrence. 1985. *Common discourse particles in English conversation*. New York: Garland.

Schourup, Lawrence. 1999. Tutorial overview: discourse markers. *Lingua* 107. 227-265.

Singleton, David. 1987. Mother and other tongue influence on learner French: a case study. *Studies in Second Language Acquisition* 9(3). 327-345.

Spencer-Oatey, Helen (ed.), 2000. *Culturally speaking: managing rapport through talk across cultures.* London: Continuum.

Spada, Nina & Patsy M. Lightbown. 1999. Instruction, first language influence and developemental readiness in second language acquisition. *Modern Language Journal* 83, 1-22.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: communication and cognition*. Oxford: Blackwell.

Stenström, Anna-Brita. 1990. Lexical items peculiar to spoken discourse. In Jan Svartik (ed.), *The London-Lund corpus of spoken English: description and research*,

137-175. Lund: Lund University Press.

Stenström, Anna-Brita.1998. From sentence to discourse: cos(because) in teenage talk. In Andreas H. Jucker & Yael Ziv (eds.), 127-146.

Stubbe, Maria & Janet Holmes. 1995. *You know, eh*, and other 'exasperation expressions': an analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication* 15 (1). 63-88.

Svartvik, Jan. 1980. *Well* in conversation In Sidney Greenbaum,Geoffrey Leech & Jan Svartvik (eds.), *Studies in English linguistics for Randolph Quirk*, 167-177. London: Longman.

Swan, Michael. 1992. Language conundrums. *EFL Journal* 46 (4). 373-377.

Swan, Michael. 1994. Language conundrums. *EFL Journal* 48 (4). 356-360.

Takahara, Paul O. 1998. Pragmatic functions of the English discourse marker anyway and its corresponding contrastive Japanese markers. In Andreas H. Jucker & Yael Ziv (eds.), 327-352.

Takahara, Paul O. 1999. Pragmatic functions of discourse markers in English and Japanese. In Jef Verschueren (ed.), *Pragmatics in 1998. Selected papers from the 6th International Pragmatics Conference* 2. 547-576. Antwerp: International Pragmatics Association.

Takahashi, Satomi. 1996. Pragmatic transferability *Studies in Second Language Acquisition* 18(2). 189-223.

Tao, Hongyin. 2003. Cong yuyin yufa he huayu tezheng kan zhidao geshi zai tanhua zhong de yanhua [Phonological, grammatical and discourse evidence for the emergence of *zhidao* construction]. *Zhongguo Yuwen [Chinese Language]* 4. 291-302.

Thompson, Geoff. forthcoming. *Conjunctive relations in discourse: a tri-functional study of six English registers*. London: Equinox.

Thompson, Sandra A. & Anthony Mulac. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Elisabeth Closs

Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, 313-329. Amsterdam, etc.: John Benjamins.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meaning in English: an example of simplification. *Language* 65. 31-55.

Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalization. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation in language*, 31-54. Cambridge: Cambridge University Press.

Tyler, Andrea. 1992. Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly* 26(4). 713-729.

Unger, Christoph J. 1996. The scope of discourse connectives: implication for discourse organization. *Journal of Pragmatics* 32. 402-438.

Van Barr, Tim. 1996. Particles. In Betty Devriendt, Louis Goossens & Johan van der Auwera (eds.), *Complex structures: a functionalist perspective*, 259-301. Berlin & New York: Mouton de Gruyter.

Van Bogaert, Julie. 2007. In Antoinette Renouf & Andrew Kehoe (eds.), *Corpus linguistics: refinements and reassessments*, 131-154. Amsterdam: Rodopi.

Wang, Hong & Yunfeng Ge. 2004. Tanhuazhe juese ji xianghu guanxi dui huayu biaojiyu shiyong yingxiang de yuyong fenxi [The influence of speakers' role and their relations on discourse marker use]. *Shangdong Nongyedaxue xuebao (shehukexue ban) [Journal of Shandong Agricultural University (Social Science)]* 6(4). 118-122.

Watts, Richard J. 1988. A relevance-theoretic approach to commentary pragmatic markers: the case of *actually*, *really* and *basically*. *Acta Linguistica Hungarica* 38. 235-260.

Watts, Richard J. 1989. Taking the pitcher to the 'well': native speakers' perception of their use of discourse markers in conversation. *Journal of Pragmatics* 13. 203-237.

Wayland, Ratree P. & Susan G. Guion. 2004. Training English and Chinese listeners to perceive Thai tones: a preliminary report. *Language Learning* 54 (4). 681-712.

Wierzbicka, Anna. 1976. Particles and linguistic relativity. *International Review of Slavic Linguistics* 1(2-3). 327-367.

Wierzbicka, Anna. 2003. *Cross-cultural pragmatics: The semantics of human interaction*. Berlin: Mouton de Gruyter.

Yang, Xiaoxia. 2007. *Cong huayu biaojiyu yuan yuyong kongneng jiaodu fenxi yanjiu you know he nizhidao [A analysis of you know and nizhidao from the perspective of the metapragmatic functions of discourse markers]*. Shanghai: Shanghai International Studies University MA thesis.

Zhang, Shuting. 2007. *Hanying biaoji duibi yanjiu [A comparative study of Chinese and English discourse markers]*. Wuhan: Central China Normal University MA thesis.

Appendix I

coding of *I mean* in the native speaker data

locations of *I mean*	functions
3-4b	assumption-correction
3-73b	assumption-correction
3-140b	assumption-correction
3-167b	assumption-correction
5-14a	assumption-correction
5-49a	assumption-correction
6-31a	assumption-correction
8-60b	assumption-correction
8-175b	assumption-correction
11-7b	assumption-correction
12-66a	assumption-correction
13-18c	assumption-correction
13-86e	assumption-correction
13-105e	assumption-correction
13-139b	assumption-correction
13-219b	assumption-correction
14-55c	assumption-correction
15-83a	assumption-correction
19-131a	assumption-correction
25-250a	assumption-correction
25-287b	assumption-correction
29-16b	assumption-correction
29-22b	assumption-correction
29-53b	assumption-correction
29-122a1+2	assumption-correction
29-242a2	assumption-correction
29-329c	assumption-correction

30-6a	assumption-correction
30-82b	assumption-correction
31-7a	assumption-correction
31-11a	assumption-correction
31-131a	assumption-correction
31-151b2	assumption-correction
32-238a	assumption-correction
35-155b	assumption-correction
36-93b	assumption-correction
39-317a	assumption-correction
40-129d	assumption-correction
43-94b	assumption-correction
43-170a	assumption-correction
45-255b	assumption-correction
50-115b	assumption-correction
52-165b	assumption-correction
53-43b	assumption-correction
54-139a	assumption-correction
54-175a1	assumption-correction
54-181b1	assumption-correction
58-248b	assumption-correction
59-209a	assumption-correction
60-55b	assumption-correction
60-71b	assumption-correction
61-355b	assumption-correction
62-76b	assumption-correction
64-64a	assumption-correction
66-42a	assumption-correction
66-56b	assumption-correction
66-104b	assumption-correction

66-133b	assumption-correction
66-138b	assumption-correction
66-203a	assumption-correction
71-206c	assumption-correction
72-49b	assumption-correction
72-218b	assumption-correction
73-166b	assumption-correction
76-119b	assumption-correction
78-180a	assumption-correction
79-11a2	assumption-correction
80-280b2	assumption-correction
81-210b1	assumption-correction
84-73a	assumption-correction
85-87b	assumption-correction
86-224a	assumption-correction
90-30b	assumption-correction
4-143b	cause
74-193a	cause
76-101b	cause
76-205b	cause
79-196a	cause
81-37a	cause
81-210b2	cause
89-168b	cause
3-114b	conclusion
23-186a	conclusion
27-205d	conclusion
31-161a	conclusion
43-50b	conclusion
48-341b	conclusion

60-222b	conclusion
63-26a	conclusion
66-125b	conclusion
70-123a	conclusion
71-313c	conclusion
73-49b	conclusion
76-125a	conclusion
3-133b1	softener of FTA
5-28a1	softener of FTA
5-189a1	softener of FTA
13-169c	softener of FTA
14-231a	softener of FTA
17-123a	softener of FTA
18-328a	softener of FTA
19-286c	softener of FTA
23-298b	softener of FTA
29-96b	softener of FTA
29-175c	softener of FTA
29-201b	softener of FTA
31-86b	softener of FTA
31-159a	softener of FTA
33-49b	softener of FTA
36-35b	softener of FTA
36-103a	softener of FTA
38-149c	softener of FTA
42-134c	softener of FTA
47-198a	softener of FTA
50-196a	softener of FTA
54-203a	softener of FTA
58-91b1	softener of FTA

58-105a	softener of FTA
58-200b	softener of FTA
59-289a	softener of FTA
66-63b	softener of FTA
66-112b1	softener of FTA
66-208b1	softener of FTA
69-281a	softener of FTA
71-124c	softener of FTA
71-261d	softener of FTA
74-148a	softener of FTA
75-65b	softener of FTA
76-139a	softener of FTA
77-257b	softener of FTA
79-115a	softener of FTA
80-63b	softener of FTA
86-156a	softener of FTA
3-82a	exemplification
3-134b	exemplification
5-79a	exemplification
5-163a	exemplification
5-179b	exemplification
7-167b	exemplification
13-34c	exemplification
13-68e	exemplification
13-191f	exemplification
13-209f	exemplification
13-217b	exemplification
13-233e	exemplification
14-157b	exemplification
15-17b	exemplification

15-228a	exemplification
17-115b	exemplification
18-329a	exemplification
19-371d	exemplification
20-159b	exemplification
23-209a	exemplification
27-80b	exemplification
37-117b	exemplification
37-146b	exemplification
37-164b	exemplification
37-267b	exemplification
42-145a	exemplification
47-173b	exemplification
48-344	exemplification
50-131b	exemplification
53-253b	exemplification
54-46b	exemplification
58-257b	exemplification
59-147a	exemplification
60-105b	exemplification
63-141a	exemplification
63-274a	exemplification
63-281a	exemplification
66-160b	exemplification
66-189a	exemplification
66-191a	exemplification
72-220b	exemplification
76-184b	exemplification
76-185b	exemplification
79-41c	exemplification

79-54a	exemplification
81-24a	exemplification
82-34a	exemplification
87-273b	exemplification
2-71b	explictness
4-140a	explictness
5-40a	explictness
6-40a	explictness
6-196a	explictness
6-242b	explictness
6-246a	explictness
7-26b	explictness
7-278b	explictness
8-68b	explictness
8-152b	explictness
9-123a	explictness
15-71a	explictness
15-160b	explictness
15-187a	explictness
18-328a	explictness
25-31a	explictness
29-257a	explictness
31-67a	explictness
38-129a	explictness
39-86b	explictness
45-32a	explictness
46-35a	explictness
50-185b	explictness
54-128a	explictness
54-129a	explictness

59-83a	explictness
66-67b	explictness
66-124b	explictness
66-146b	explictness
66-165b	explictness
68-294c	explictness
69-321b	explictness
71-258d	explictness
72-105a	explictness
72-111b	explictness
75-113b	explictness
76-12a	explictness
78-197b	explictness
80-224b	explictness
84-49b	explictness
84-151b	explictness
69-273a	hesitation marker
2-108b	hesitation marker
3-133b2	hesitation marker
45-258b	hesitation marker
48-74b	hesitation marker
56-252b	hesitation marker
56-269b	hesitation marker
60-62b	hesitation marker
66-200b	hesitation marker
2-12c	interactional repair
3-121b	interactional repair
4-11b	interactional repair
4-97a	interactional repair
4-98a	interactional repair

5-89a	interactional repair
5-96a	interactional repair
12-123b	interactional repair
15-14a	interactional repair
16-236a	interactional repair
24-115a	interactional repair
26-341c	interactional repair
29-139a	interactional repair
29-271a	interactional repair
38-80c	interactional repair
46-417a	interactional repair
51-138a	interactional repair
54-175a2	interactional repair
58-107a	interactional repair
58-217b	interactional repair
61-134b	interactional repair
72-44b	interactional repair
7-288b	justification
2-89b	justification
3-67b	justification
3-119b	justification
4-87a	justification
5-28a2	justification
5-30a	justification
5-172a	justification
5-189a2	justification
6-138b	justification
6-144b	justification
6-317b	justification
7-79b	justification

7-87a	justification
10-155a	justification
10-190b	justification
10-262b	justification
11-30a	justification
11-119a	justification
11-237a	justification
12-29b	justification
12-117c	justification
12-124c1	justification
12-124c2	justification
12-208a	justification
12-227b	justification
13-190f	justification
15-21b	justification
15-58b	justification
15-100a	justification
15-131a	justification
15-145a	justification
16-165a	justification
16-327e	justification
17-50a	justification
18-324a1	justification
18-333a	justification
19-288a	justification
19-289a	justification
19-339d	justification
19-360c	justification
21-32a	justification
22-57a	justification

23-40b	justification
23-53a	justification
23-88a	justification
23-92a	justification
23-325a	justification
24-79a	justification
26-188a	justification
27-72b	justification
27-131c	justification
27-184d	justification
27-188d1+2	justification
27-195d	justification
28-262a	justification
29-27b	justification
29-116a	justification
29-140a	justification
29-159a	justification
29-169a	justification
29-176c	justification
29-183a	justification
29-202b	justification
29-241a	justification
29-242a1	justification
30-64c	justification
30-89a	justification
31-40b	justification
31-41b	justification
31-45a	justification
31-55a	justification
31-56a	justification

31-87b	justification
31-102b	justification
31-139b	justification
31-160a	justification
33-28b	justification
33-95b	justification
33-98b	justification
34-136b	justification
34-164b	justification
35-186a	justification
36-38a1	justification
36-38a2	justification
36-80a	justification
37-178b	justification
37-258a	justification
37-265b	justification
38-21a	justification
38-77a	justification
38-275c	justification
39-289a	justification
39-301b	justification
39-304a	justification
41-239a	justification
41-387a	justification
42-171c	justification
43-47b	justification
43-95b	justification
43-121a	justification
43-275a	justification
44-351a	justification

44-386a	justification
45-87b	justification
45-92b	justification
48-72a	justification
50-221b	justification
52-162a	justification
54-3a	justification
56-186a	justification
56-261b	justification
57-211a	justification
57-217a	justification
57-225a	justification
59-103a	justification
59-304b	justification
60-64b	justification
60-68b	justification
60-74b	justification
60-130b	justification
61-111b	justification
61-127b	justification
61-169b	justification
61-268b	justification
61-348b	justification
62-19b	justification
62-199a	justification
63-100a	justification
63-245a	justification
63-246a	justification
64-47a	justification
65-328e	justification

66-77a	justification
66-122b	justification
66-164b	justification
66-173b	justification
66-178b	justification
66-208b2	justification
66-213a	justification
67-29a	justification
67-166b	justification
67-178b	justification
68-27a	justification
68-34a	justification
68-103a	justification
68-160b	justification
69-86a	justification
69-149a	justification
69-161b	justification
69-222b	justification
69-240a	justification
69-276a	justification
69-290a	justification
71-61b	justification
71-232c	justification
71-289c	justification
72-199b	justification
72-232b	justification
73-12b	justification
73-15c	justification
73-21a	justification
73-102b	justification

74-25b	justification
75-21b	justification
75-118b	justification
76-135b	justification
78-148a	justification
78-160b	justification
79-34a	justification
79-57a	justification
79-63c	justification
79-102a	justification
79-140a	justification
79-227a	justification
80-88b	justification
80-153a1+2	justification
80-171b	justification
81-8a	justification
81-31a	justification
81-36a	justification
81-147b	justification
81-198a	justification
81-285b	justification
82-26a	justification
82-151b	justification
83-71a	justification
84-2a	justification
88-151a	justification
90-156d	justification
90-217a	justification
5-23a	quotative
80-280b1	quotative

3-70b	reason
5-8a	reason
5-42a	reason
5-52a	reason
5-220b	reason
10-48b	reason
10-282b	reason
14-137b	reason
20-68c	reason
27-4b	reason
29-240a	reason
34-155b	reason
37-260a	reason
39-231b	reason
41-123a	reason
50-43b	reason
50-48b	reason
50-114b	reason
51-40b	reason
54-23a	reason
70-267b	reason
76-120b	reason
77-58b	reason
90-226a	reason
3-25b	reformulation
11-21a	reformulation
15-225a	reformulation
24-138a	reformulation
31-192b	reformulation
39-285b	reformulation

46-363a	reformulation
52-176a	reformulation
54-190b	reformulation
62-80b	reformulation
63-130a	reformulation
66-64b	reformulation
72-201b	reformulation
72-229a	reformulation
76-202a	reformulation
77-293c	reformulation
1-125b	restart
3-13b	restart
3-75a	restart
4-132a	restart
5-31a	restart
5-58a	restart
5-174a	restart
6-99b	restart
7-174b	restart
10-24b	restart
12-167b	restart
12-196a	restart
18-336b	restart
18-343b1+2	restart
19-281c	restart
20-156d	restart
21-133c	restart
23-207a	restart
24-99a	restart
24-114a	restart

24-146a1	restart
26-340c	restart
29-117a	restart
29-134a	restart
29-189d	restart
29-330c	restart
31-65a	restart
31-151b1	restart
34-12a	restart
34-14b	restart
34b-125b	restart
34-161b	restart
36-38a4	restart
37-33b	restart
37-271b	restart
38-125a	restart
43-102a	restart
43-225b	restart
45-90b	restart
45-231b	restart
45-298a	restart
45-310b1	restart
45-314b	restart
45-321a	restart
45-322a	restart
46-45a	restart
47-97b	restart
48-97a	restart
48-340b	restart
48-345b	restart

49-41c	restart
50-87b	restart
52-157b	restart
52-164b	restart
54-136a	restart
54-180b	restart
54-181b2	restart
55-181a	restart
56-73a	restart
56-187a	restart
56-244b	restart
58-91b2	restart
59-256b	restart
60-18b	restart
60-75b	restart
60-139b	restart
60-163b	restart
62-67a	restart
62-96a	restart
63-275a	restart
64-37c	restart
66-112b2	restart
66-118b	restart
66-186a	restart
69-191a	restart
71-184b	restart
72-189b	restart
78-123b	restart
79-6a	restart
79-11a1	restart

79-97a	restart
79-98a1	restart
79-98a2	restart
81-39a	restart
83-196a	restart
84-40b	restart
84-232c	restart
85-171a	restart
89-6a	restart
89-22a2	restart
37-47b	result
66-57b	result
10-278b	resumption
12-49a	resumption
21-167c	resumption
21-169b	resumption
27-214d	resumption
32-211d	resumption
38-267a	resumption
43-9a	resumption
46-120a	resumption
62-186b	resumption
63-171a	resumption
71-226c	resumption
78-143a	resumption
84-56a	resumption
2-92b	indicating speaker attitude
3-120b	indicating speaker attitude
3-161b	indicating speaker attitude
5-15a	indicating speaker attitude

5-158b	indicating speaker attitude
5-171a	indicating speaker attitude
6-149b	indicating speaker attitude
6-227a	indicating speaker attitude
7-78b	indicating speaker attitude
8-240a	indicating speaker attitude
10-165b	indicating speaker attitude
10-230a	indicating speaker attitude
10-279b	indicating speaker attitude
13-220b	indicating speaker attitude
16-231d	indicating speaker attitude
19-110e	indicating speaker attitude
19-364a	indicating speaker attitude
22-173d	indicating speaker attitude
22-210d	indicating speaker attitude
22-295d	indicating speaker attitude
23-238a	indicating speaker attitude
23-239a	indicating speaker attitude
27-71b	indicating speaker attitude
27-114b	indicating speaker attitude
27-201d	indicating speaker attitude
29-6b	indicating speaker attitude
29-57a	indicating speaker attitude
29-126a	indicating speaker attitude
29-328c	indicating speaker attitude
31-145b	indicating speaker attitude
35-4b	indicating speaker attitude
37-190b	indicating speaker attitude
37-198b	indicating speaker attitude
38-83a	indicating speaker attitude

39-309a1	indicating speaker attitude
44-385a	indicating speaker attitude
44-389a	indicating speaker attitude
45-119b	indicating speaker attitude
54-104a	indicating speaker attitude
56-139b	indicating speaker attitude
57-164b	indicating speaker attitude
63-74a	indicating speaker attitude
66-73b	indicating speaker attitude
72-198b	indicating speaker attitude
76-136b	indicating speaker attitude
78-162a	indicating speaker attitude
80-81b	indicating speaker attitude
85-25b	indicating speaker attitude
85-327b	indicating speaker attitude
85-366a	indicating speaker attitude
86-169a	indicating speaker attitude
3-92a	summarisation
5-7a	summarisation
10-246b	summarisation
31-162a	summarisation
58-274b	summarisation
60-113b	summarisation
60-138b	summarisation
71-243d	summarisation
75-122b	summarisation
75-160b	summarisation
3-126b	transactional repair
4-2a	transactional repair
4-61a	transactional repair

5-26a	transactional repair
5-106b	transactional repair
6-89a	transactional repair
11-115c	transactional repair
13-13c	transactional repair
13-37c	transactional repair
13-250b	transactional repair
14-144b	transactional repair
16-356e	transactional repair
19-120c	transactional repair
19-362a	transactional repair
21-64c	transactional repair
21-277d	transactional repair
21-311c	transactional repair
24-146a2	transactional repair
27-183d	transactional repair
29-212b	transactional repair
31-185b	transactional repair
34-149b	transactional repair
38-51c	transactional repair
39-309a2	transactional repair
44-365b	transactional repair
54-157a	transactional repair
66-102b	transactional repair
76-74b	transactional repair
84-20a	transactional repair
84-154b	transactional repair
87-143b	transactional repair
88-125a	transactional repair

Appendix II	
Coding of *you know* in the native speaker data	
locations of *you know*	functions
3-142b	approximator
10-278b	approximator
13-253b1	approximator
13-253b2	approximator
23-59b	approximator
26-23a	approximator
31-152b2	approximator
31-154b2	approximator
64-90b	approximator
76-19b	approximator
84-35a	approximator
90-39d	approximator
90-226a2	approximator
52-165b	assumption-correction
3-126b	indicating speaker attitude
3-161b	indicating speaker attitude
10-47b	indicating speaker attitude
10-261b	indicating speaker attitude
10-262b	indicating speaker attitude
12-124c	indicating speaker attitude
12-142b	indicating speaker attitude
12-155c2	indicating speaker attitude
13-52e	indicating speaker attitude
13-174b	indicating speaker attitude
14-208b	indicating speaker attitude
16-284d	indicating speaker attitude
19-102a	indicating speaker attitude

19-364a	indicating speaker attitude
20-99a	indicating speaker attitude
23-213a	indicating speaker attitude
23-238a	indicating speaker attitude
28-7b	indicating speaker attitude
29-46a	indicating speaker attitude
29-150a	indicating speaker attitude
31-86b3	indicating speaker attitude
31-193b	indicating speaker attitude
34-175b	indicating speaker attitude
35-70b	indicating speaker attitude
35-93b	indicating speaker attitude
37-160b	indicating speaker attitude
42-200c	indicating speaker attitude
42-342b	indicating speaker attitude
44-385a	indicating speaker attitude
45-98b	indicating speaker attitude
45-99b	indicating speaker attitude
47-126b	indicating speaker attitude
48-9c	indicating speaker attitude
48-41b	indicating speaker attitude
48-137a	indicating speaker attitude
48-156c	indicating speaker attitude
48-192c	indicating speaker attitude
49-59b	indicating speaker attitude
50-82b	indicating speaker attitude
50-259b	indicating speaker attitude
52-18b	indicating speaker attitude
52-69b3	indicating speaker attitude
52-181a	indicating speaker attitude

52-193a	indicating speaker attitude
54-58b	indicating speaker attitude
57-136b	indicating speaker attitude
57-139b	indicating speaker attitude
57-225a	indicating speaker attitude
57-244b	indicating speaker attitude
58-158a1	indicating speaker attitude
58-217b2	indicating speaker attitude
60-113b3	indicating speaker attitude
62-40a	indicating speaker attitude
62-48a	indicating speaker attitude
66-73b	indicating speaker attitude
66-104b1	indicating speaker attitude
66-118b	indicating speaker attitude
66-170a	indicating speaker attitude
68-225c	indicating speaker attitude
71-265b	indicating speaker attitude
71-349a2	indicating speaker attitude
73-113c	indicating speaker attitude
75-119b	indicating speaker attitude
75-120b1+2	indicating speaker attitude
75-121b	indicating speaker attitude
76-114b	indicating speaker attitude
77-176b	indicating speaker attitude
81-117b2	indicating speaker attitude
81-278b	indicating speaker attitude
82-52a	indicating speaker attitude
84-105a	indicating speaker attitude
84-153b	indicating speaker attitude
84-169b	indicating speaker attitude

87-259a	indicating speaker attitude
90-209a	indicating speaker attitude
90-215a	indicating speaker attitude
49-270b	introducing background information
72-142b	introducing background information
79-49c	introducing background information
81-276b	introducing background information
82-113b	introducing background information
10-265b1	introducing background information
14-105b	introducing background information
15-5a	introducing background information
33-33a	introducing background information
52-204a2	introducing background information
90-194c	introducing background information
10-22a	cause
11-246b	cause
31-187b	cause
52-68b	cause
58-159a	cause
60-139b1	cause
60-139b2	cause
8-290b	indicating unspoken message to be completed by the hearer
13-111e	indicating unspoken message to be completed by the hearer
13-246b	indicating unspoken message to be completed by the hearer
16-153e	indicating unspoken message to be completed by the hearer
20-135c	indicating unspoken message to be completed by the hearer
22-58a	indicating unspoken message to be completed by the hearer
45-285b2	indicating unspoken message to be completed by the hearer
46-65c2	indicating unspoken message to be completed by the hearer
52-176a2	indicating unspoken message to be completed by the hearer

66-184b	indicating unspoken message to be completed by the hearer
71-353a	indicating unspoken message to be completed by the hearer
33-105b	indicating unspoken message to be completed by the hearer
34-136b	indicating unspoken message to be completed by the hearer
49-237b	indicating unspoken message to be completed by the hearer
52-109a	indicating unspoken message to be completed by the hearer
71-207c	indicating unspoken message to be completed by the hearer
71-317c	indicating unspoken message to be completed by the hearer
81-119b1	indicating unspoken message to be completed by the hearer
1-122b	conclusion
3-144b	conclusion
3-147b	conclusion
10-39b	conclusion
31-124b2	conclusion
35-187a	conclusion
62-72b	conclusion
64-76c	conclusion
79-88c	conclusion
89-62a	conclusion
9-308a	seeking confirmation
15-163a	seeking confirmation
3-9b	exemplification
3-94a	exemplification
3-134b	exemplification
5-179b	exemplification
9-159a	exemplification
10-41b1	exemplification
10-217a	exemplification
10-219a	exemplification
10-254b	exemplification

10-257b	exemplification
10-258b1	exemplification
10-258b2	exemplification
14-151c1	exemplification
14-230b	exemplification
14-232a	exemplification
32-232a	exemplification
35-43b	exemplification
36-224b	exemplification
37-41a	exemplification
37-163b	exemplification
37-187b	exemplification
49-282b	exemplification
52-31b	exemplification
52-33b	exemplification
52-34b	exemplification
52-35b	exemplification
58-257b	exemplification
60-22b	exemplification
60-134b	exemplification
60-202b	exemplification
63-244a	exemplification
64-96c1	exemplification
64-96c2	exemplification
66-182b	exemplification
76-198b	exemplification
79-86c	exemplification
80-282b	exemplification
82-41a1	exemplification
82-41a2	exemplification

82-96b1	exemplification
88-115b	explicitness
2-71b	explicitness
5-40a2	explicitness
7-136c	explicitness
9-61a	explicitness
9-109a	explicitness
9-158a	explicitness
9-267a	explicitness
9-342a	explicitness
10-38b	explicitness
13-221b	explicitness
14-2a	explicitness
14-168b	explicitness
15-261b	explicitness
24-126a	explicitness
30-64c	explicitness
31-124b1	explicitness
33-92a	explicitness
35-162b	explicitness
36-82b	explicitness
37-150b	explicitness
39-314a	explicitness
40-32d	explicitness
41-30b	explicitness
42-126b	explicitness
43-198a	explicitness
45-96b	explicitness
52-24b2	explicitness
54-1a	explicitness

55-212b	explicitness
56-175b	explicitness
58-15c1	explicitness
58-161a	explicitness
58-258b	explicitness
63-278a	explicitness
64-33a1	explicitness
64-103c	explicitness
66-146b	explicitness
67-184b	explicitness
72-232b	explicitness
77-373b	explicitness
84-180a	explicitness
85-11b	explicitness
87-143b	explicitness
88-58b	explicitness
88-65b	explicitness
13-168c	softener of FTA
55-225b	softener of FTA
58-110b	softener of FTA
62-69b	softener of FTA
35-88b	softener of FTA
8-201a	softener of FTA
30-202c	softener of FTA
31-86b1	softener of FTA
31-175b	softener of FTA
48-164a	softener of FTA
66-168a2	softener of FTA
68-154a	softener of FTA
78-41b	softener of FTA

3-56b1	hesitation marker
3-133b1	hesitation marker
3-133b2	hesitation marker
3-141b	hesitation marker
5-14a	hesitation marker
5-153a	hesitation marker
16-114b	hesitation marker
16-327e	hesitation marker
18-276a	hesitation marker
19-62a	hesitation marker
19-127a	hesitation marker
29-257a	hesitation marker
31-98b	hesitation marker
31-149b	hesitation marker
31-155b1	hesitation marker
31-155b2	hesitation marker
31-186b	hesitation marker
32-116d	hesitation marker
34-19b2	hesitation marker
45-329b	hesitation marker
48-74b	hesitation marker
48-226b	hesitation marker
49-4b	hesitation marker
52-29b	hesitation marker
52-176a1	hesitation marker
52-202a1	hesitation marker
56-19a	hesitation marker
56-269b	hesitation marker
58-15c2	hesitation marker
58-249b	hesitation marker

60-11b	hesitation marker
60-139b3	hesitation marker
66-97b	hesitation marker
67-257b	hesitation marker
80-153b	hesitation marker
81-92b	hesitation marker
82-96b3	hesitation marker
83-150a	hesitation marker
87-254a	hesitation marker
88-12a	hesitation marker
88-160b	hesitation marker
90-72d	hesitation marker
90-218a2	hesitation marker
90-241a	hesitation marker
4-97a	interactional repair
58-217b1	interactional repair
3-26b	justificaiton
5-40a1	justificaiton
5-183b	justificaiton
10-190b	justificaiton
17-329b	justificaiton
18-82a	justificaiton
18-87a	justificaiton
19-370a	justificaiton
21-2a	justificaiton
27-209c	justificaiton
29-56a	justificaiton
29-159a	justificaiton
30-315a	justificaiton
31-40b	justification

31-56a	justification
31-102b	justification
31-154b1	justification
32-52c	justification
32-88d	justification
34-56b	justification
35-98b	justification
36-199b	justification
36-216b1	justification
37-219b	justification
43-203b	justification
44-380a	justification
45-8b	justification
45-285b1	justification
48-95c	justification
48-379b	justification
49-56b	justification
54-124a	justification
59-30b	justification
62-151b	justification
69-207a	justification
69-275a	justification
71-349a1	justification
73-75b	justification
78-144a	justification
81-19a	justification
84-112a	justification
1-101a	indicating marked expression
3-6b	indicating marked expression
3-56b2	indicating marked expression

3-139b	indicating marked expression
8-119a	indicating marked expression
9-96a	indicating marked expression
9-268a	indicating marked expression
10-265b2	indicating marked expression
13-249b	indicating marked expression
14-177b	indicating marked expression
19-79a	indicating marked expression
19-137a	indicating marked expression
20-44b	indicating marked expression
22-170d	indicating marked expression
25-326b	indicating marked expression
27-72b	indicating marked expression
31-38b1	indicating marked expression
31-49b	indicating marked expression
31-52b	indicating marked expression
31-90b	indicating marked expression
31-101b	indicating marked expression
31-152b1	indicating marked expression
41-355b	indicating marked expression
45-97b	indicating marked expression
49-26b	indicating marked expression
52-7b	indicating marked expression
52-12b	indicating marked expression
52-25b	indicating marked expression
52-78a	indicating marked expression
52-110a	indicating marked expression
52-114a	indicating marked expression
52-129a	indicating marked expression
52-133a	indicating marked expression

52-202a2	indicating marked expression
52-206a	indicating marked expression
56-24a	indicating marked expression
57-157b	indicating marked expression
57-191b	indicating marked expression
58-42b	indicating marked expression
62-1a	indicating marked expression
63-131a	indicating marked expression
63-173a	indicating marked expression
63-188a3	indicating marked expression
66-165b	indicating marked expression
80-233a	indicating marked expression
82-71a	indicating marked expression
87-82b	indicating marked expression
88-114a	indicating marked expression
29-80a	indicating the coming message to be evaluated
31-38b2	indicating the coming message to be evaluated
31-123b	indicating the coming message to be evaluated
31-144b	indicating the coming message to be evaluated
31-145b	indicating the coming message to be evaluated
33-168b	indicating the coming message to be evaluated
34-164b1	indicating the coming message to be evaluated
34-164b2	indicating the coming message to be evaluated
37-168b	indicating the coming message to be evaluated
37-211b	indicating the coming message to be evaluated
39-43a	indicating the coming message to be evaluated
41-74b	indicating the coming message to be evaluated
43-254b	indicating the coming message to be evaluated
45-6b	indicating the coming message to be evaluated
47-148b	indicating the coming message to be evaluated

48-357c	indicating the coming message to be evaluated
48-361c	indicating the coming message to be evaluated
49-14b	indicating the coming message to be evaluated
49-22b1	indicating the coming message to be evaluated
52-5b	indicating the coming message to be evaluated
52-17b2	indicating the coming message to be evaluated
52-100a	indicating the coming message to be evaluated
52-111a	indicating the coming message to be evaluated
52-147b	indicating the coming message to be evaluated
52-151b	indicating the coming message to be evaluated
52-204a1	indicating the coming message to be evaluated
53-49b	indicating the coming message to be evaluated
56-17a	indicating the coming message to be evaluated
60-219b1	indicating the coming message to be evaluated
64-74c	indicating the coming message to be evaluated
64-112b	indicating the coming message to be evaluated
66-181b	indicating the coming message to be evaluated
71-109a	indicating the coming message to be evaluated
71-152a	indicating the coming message to be evaluated
75-26b	indicating the coming message to be evaluated
79-98a	indicating the coming message to be evaluated
79-197a	indicating the coming message to be evaluated
80-147b	indicating the coming message to be evaluated
80-292b	indicating the coming message to be evaluated
80-293b	indicating the coming message to be evaluated
80-294b	indicating the coming message to be evaluated
81-16a	indicating the coming message to be evaluated
82-34a1	indicating the coming message to be evaluated
82-34a2	indicating the coming message to be evaluated
82-114b	indicating the coming message to be evaluated

83-82a	indicating the coming message to be evaluated
83-86b	indicating the coming message to be evaluated
83-113b	indicating the coming message to be evaluated
83-152b	indicating the coming message to be evaluated
90-193c	indicating the coming message to be evaluated
90-205c	indicating the coming message to be evaluated
90-245a	indicating the coming message to be evaluated
3-58b	indicating the coming message to be evaluated
6-210a	indicating the coming message to be evaluated
8-49a	indicating the coming message to be evaluated
8-78a	indicating the coming message to be evaluated
8-175b	indicating the coming message to be evaluated
10-239a	indicating the coming message to be evaluated
12-81b	indicating the coming message to be evaluated
12-155c1	indicating the coming message to be evaluated
14-136b	indicating the coming message to be evaluated
14-151c2	indicating the coming message to be evaluated
14-212b	indicating the coming message to be evaluated
15-23b2	indicating the coming message to be evaluated
15-80a	indicating the coming message to be evaluated
19-281c	indicating the coming message to be evaluated
24-155b	indicating the coming message to be evaluated
3-136b	quotative
5-23a	quotative
10-263b	quotative
10-266b	quotative
10-267b	quotative
12-138c	quotative
18-286a	quotative
31-86b2	quotative

37-29b	quotative
41-321a	quotative
41-322a	quotative
49-17b	quotative
49-55b	quotative
50-115b	quotative
52-26b	quotative
53-340b	quotative
54-43b	quotative
54-152a	quotative
62-153b	quotative
63-188a1	quotative
64-99c	quotative
71-233c	quotative
71-344a	quotative
73-158b	quotative
80-280b	quotative
85-67b	quotative
90-162a	quotative
90-213a	quotative
90-218a1	quotative
3-69b	reason
3-160b	reason
5-7a	reason
5-119b	reason
10-46b	reason
10-48b	reason
15-240b	reason
27-168c1	reason
34-226b	reason

40-192d	reason
48-264c	reason
50-135b	reason
52-69b1	reason
52-69b2	reason
82-26a	reason
82-38a	reason
12-157c	reformulation
15-179b	reformulation
16-152e	reformulation
22-154d	reformulation
25-248a	reformulation
37-138b	reformulation
41-324a	reformulation
52-28b	reformulation
54-190b	reformulation
64-33a2	reformulation
80-163b	reformulation
82-11a	reformulation
84-84b	reformulation
3-89a	restart
3-91a	restart
5-244b	restart
10-41b2	restart
13-138b	restart
14-228b	restart
15-23b1	restart
15-177b	restart
31-2a	restart
33-190a	restart

36-80b	restart
36-216b2	restart
39-317a	restart
40-99d	restart
42-159b	restart
46-65c1	restart
49-21b	restart
49-25b	restart
49-276b	restart
52-13b	restart
52-24b1	restart
58-158a2	restart
60-62b	restart
60-113b2	restart
60-133b	restart
62-93a	restart
62-95a	restart
64-56a	restart
66-168a1	restart
67-206b	restart
72-189b	restart
75-49b	restart
78-84a	restart
79-216b	restart
80-61b	restart
82-96b2	restart
84-34a1+2	restart
84-55a	restart
90-211a	restart
13-108e	result

15-95a	result
19-133a	result
57-184a	result
62-78b	result
65-220c	result
84-94b	result
13-247b	indicating the most likely event
15-28b	indicating the most likely event
27-165c	indicating the most likely event
30-85a	indicating the most likely event
34-211b1	indicating the most likely event
34-211b2	indicating the most likely event
39-90b	indicating the most likely event
52-17b1	indicating the most likely event
54-205a	indicating the most likely event
58-141a	indicating the most likely event
60-87b	indicating the most likely event
63-188a2	indicating the most likely event
66-78a	indicating the most likely event
71-21a	indicating the most likely event
76-88b	indicating the most likely event
81-117b1	indicating the most likely event
84-146b	indicating the most likely event
90-226a1	indicating the most likely event
90-244a	indicating the most likely event
10-247b	summarization
10-258b3	summarization
34-224b	summarization
36-69a	summarization
45-275a	summarization

60-86b	summarization
60-113b1	summarization
60-219b2	summarization
79-46c	summarization
82-70a	summarization
83-22a	summarization
5-195b	introducing a new topic
9-1a	introducing a new topic
23-13a	introducing a new topic
29-358a	introducing a new topic
46-428c	introducing a new topic
48-278b	introducing a new topic
49-259c	introducing a new topic
52-14b	introducing a new topic
52-20b	introducing a new topic
52-71b	introducing a new topic
55-137b	introducing a new topic
63-1a	introducing a new topic
84-83b	introducing a new topic
85-100b	introducing a new topic