Culture-Sensitive Taxonomy of Response Tokens: Moving from Listenership to Speakership

Keiko Tsuchiya, MA, BA.

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

September 2010

Acknowledgements

I would like to thank Svenja Adolphs, who has patiently guided me over the past four years as a skillful and generous supervisor and mentor. Without Svenja's encouragement and faith in my work, this thesis would not have been completed. I am grateful to Ronald Carter, whose work on spoken grammar and social interactional context has a great influence on this study.

The completion of this thesis is largely due to support of many people. Special thanks to Linda Taylor for her continuous support since my MA study; to Diane Schmitt, for introducing me to Svenja at the starting point of my PhD study; to Katie Head, for her great help in collecting the research data; to Mike Handford, for being a resourceful mentor at home; to Dawn Knight for her relevant and useful instructions; to Steve Kirk and Chloe Harrison for proof reading my thesis; to Martin Brown and school staff for their efforts on sorting out the administrative issues.

I would like to extend my gratitude to late Yoichi Tanabe and Michiko Nakano at Waseda University, who first started me on this journey of discovery and let me have this precious experience in the UK. Thanks also to Kumiko Murata and Toshihiko Suzuki for their insightful comments on issues related to this thesis.

Finally, love and thanks to my family and friends both in Japan and in the UK for their invaluable support throughout the project.

Abstract

This thesis compares conversations between British tutors and British students, and conversations between British tutors and Japanese students, in English in order to investigate differences and similarities in their listenership behaviour in relation to the use of response tokens in the context of academic supervision sessions.

A new method for conversation analysis to synthesise visual data with verbal data on timeline has been established. The concept of *leadtime*, which is a time scale to measure a distance between a point where a particular response token is uttered to a point where a turn transition occurs, has been introduced to implement the research method. Approaches in conversation analysis, roles of context, and intercultural communication are reviewed in this thesis. In addition, participants' assumptions of framing and turn-taking structure in conversation and self-expressions in listenership with reflection of their cultural values in interlanguage settings have been taken into consideration. The results from the preliminary research are summarised as follows:

1. Similarities in use of strategies for framework shifts, such as increase and decrease of response tokens before floor-taking, and multi-functional nature of hand gestures, such as hand gestures used for speaker change and metaphoric signs, have been recognised between the British-British conversations and the British-Japanese conversation.
2. L1 (first language) transfer has been observed in the Japanese students' use of response tokens, such as their constant use of head nods at a particular pace.

These findings highlight areas for further research and application in intercultural communication.

Table of contents

List of abbreviations

CA	Conversation Analysis
CN	Continuers
CV	Convergence tokens
EN	Engaged response tokens
FG	Floor giving
FS	Floor seeker
FT	Floor-taking
HGs	Hand gestures
HNs	Head nods
IR	Information receipt tokens
LS	Listenership
SD	Standard Deviations
SE	Standard Errors
TCU	Turn Construction Unit
TRP	Transition Relevance Place
TTP	Turn Transition Point

List of tables, figures and transcripts

Tables

Figures

Transcripts

Chapter 1 Introduction

1.0 Introduction

In this thesis, the main focus is placed on comparing British-British conversation and British-Japanese conversation in English in relation to their turn-taking structure in a context of academic supervision sessions. Listenership patterns with reference to the use of response tokens will be investigated.

Listenership behaviour has been an objective of linguistic research since the early 1970s (Adolphs 2008, Duncan 1974, Heritage 1997, LoCastro 1987, Maynard 1990, McCarthy 1998, O'Keeffe & Farr 2003, Sacks 1992, Yngve 1970). Although Chomsky's grammarian view was dominant in linguistics at that time, Yngve (1970: 142) showed his interest in functions of discourse and describes the organisation of conversation, in other words *turn-taking*, as follows:

> When two people are engaged in conversation, they generally take turns. First one person holds the floor, then the other. The passing of the turn from one party to another is nearly the most obvious aspect of conversation.
> (Yngve 1970: 567-568)

The concepts of turn-taking and the *floor* of conversation were developed in his work. Based on previous studies, such as the latter, on conversation, the current study attempts to investigate turn-taking structure placing focus on response tokens.

Yngve (1970) also introduces the term *backchannels*. Backchannels have many other names such as *response tokens* (Gardner 2002, O'Keeffe et al. 2007),

listener response (Maynard 1990) and *minimal response* (Fellegy 1995) and the definitions of backchannels vary from study to study (McCarthy & O'Keeffe 2004). In this thesis, I will use the term response tokens by employing O'Keeffe et al.'s (2007) terminology, and take a broad definition of response tokens as described by Duncan (1974), which includes *verbal response tokens* (*right* and *mm*), *non-verbal response tokens* (hand gestures, gaze, nods and silence), and *forms* such as completion of a prior turn. The use of response tokens by listeners of different status and the transition from listener to speaker is central in the current study.

In O'Keeffe et al. (2007: 142), *listenership* is defined as 'the active, responsive role that listeners have in conversation', and response token is a term to refer to 'the many vocal, verbal and non-verbal non-floor-holding devices that a listener may use to respond to the floor-holding message in a conversation'. The reason why it is important to focus on listenership and response tokens is articulated in the following quotation.

> Without response tokens, conversations, even the most business-like and utilitarian ones, would be lacking in terms of the social relationship between speakers. That is, an interaction without response tokens may achieve its goal, but it may not achieve any level of relational bonding between interactants.
>
> (O'Keeffe et al. 2007: 156)

As described in the quotation above, some of the functions of response tokens might be linked with relational and interactional aspects of conversation. McCarthy (2002)

also reports a comparison between response tokens in British and American English in everyday conversation, and concludes by highlighting the importance of good listenership in conversation as social interaction.

> […] 'good listenership' seems to demand more than just acknowledgement and transactional efficiency, and listeners orientated towards the creation and maintenance of sociability and affective well-being in their responses.
> (McCarthy 2002: 69)

Response tokens seem to play a crucial role in achieving good listenership in conversation, which concerns transactional business in conversation and is related to relational/interactional issues. However, not much research has been undertaken on good listenership in relation to use of response tokens, and this study aims to address this by investigating the forms, placement and 'multi-functional nature' of response tokens in relation to their relational/interactional functions in conversation.

This study is based theoretically on the strand of functionalism in linguistics, and in order to situate it in the history of linguistics, key theories in linguistics are reviewed here. In de Saussure's (1972[1983]) major contribution of structuralism and Chomsky's (1971) theory of Universal Grammar, language was studied as an abstract object separated from reality. This trend was altered after the introduction of *communicative competence* as an object of linguistic study by Dell Hymes. Hymes (1974 [1989]) distinguished language structure from language use, and suggests that both of them can be an objective of study in linguistics. In the UK, functional

linguists in the twentieth century, such as Firth (1934), Malinowski (1923) and Halliday and Hasan (1985), developed the idea that language acts and functions in a context. Further, Malinowski (1923), in the early twentieth century, conducted research on a primitive language in an African tribe and raised awareness of the relationship between language and the context which is realised in the culture;

> What I have tried to make clear by analysis of a primitive linguistic text is that language is essentially rooted in the reality of the culture, the tribal life and customs of a people, and that it cannot be explained without constant reference to these broader contexts of verbal utterance.
> (Malinowski 1923: 305)

The stance in linguistic research that language is analysed *within its context* rather than as an abstract objective is important for my research, for it is based on the same premise. Soon after Malinowski's notion of context was introduced, Firth (1934) developed the concept of routine of language use and argued:

> It is true that just as contexts for a word multiply indefinitely, so also situations are infinitely various. But after all, there is the routine of day and night, week, month, and year. And most of our time is spent in routine service, familial, professional, social, national. Speech is not the boundless chaos.
> (Firth 1934: 28)

Following this, it is therefore expected that interlocutors' cultural values and their identities might be reflected in routinised use of language. This premise is considered in my research.

Following the trend of functionalism in linguistics, several approaches to analyse language and context were explored in linguistic research. On the one hand conversation analysts, based at the University of California Los Angels (UCLA), investigated the rules of conversation and established an approach for conversation analysis, placing particular emphasis on the sequence of interaction (Sacks et al. 1974, Schegloff 2007). Units of paired utterances in conversation, referred to as *adjacency pairs*, which are used to analyse these sequences of conversation, include greeting-greeting, request-response and question-answer. Completion and incompletion of such adjacency pairs are integral cues for analysing 'social interaction' (Schegloff 2007) (see Section 2.1.2). On the other hand, discourse analysts, on the other hand, attempted to categorise acts of speech in specific situations, such as classroom interaction (Coulthard 1977, Sinclair & Coulthard 1975, Stubbs et al. 1979), work place interaction (Clyne 2003, Yamada 1997) and casual conversations (Burton 1981, Francis & Hunston 1992). Furthermore, some discourse analysts investigated the *discourse framework*, which is a patterned organisation of conversation (Baker et al. 2001, Tannen 1984), basing such studies on the context of a Thanksgiving dinner (Tannen 1984) and an IT helpline conversation (Baker et al. 2001) (see Section 2.1.3.3).

Following the strand of systemic functional linguistics (Halliday 1978, Halliday & Hasan 1976, Halliday & Hasan 1985), research on the use of language in naturally occurring conversation has been carried out by McCarthy and Carter (Carter

2004, McCarthy & Carter 1994). A model of a social interactional map was introduced by McCarthy and Carter (1994), which classifies contexts into fifteen categories depending on context type and goal type (see Section 2.2.2).

Conversation analysis will be reviewed in detail at a later stage (see Section 2.1), as well the relationship between language and context in the literature review and the discussion section (see Section 2.2 and Section 6.1).

1.1 Intercultural communication

We may hypothesise that Japanese learners of English might transfer their conversation styles and discourse strategies used in Japanese conversation to English conversation, since languages embrace a particular culture, and cultural values are reflected in routines of language use (Maynard 1990, Maynard 1997a, Turner & Hiraga 1996). Learning a language may lead the learner to attempt to integrate themselves within the *speech community* of the target language by adjusting their own cultural identities and routines from L1 (first language) to the target culture. According to Hymes (1972), a speech community is described as follows:

> Tentatively, a 'speech community' is defined as a community sharing rules for the conduct and interpretation of speech, and rules for the interpretation of at least one linguistic variety.
> (Hymes 1972:54)

This means that there is a domain referred to as a speech community, where the use of language is interpreted under certain rules shared by people inside the community. There are, however, occasions where people from different speech communities

encounter each other. One of these instances occurs during language learning in a host community. By learning a second or foreign language, learners are, either consciously or unconsciously, trying to cross the border between the speech community of their first language and the target language(s).

Kasper (1993:3) defines *interlanguage pragmatics* as 'the study of nonnative speakers' use and acquisition of linguistic action patterns in one second language (L2)'. Good communicators know not only grammar or vocabulary but also strategies to 'convey' their intentions effectively in order to establish a good relationship with participants in conversation. Kasper (1993: 10) recognises such pragmatic transfer from language learners' first language to a second language, and distinguishes *positive transfer* from *negative transfer*. In her definition, positive transfer is not a problem since 'pragmatic behaviours or other knowledge displays consistent use across L1 [first language], IL [interlanguage], and L2 [second language]', however, negative transfer might cause 'risk to communicative success' because of 'the influence of L1 pragmatic competence on IL pragmatic knowledge that differs from the L2 target'. I will consider both negative and positive transfer in interlanguage communication in relation to turn-taking structure in the current study. The differences and similarities in spoken discourse in British English and Japanese, and to what extent those differences might affect Japanese learners' being successful users of English are central to my interest. 'Successful users of English' (Prodromou 2005) here refers to how successful speakers can construct contexts and identities as well-balanced language users, rather than referring to how they approximate their conversation styles to the target language. Kramsch (2008) refers to this ability as *symbolic competence* (see Section 6.2.2).

In order to analyse listenership in British-British conversation in comparison with British-Japanese conversation, intercultural communication and interlanguage pragmatics will be considered in this study, and emphasis will be placed on turn management strategies and the cultural rationales behind them. For data, two types of face-to-face dyadic conversations in English have been video-recorded and analysed with time-based multimodal annotation software:

1. British tutor-British student dyad conversation,
2. British tutor-Japanese student dyad conversation.

Native speakers of British English and Japanese advanced learners of English are targeted in this thesis. The terms 'native speakers of English' and 'standard British English' raise controversial issues (Pennycook 2001, Pennycook 2006, Prodromou 2005). Pennycoock (2009) introduces three dimensions of English use: (1) *language resource* as one of the 'communicative repertoires', (2) *language context* as English use 'in time and space', and (3) *speaker location* as 'a language connected to certain desires and ideologies' (ibid: 204-205). Use of English is treated as the second dimension in the current study. Native speakers of British English here refer to people who grew up and spent most of their lives in the UK using British English as a medium of communication. Differences and similarities in strategies of placing response tokens in order to initiate speakership in these two dyad conversations are compared.

1.2 A time-related corpus-based approach

I developed a new multi-modal framework for analysing active listenership which

includes verbal and non-verbal components. The two areas of the current study which make original contribution are as follows. The first area analyses physical movements in conversation alongside verbal utterances, and, to this end, I will examine the use of visual response tokens, such as hand gestures and head nods, in addition to vocal signals; the second area in turn is the development of a time-related corpus approach.

Original contributions:

a) Analysis of visual data and audio data: analysing visual response tokens, such as head nods and hand gestures with verbal response tokens;
b) A time-based approach for analysing turn structure: time-related transcripts of conversation data will be used for analysis. The concept of *leadtime*, will be introduced and applied for analysing floor exchanges with the use of response tokens in relation to the timeline.

By using time-related data, I have conceptualised a timescale to analyse turn-taking structure. This scale will be referred to as leadtime.

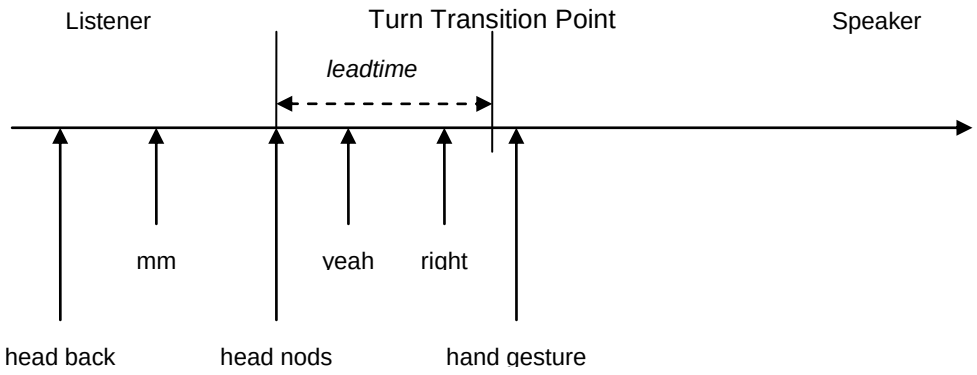


Figure 1.2-1 The concept of leadtime

On a continuous timeline, a particular interlocutor has periods of being a listener and periods of being a speaker, and various types of response tokens are allocated on the timeline in reference to a turn transition point from listener to speaker, as illustrated above. A *turn-transition point* (TTP) can be differentiated from a *transition-relevance place* (TRP): a TRP is a possible place for turn transition (Duncan 1972, Sacks et al. 1974, see Chapter Two), whereas a TTP is a point where actual turn exchanges occur.

From this assumption, I have developed the concept of leadtime. Leadtime is applied to both listener status and speaker status in order to measure the length of time of a speaker/listener status with the turn transition point as a datum point. The datum point is described as 0 in leadtime. Leadtime is used to describe the time distance between the point where a response token or a discourse marker is used and the floor transition point (see Section 3.1.6).

With this concept of leadtime, a model for analysing turn-taking will be suggested. In previous research on analysing turn-taking, no timeline was used, and therefore the length of time of the speaker/listener status and the time distance between a response token and turn-taking point were not able to be measured. This new model fills a gap between the limitation of previous methods and the current needs for analysing turn-taking with the concept of time.

Although the main focus of the current research compares British-British conversation and British-Japanese conversation to investigate differences and similarities in their turn-taking structure from a cross-cultural aspect, there is a gap between the level of current knowledge of conversation styles available to us in existing research and the level of knowledge to discuss cross-cultural issues in relation to turn-taking structure. At this point, we do not even know which aspects of

the interactions are relevant in co-constructing speaker/listener transitions, which means that we cannot begin to compare cross-cultural interactions of this type in a meaningful way. For this reason, an exploratory study comparing British-British conversation and British-Japanese conversation will be conducted, which helps to build a bridge across the gap and suggest directions of future research.

In order to investigate differences and similarities in the use of vocal and visual response tokens in relation to floor exchanges, it is necessary to establish research methods for conversation analysis placing focus on listenership. Generating and implementing research methods for conversation analysis with time-related data will be one of the main areas to be explored during the course of the current study.

1.3 Research question and research design

The main study question in this thesis is as follows.

> What are the differences and similarities between British-British conversation and British-Japanese conversation in English in a context of academic supervision sessions?

There are three aspects used to compare British tutor-British student conversation and British tutor-Japanese student conversation in the current study:

> Aspect 1: Turn-taking structure,
>
> Aspect 2: Use of verbal response tokens,
>
> Aspect 3: Use of visual response tokens.

In order to address the main study question, features of turn-taking structure with use of verbal and visual response tokens, such as *erm, yeah, mm, mhm, head nods* and *hand gestures*, will be analysed. British-British conversation and British-Japanese conversation are compared in order to analyse differences and similarities in their turn-taking structure from an aspect of intercultural communication. As a prior task, a method for analysing floor exchanges needs to be established (see Section 1.2).

To address the current study question, an exploratory study of comparing British-British conversation and British-Japanese conversation will be designed with the following goals:

a) To establish a method for analysing floor exchanges to compare British-British conversation with British-Japanese conversation by introducing time-related transcripts with multimodal annotations,
b) To indicate some preliminary results based on the analysis on use of vocal/visual response tokens in relation to turn-taking structure,
c) To highlight areas for further research.

These aims will be addressed through the following steps:

a) A comparison of the numbers of floor exchanges and floor length between British-British conversation with British-Japanese conversation in English;
b) A comparison of the numbers of vocal and visual response tokens, such as *erm, yeah, mm, mhm, head nods* and *hand gestures*, in British-British

conversation with British-Japanese conversation in English in reference to the timeline;

c) Qualitative analysis of response tokens in order to identify turn-structural episodes in British-British conversation, and compare the turn-structural episodes with those in British-Japanese conversation in English.

The current study is new and innovative in that it uses a time-related corpus-based approach. A corpus is ‘a collection of pieces of language, selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language’ (Sinclair 1996). In the early stage of corpus linguistics, corpora of written language were mainly used in research on forms and lexical items for pedagogic purposes (Biber 2006, Garside & Leech 1982, Renouf 1984). Corpus analysis with spoken language data has been applied to pragmatic research (Adolphs 2008, Adolphs et al. 2004, O'Keeffe & Adolphs 2008, Stubbs 1996) and integrated with discourse and conversation analysis recently (Biber et al. 2007, Evison & McCarthy 2010, Handford 2007, O'Keeffe 2004, O'Keeffe 2006, Tao 2003). This current research integrates a corpus-based study of spoken English with conversation analysis.

Due to the improvement in information technology in the past two decades, there are many software programmes available for linguistic researchers to investigate language use with multimodal annotations and transcripts along a timeline. The current research would not be possible without the aid of this technology. This is one of the reasons why the current study has been realised now. Two multimedia annotation software packages are considered in the current research; Transana, which is a conversation analysis software written by Chris Fassnacht at the University of

Wisconsin, Madison, and Digital Replay System (DRS), which has been developed in the School of Computer Science & IT at the University of Nottingham, respectively. DRS in particular allows a researcher 'to replay, manage, annotate and visualize that [time-based] data' (French et al. 2006).

Based on the transcripts annotated and time-stamped by these annotation tools, a time-related multimodal corpus is here developed, which includes not only audio data but also visual data in reference to a timeline. There are three reasons for using a time-related multimodal corpus analysis as a central method in the current study project; (1) authenticity of the data, (2) ease of analysis with data searchablility, and (3) capability of analysing both audio and visual data synchronously (see Section 3.2.2).

1.4 Overview of the study

This thesis is divided into seven chapters. Following the introduction in the present chapter, existing research and theories on conversation analysis, context and genre, response tokens, and intercultural communication are explored in Chapter Two. This will include an overview of methods in conversation analysis and features of response tokens in English conversation and Japanese conversation.

Chapter Three deals with research methodology and with establishing a time-related corpus-based approach, as well as conducting the pilot study with two 10-minute conversations: a British tutor-British student conversation and a British tutor-Japanese student conversation. The concepts of TTP and leadtime are described in detail in Chapter Three. With these concepts, elements and alignments of the time-related transcripts are developed for the pilot study. Findings from the pilot study are also reported in Chapter Three.

Chapter Four links the pilot study to the main study. The main study is an extension of the pilot study. The scope of the main study is defined and two approaches to be taken in the main study are described; global pattern analysis and turn-structural analysis. In the global pattern analysis, length of speaker status and the numbers of targeted response tokens are examined quantitatively by implementing the concept of leadtime. The turn-structural analysis examines the features of turn-taking in reference to forms and placement of response tokens.

Chapter Five draws on the time-related spoken corpus data established in the previous chapter and reports the findings from the data analysis in the main study. For the scalability of the research, four 39-minute sets of conversation data are analysed in the main study with the method developed through the pilot study. Placements of targeted response tokens are summarised in the global pattern analysis. Seven turn-structural episodes based on Ohama (2006) are introduced in the turn-structural analysis. The use of these turn-structural episodes by each participant is investigated.

Chapter Six extends the discussion of the findings from the main study with the rationales and possible theoretical interpretations of the participants' choice of response tokens and the use of turn-structural episodes in the British-British conversations and the British-Japanese conversations. The concepts of contextualisation and multiple identities are considered with reference to interlanguage pragmatic perspective.

This thesis ends with a summary of the research and the limitations of the research. Areas highlighted from the research will also be discussed with relevance to possible research.

Chapter 2 Literature review

2.0 Introduction

In this chapter, a literature review will be conducted on four issues in linguistic theories: (1) conversation analysis, (2) context and genre, (3) response tokens and listenership, and (4) intercultural communication. Some key concepts of CA, such as the floor of conversation, TRP and *speaker selection* were applied to the analysis in the current study. It is assumed that participants' language use is affected by the context where conversations occur and, at the same time, that the participants construct the context with their language use. This two-way relationship is also taken into consideration in the current study, and based on previous research on functions and placement of response tokens, turn-taking structure is analysed using the new research model with leadtime. The participants' use of turn taking structure will also be discussed from perspectives of intercultural communication at a later stage (see Chapter Six).

2.1 Conversation analysis

This chapter will review previous research on conversation analysis (CA), which was developed in the early 1970s by Harvey Sacks with Emmanuel Schegloff and Gail Jefferson (Heritage 1984b). CA is broadly defined, as 'the study of talk and other forms of conduct (including the disposition of the body in gesture, posture, facial expression, and ongoing activities in the setting) in all forms of talk in interaction' (Schegloff et al. 2002: 3). Heritage (1984b) defines CA with emphasis on ordinary speakers' language use:

> Conversation Analysis – like the other research streams of

> ethnomethodology – is concerned with the analysis of the *competences* which underlie ordinary social activities. Specifically it is directed at describing and explicating the competences which ordinary speakers use and rely on when they engage in intelligible, conversational interaction.
>
> (Heritage 1984: 241)

In CA, authentic conversation data is analysed in order to give a sociological explanation for conversation interaction among speakers.

In an article titled 'A simplest systematics for the organization of turn-taking for conversation' (Sacks et al. 1974), features of turn-taking in conversation were described. Since then, CA has emerged as a prevalent and developing discipline. In Sacks et al., attempts were made to investigate a system and features of turn-taking which is assumed to be a 'basic' organisation for sequence in conversation (ibid: 700), and, based on six years of observation of 'tape recordings of natural conversation' (ibid: 698), fourteen items which describe the characteristics of turn-taking were listed as follows:

1. Speaker-change reoccurs, or at least occurs
2. Overwhelmingly, one party talks at a time
3. Occurrences of more than one party talks at a time are common, but brief
4. Transitions (from one turn to a next) with no gap and no overlap are common. Together with transitions characterized by slight gap

or slight overlap, they make up the vast majority of transitions

5. Turn order is not fixed, but varies
6. Turn size is not fixed, but varies
7. Length of conversation is not specified in advance
8. What parties say is not specified in advance
9. Relative distribution of turns is not specified in advance
10. Number of parties can vary
11. Talk can be continuous or discontinuous
12. Turn-allocation techniques are obviously used. A current speaker may select a next speaker (as when he addresses a question to another party); or parties may self-select in starting talk
13. Various 'turn-constructional units' are employed; e.g., turns can be projectedly 'one word long', or they can be sentential in length
14. Repair mechanisms exist for dealing with turn-taking errors and violations; e.g. if two parties find themselves talking at the same time, one of them will stop prematurely, thus repairing the trouble

(Sacks et al. 1974:700-701)

The first and second items describe the simple fact that turn-taking occurs or reoccurs in conversation, and that the participants in conversation generally occupy both a speaker role and a listener role in turn. The third and fourth concern the notions of *overlap* and *pause* in conversation. From the fifth to the tenth item, the arbitrary nature of turns in terms of the order, size, length, meanings and participations are mentioned. In the eleventh item, the opening and closing of turns are recognised and

the linguistic systems of turns at speaker change are the focus of the twelfth item. In the thirteenth item, the term *turn-constructional unit* (TCU) is introduced, which are organised by various linguistic aspects such as grammar (sentence and clause), phonetic units and 'a recognizable action in context' (Schegloff 2007: 4). The final item describes one of the features seen in turn-taking, namely *repair*, which is taken as a remedy of problems in conversation (Hosoda 2006, Schegloff 1992, Schegloff 1987). Studies of the functions of these processes of repair and *expansion* in conversation are reviewed in Section 2.1.3.1., and are applied in relation to turn-taking structure in the analysis of this study.

Ten Have (2001) attempts to separate Sacks's CA, which is referred to as 'pure' CA, from the application of the findings of 'pure' CA for studying institutional interaction. The latter is referred to as 'applied' CA (Ten Have 2001:3); a category which this study falls within. Conversational interactions in a pedagogic context will be investigated based on the findings of CA and the notion of context in previous research.

2.1.1 Turn-taking organisation

Turn organisation, turn-taking and TCU are taken as fundamental elements of conversation and described in detail in Schegloff's (2007) work. As discussed briefly above, a TCU is a part of speech which can be a *dependent* turn. According to Schegloff, there are three elements which enable us to recognise TCUs; grammar, a phonological unit, and 'a recognizable action in context' (ibid: 4). A single turn often includes 'more than one TCU' (ibid), and, further, the point of transition to the next speaker is referred to as TRP. Two instances of the turn transition are described as follows:

> First, a just-prior speaker can have selected them as next speaker by addressing them […]. Second, if no one has been so selected by a/the prior speaker, then anyone can self-select to take the next turn and does that by starting to fashion a first TCU in the turn-space […].
>
> (Schegloff 2007: 4)

The question of how people initiate or terminate turns in conversation has also been investigated by Sacks, Schegloff and their followers.

Schegloff (1987) highlights a tendency for turn beginnings to be recycled in spoken English conversation, as shown in the example below:

> R: Well the uhm in fact they must have grown a culture, you know, they must've – I mean how long – he's been in the hospital for a few days, right?
>
> Takes a | bout a week to grow a culture
>
> K: | I don't think they grow a I don't think they grow a culture to do a biopsy.
>
> (Schegloff 1987: 75)

He states that a 'precise relationship' can be observed between the overlap of the prior and the new turn and the recycled turn beginning (ibid). In other studies, Antaki (2002) investigated telephone conversations and found that the turn initial items with 'high-grade assessments' (ibid: 5) such as *lovely* and *brilliant* have the function of closing the sequence; and Lerner (2002: 226) reports people's use of *turn sharing*, in

other words ‘choral co-production’ of a turn, which is one of the characteristics observed in turn-taking.

8 B: you know what my teacher was gone

9 for a week, she went

10 (.)

11 [she’s in the hospita:l

12 → C: [(she’s in the) hospital:l

13 D: mmm, mm

14 B: she has an [operation

15→ C: [(opera)tion

(ibid)

In the transcription above, choral co-production can be observed in line 12 and line 15, which can be used ‘to exhibit understanding, affiliation, and agreement with a current speaker’ (ibid: 250). Not only co-production of words or a part of a word but also ‘gestural matching’ was noted in the study, which serves to emphasise the listener’s understanding of what the current speaker is saying (ibid: 245).

2.1.2 Sequence organisation

In ‘Lectures on Conversation’, Harvey Sacks (1992: 254) introduced two rules for conversation: ‘one party talks at a time’, and ‘sequencing’ of turns, respectively. By articulating these rules, he does not mean that people never observe more than one party talking at a time, rather that ‘A-B-A-B would characterize any two-party conversation as a natural law’ and that ‘a good deal of the time one party only is

talking , the other party is listening' (ibid). Following on from these rules, several techniques to keep a sequence of conversation have been observed and described by Sacks.

2.1.2.1 Speaker-selection

Sacks (1992) introduced two techniques for next speaker-selection. One is that the current speaker selects a next speaker by addressing their names or questioning; a technique referred to as *other-selection*.

A: Are you fed up?

B: To the neck

(Sacks 1992: 676)

In the transcription above, A is asking a question of B and by doing so selects B as a next speaker. Sacks refers to this kind of string of utterances where the first pair requires the second pair to complete the sequence of conversation as *paired-utterance classes*, which is nominalised as an adjacency pair by Schegloff (2007). The notion that the first pair of paired utterance classes requires the second pair to complete the sequence is important in order to understand the *second* technique for next speaker selection. The details of adjacency pairs will be discussed further in Section 2.1.3.

The second technique for the selection of next speaker is referred to as *self-selection*. People can select themselves as the next speaker in conversation under certain conditions: for example, when the previous speaker asks a question without selecting a next speaker, a participant can choose himself or herself as a next speaker by his or her utterance (Sacks 1992: 676). These two types of speaker selection were

applied to the analysis in the current study.

Sacks also raises awareness of silence and eye-monitoring although he treats them as 'rather small issue[s] with regard to selection' (Sacks 1992: 672) and only gives a brief description on these areas. These areas, however, have recently received a lot of attention from linguists, and the relationship between silence and initiating a turn at the transition point from listener to speaker is one of the main issues to be investigated here.

2.1.2.2 Turn and floor

Through the consideration of multi-party conversations, moreover, Sacks (1992) recognises the phenomenon of the floor which he differentiates from a turn in conversation. Furo (2001) also studies floor control and divides turns into two types, *floor-taking turns* and *non-floor-taking turns*. Non-floor-taking turns include six types: (1) response token, (2) reactive expressions, (3) repetitions, (4) collaborative finishes, (5) laughter, and (6) short statement.

Sacks (1992) introduces the term *floor seeker* to refer to actions of participants who are trying to take the floor of the conversation. Sacks points out that particular sentences for story telling and general sentences in a particular situation can be floor seeking. A story telling sequence, such as 'I was at the police station this morning', and a general sentence, such as 'You know what happened to me last night?' can function as a floor seeker (Sacks 1992: 680). This function of floor seeking is assumed to be one of the features related to turn-taking structure. Analysis using the concept of floor seeking will be conducted in the pilot study in Chapter Three.

2.1.2.3 Two-party and multi-party conversation

When Sacks (1992) discusses speaker selection, multi-party conversation is central to his argument. He states that speaker selection techniques will be a 'small issue' in two-party conversation since the second pair parts of paired-utterance classes will be answered by the other speaker. On the basis of my analysis, this is only partly true. This may or may not be the case since participants in a conversation do not always follow a defined order of floor exchanges even in a dyad conversation. If we imagine a situation where the current speaker tries to keep the floor of the conversation in a dyad conversation, and at the same time the other participant is seeking the floor to give his or her opinion, it can be said that speaker selection becomes an important issue even in a two-party conversation. In the current study, two-party conversations will be targeted partly to simplify the research condition by reducing the number of participants.

2.1.2.4 Institutional talk and interactional asymmetries

Heritage (1997: 236) examines some institutional interaction such as the discourse between doctors and patients, sales persons and customers, and teachers and students, and highlights four types of asymmetries in institutional talk: (1) participation, (2) 'knowhow' about the interaction and institution in which it is embedded, (3) knowledge, and (4) rights to knowledge. In terms of participation, Heritage (1997) finds that there is a contrast between 'the symmetries of ordinary conversation and the asymmetries of institutional discourse' (ibid: 237), which means people have equal participation in daily conversation more often than in institutional or professional interaction. This kind of situation can be observed in institutional talk since a participant who knows 'the specific institutionality of interactions' (ibid: 236) in a

particular context more than the other participants will take the initiative in conversation. Therefore participation in the conversation tends not to be equal for every participant. Such differences will be evident when the data in the current study is analysed in Chapter Three.

2.1.3 Actions in turn-taking

Emanuel Schegloff (2007: 9) defined 'sequences' in conversation as 'courses of action implemented through talk'. Sequence organization is defined as follows:

> […] the organization of courses of action enacted through turns-in-talk – coherent, orderly, meaningful successions of 'sequences' of actions or 'moves'. Sequences are the vehicle for getting some activity accomplished.
>
> (Schegloff 2007: 2)

A sequence of conversation is constructed by completion and incompletion of adjacency pairs, which are are thought to be the minimum unit to build up courses of action in conversation. Adjacency pairs are :

> (a) composed of two turns
>
> (b) by different speakers
>
> (c) adjacently placed; that is, one after the other
>
> (d) these two turns are relatively ordered; that is, they are differentiated into 'first pair part' (FPPs, or Fs for short) and 'second pair part' (SPPs or Ss for short) […]

> (e) pair-type related; that is, not every second pair part can properly follow any first pair part. Adjacency pairs compose pair 'types'; types are exchanges such as greeting-greeting, question-answer, offer-accept/decline, and the like. […].
>
> (Schegloff 2007: 13)

Even in a simple interaction like the conversation shown below, an adjacency pair in a greeting-greeting type can be observed.

> A: Hello doctor.
>
> B: Hello Anna. What can I do for you tonight?
>
> (From British National Corpus, Davies 2004)

What A is saying can be recognised as a FPP, and the SPP by B follows A's utterance. In this study, attempts to conceptualise the discourse organisation of conversation above the level of adjacency pairs will be made based on Sacks and Schegloff's theories in conversation analysis.

2.1.3.1 Expansion and repair

Schegloff (2007) defines several types of expansion, such as *pre-expansion*, *insert-expansion* and *post-expansion*, depending on the differences in the location of an extended sequence which is added to 'a base adjacency pair'. The transcript below shows an example of insert-expansion.

> 1 Bet: Was last night the first time you met Missiz Kelly?

2 (0.1)

3 Mar: Met whom?

4 Bet: Missiz Kelly.

5 Mar: Yes.

(Schegloff 2007: 97)

In line 3, Mar, the second speaker, extends the sequence by using a question 'Met whom?' According to Schegloff, expansions can also be seen as *repair* in conversation, a term which can be described as 'efforts to deal with trouble-sources or repairables – marked off as distinct within the ongoing talk' (ibid: 101). This concept of expansion and repair will not be central in the current research; however, they will be worth noting the issue as a feature of sequence organisation.

2.1.3.2 Discourse marker *oh* and expansion

As an example of the use of discourse markers in relation to sequence, the functions of *oh* will be reviewed here. Schiffrin (1987) investigates the functions of discourse markers such as *oh*, *well*, and *you know*, and states that the discourse marker *oh* has functions of information management such as repairs, acknowledgement and recognition although '*oh* is traditionally viewed as an exclamation or interjection' (ibid: 73). Heritage (1984a) also considers the functions of *oh* in relation to sequencing in conversation and introduces the term *change-of-state token* which expresses the speaker's shift from 'non-knowing to now-knowing' (Schegloff 2007: 118). Schegloff (2007) also points out another function of the change-of-state token *oh* in relation to a sequence:

> By registering a state-changing receipt of information, free-standing *oh* can serve as a possible sequence-closing, third position turn. (Schegloff 2007: 119).

As Heritage and Schegloff observe, the discourse marker *oh* has a contribution to sequence organisation and the free-standing *oh* works as closure of sequence. Finally, the notion that a particular discourse marker functions as a specific sequence organiser will be important here since it can be a cue to investigate the techniques used at the shift from listener to speaker in conversation.

2.1.3.3 Discourse framework

Baker, Emmison and Firth (2001) analyse the organisation of calls to a software helpline, which is contextualised in a specific institutional context, and extract features of their interaction. Here is a sample conversation taken from the software helpline calls:

1	CT	how can I help you?
2		(0.4)
3	C	erm, I've installed (.) office ninety-seven?
4		(0.8)
5		.hh and (.) erm my negative figures are different
6		(0.8)
7		In excel (.) from this time=an' I think it's
8		somewhere in the setup that I haven't-
9		(0.8)

10 selected something

11 CT the- when you have negative numbers in your cells? (.)

12 erm how are they displayed?

* CT= the call taker, C=the caller

(Baker, Emmison & Firth 2001:45)

Based on the transcription of the helpline calls, a typical sequence of the call opening in software helpline encounters is described as follows:

CT [how can I help you]

C [.hh erm]

C [I've been installing product x]

CT [+/- yeah, okay]

C [and + specific domain of y]

CT [+/- yeah, okay]

CT [and/but]

CT [something is happening that should not happen]

CT [+/- substantive comment or question]

(Baker, Emmison & Firth 2001:53)

Not only are the findings from the study but also the methods for the analysis pertinent to this study. This is one of the methods to analyse turn-taking structure apart from CA, and will also be considered in this study. A chronicle order of the elements in IT helpline was described in the study.

Saft (2007) studies conversation in a Japanese faculty meeting in a university and reports that there were two frameworks apparent in the faculty meeting: a reporting framework and a discussion framework. In the transcript below, the participants of the faculty meeting have a discussion concerning the equipment they have bought and how they might ensure its maintenance. The chair is reporting the issue in the first few lines, and then Kida, one of the participants, cuts in with *iya iya sore ga ne* in line 16, which can be translated as *no no the thing is*.

12 Chair: *gakumuka da tte kanri suru na n tte koto wa*
school affairs COP QT manage COP NUM QT thing TM

13 *dare mo ittenai yo tte koto na n de(.) soko de*
nobody not say FP QT manage COP NOM COP that place LOC

14 *chotto takano-san to(.) suttamonda shuchatta n desu*
little Mr. Takano with big fuss did NOM COP

15 *kedo tonikaku[*
but anyway

16 Kida: *[iya iya sore ga ne*
no no that S FP

17 Chair: *un*
uh-huh

18 Kida: *gakumuka no yosan de kau mono wa ano=*
school affairs LK budget COP buy thing TM that

(Saft 2007: 48-49)

Keys: Utterances in Italic = original Japanese transcript, COP = various forms of copula verb be, QT = quotative marker, TM = topic marker, FP = final particle, NOM = nominalizer, LOC = locative, S = subjective marker, LK = linking nominal-occurs between two nouns

Kida's utterance in line 16 is followed by the chair's response token *un*, which is an equivalent of *uh-huh*. This response token leads the shift from a reporting framework

to a discussion framework in the faculty meeting. In a reporting framework, the chair is 'reporting' the matters to share dominantly in the meeting, while in a discussion framework participants in the meeting are involved in discussions on the issues raised. Here, the use of response tokens plays an important role in the transition from reporting framework to discussion framework. This discourse framework analysis method has been taken as a model of conversation analysis in the current study.

2.1.4 Preference in turn-taking

Sacks (1987) reports on preference of agreement in the second pair part, meaning that people tend to avoid direct rejection of a proposal or offer in conversation. In the conversation below, B could answer *no* instead of *that's where I was born*, however, B chose the utterance to show an 'agreement response':

A: That where you live? Florida?

B: That's where I was born.

(Sacks 1987: 60)

Drew (1984) also investigates one particular feature of discourse, *reportings*, and finds that reportings have several functions. People tend to use reportings, for instance, when they want to reject a proposal or offer an invitation, as in the following exchange:

I: How about the following weekend?

(0.8)

C: That's the vacation, isn't it?

I: All right

(Drew 1984:134, transcription is simplified by me)

When people wish to decline an invitation or reject a proposal, people tend to report circumstances or activities as reasons for declining or rejecting (ibid: 146).

Even in daily conversation as in the example above, people are continuously making decisions on how to articulate their intention with more 'preferred' turns, words, tones, and gestures in the context they are in – either consciously or unconsciously. The design of turn-taking will also be an element to be considered in terms of the organisation of a conversation in a 'preferred' manner. It might be true that people prefer agreement over disagreement when they initiate turns in English conversations, for it is assumed that people in a discourse community tend to choose agreement and avoid disagreement more than other discourse community. This is not only a matter of 'direct and indirect' speech but also a matter of how to initiate turns to organise a smooth sequence in a particular discourse community.

In CA, 'preference' in sentence structure or expressions to decline offers/proposals has been studied (Drew 1984). I shall take preference in conversation as a broad meaning including turn exchanges, such as timing of taking turns or length of speaker/listener status. How people manage turn exchanges with their preferences will be investigated in the current study.

2.1.5 Conversational gestures

A review of research on conversational gestures will be described in this section since hand gestures and head movements are a further focus of the current study. In terms of hand gestures, Goldin-Meadow (1999) highlights the following characteristics:

> Gesture provides speakers with another representational format in addition to speech, one that can reduce cognitive effort and serve as a tool for thinking. Gesture also provides listeners with a second representational format, one that allows access to the unspoken thoughts of the speaker and thus enriches communication.
>
> (Goldin-Meadow 1999: 428)

Goldin-Meadow (1999) categorises hand gestures according to their functions. There are four types: *iconic gestures*, which describe a picture that the speaker has in mind such as pouring water into a glass; *metaphoric gestures*, which are more abstract than iconic gesture and describe speakers' thoughts or idea; *beat gestures*, which can be used to emphasise what the speaker is saying 'along with the rhythmical pulsation of speech'; and *deictic gestures*, in other words, pointing gestures. Although the categorisation was based mainly on speakers' hand gestures, *listeners*' hand gestures can also be considered as conversational gestures.

Schegloff (1984) raised awareness of the importance of gestures in conversation and analysed the functions of hand gestures in conversation. Although gestures are normally used by speakers in conversation to support their verbal description of an idea, Schegloff (1984: 271) reported three types of hand gestures used by nonspeakers: (1) to show intention to be a next speaker, (2) 'in lieu of talk' which is used by the listener to communicate without interrupting the current speaker, and (3) to interrupt the current speaker. The issue he raises out is significantly related to turn-taking organisation: according to Schegloff, gestures can be used for initiating

a turn by listeners and taking back a turn from an interrupter, which can be interpreted as gestures functioning as floor seeker. Moreover, gestures can be a kind of TCU since gesture can be used to communicate the listener's intention, although it is not verbal but visual communication. Kendon (1972) also investigates the features and functions of conversational gestures.

> Most of the work on the relation between body motion and speech has been concerned with how body motion may express aspects of what the speaker is saying […], how it may express additional usually unconscious thoughts or feelings. It would appear, however, that a prior task should be a description of how body motion that accompanies speech is organized, and how it is related to the patterning of speech.
>
> (Kendon 1972: 179)

Kendon (1972) categorises combinations of hand gestures and head movements in great detail, such as 'forearms rotates, fingers extd [extend] and abducted' and 'left forearm raised, lowered, palm open', to match each motion with each speech unit. The total length of data was about one and a half hour, and only one participant's utterances and movements were filmed for the research. The data, however, is limited in size and the matching of the visual and audio data was carried out only by observation. As Kendon also points out, the research in describing the relationship between body motion and the speech unit was successful, whereas the integration between the body motion and the flow of speech based on time could not be analysed.

Further precise analysis might have been difficult without the more developed technology we currently have: for example, with multimodal annotation software packages, such as Transana and DRS (see Section 3.1.4), body movements can be analysed alongside verbal data. A time-related corpus approach was therefore applied to the current study.

In a recent study using modern technology, Davies and Vaks (2001) reported an interactive computer project for detecting people's head gestures such as head nods and head shakes. A real-time face detection tool, the IBM PupilCam, was used for the research and succeeded in monitoring face movements which were signalled back to the computer. Knight et al. (2006) also reported on the HeadTalk project, where the research focus was placed on 'the characteristics of a specific "semiotic channel"; that of head nods' (ibid: 2). The relationship between head nods and response tokens was investigated with video-recorded data and a detection tool for head nods.

Four roles of response tokens in discourse, namely continuers, convergence tokens, engagement response tokens, and information receipt tokens, are described by O'Keeffe and Adolphs (Knight et al. 2006, O'Keeffe & Adolphs 2008).

> (1) Continuers [CN]: Maintaining the flow of discourse.
>
> (2) Convergence tokens [CV]: Markers of agreement/convergence.
>
> (3) Engaged response tokens [EN]: Markers of high engagement where addressee(s) respond on an affective level to the content of the message.
>
> (4) Information receipt tokens [IR]: Markers of points in the discourse where adequate information has been received.
>
> (O'Keeffe & Adolphs 2008 : 84)

In this investigation, the relationship between the length of head nods and their functions were analysed. The findings show that short head nods, rather than long head nods, function as response tokens, which tend to have additional discoursal meanings. This project shows a new approach to analysing the relationship between body movements and functions of response tokens with modern technology; for there appears to be limitless potential in the use of technology for analysing conversational gestures, and therefore it is important to develop protocols of how technology is to be used for this purpose. Establishing appropriate research methods using modern technology effectively will be an increasingly necessary task for researchers in the future.

2.1.5.1 A classification of conversational gestures

Conversational gestures need to be categorised for the current study since not only verbal data but also visual data are analysed. There are various ways to categorise visual data, such as body movements in conversation. One can differentiate, for example, shallow head nods and deep head nods according to how much people move their head (Knight et al. 2006, Knight et al. 2009). One can, as another example, divide the types of hand gestures into more than twenty clusters (Kendon 1972, Kendon 2004) depending on how much space people use to produce the hand gestures, or according to the functions of the hand gestures. In the current research, however, I classify conversational gestures into six basic types; head nods, head shake, head turning, head back, hand gestures and self-comfort as shown in the table below. Foot movement and facial expression were removed from the research focus here since the video-recorded data I used for the current study did not sufficiently capture both participants' foot movement and facial expression. Other conversational gestures were

annotated in the pilot study although the focus was narrowed in the main study.

Table 2.1.5-1 A classification of conversational gestures

Conversation gestures	head nods	Vertical head movement
	head shake	Horizontal head movement
	head turning	Head is moving towards speaker
	head back	Head is moving back from speaker
	hand gesture	Hand gestures help verbal description
	self-comfort	Hand gestures show self-comfort ex) scratch hair, hug oneself by arms

2.1.6 Transcribing conversation

Ten Have (2001) states the importance of transcription in CA by exemplifying a transcription as a kind of translation.

> […] a transcription might be seen as a *translation*, made for various practical purposes, of the actually produced *speech* into version of the standardized *language* of that particular community, with some selective indication of the actual speech production.
>
> (ten Have 2001: 76)

Various kinds of transcription conventions have been developed to describe not only spoken words but also *uttered* sounds, spaces/silences, overlapped speech and sounds, pace, stretches, stresses, volume and metadata in talk noted by transcribers. Using those conventions, a conversation may be transcribed as below:

1 Maude: I says well it's funny: Mizssi:z uh:↑Schmidt ih you'd
2 think she'd help<.hhh Well (.) Missiz Schmidt was the
3 one she: (0.2) assumed respo:nsibility for the three
4 specials.
5 (o.6)
6 Bea: Oh↓*::,°°M-hm,°°=
7 Maude: =Maybe: ↑told me this.

(ten Have 2001: 90)

In the transcription here, the underscore expresses stresses, and the colons show stretches of the sounds, whereas the arrow in line 1 indicates 'higher pitch' and the one in line 6 indicates lower pitch. 'M-hm' in line 6 is denotes a quieter sound than the surrounding talk, and 'Maybe' in line 7 is following the prior utterance without a gap.

More simplistic transcription conventions are applied to CANCODE (Cambridge and Nottingham Corpus of Discourse in English) since CANCODE is an orthographic transcription without prosodic information.

<$01><$02>	Speaker codes. Each speaker is numbered with separate codes.
[ ]	Extralinguistic information. This includes laughter, coughs.
+	Interrupted sentence.
=	Unfinished words.

(Adolphs 2006: 134-135)

In the current research, transcriptions will be modelled on the conventions of the

CANCOCDE data since 'prosodic' features will not be central. Functions of discourse and patterns of turn-taking structures will also be annotated in the transcripts (see Section 3.2.2).

2.2 Conversation in context

This section looks at the relationship between language and context by reviewing previous research on context. In my research, conversations in academic tutorial settings are explored with emphasis on their use of response tokens in relation to turn-taking structure. The section aims to clarify where the context chosen and the object of my research fit in the theoretical scheme established by the strand of linguistic research on context.

In discourse analysis and systemic functional linguistics (SFL), context is conceptualised as a pre-established 'bucket', which configurates participants' actions. On the other hand, CA treats context as 'both the project and product of the participants' own actions and therefore as inherently locally produced and transformable at any moment' (Drew & Heritage 1992: 19). The current study is positioned somewhere in between these two strands. Context is viewed as a social and cultural container which affects interlocutors' behaviour in conversation, and also as a renewable entity through interactions.

2.2.1 Context, genre and register

Halliday and Hasan (1985) introduce the term *context* to refer to the setting of conversation. Context in SFL is divided into three components; field, tenor and mode. In addition, Hasan (1985) introduces the concept of *contextual configuration*, which is a system to regulate the speech act performance suitable for the social and cultural

context given. In this system, possible outcomes from contextual configuration are referred to as *generic structure potential* (GSP). Not only the environment where the text occurs but also the cultural values and norms seem to affect construction of contextual configurations.

Swale (1990) explores the concept of genre and takes the definition of genre as 'the "type" of communicative event' (ibid: 39, original author's emphasis) such as jokes, stories and lectures. Couture (1986: 80) separates *genres* from *registers*, and defines registers as 'collections of certain lexical choices and conventional syntactic arrangement' associated with discourse situations, and genres as 'conventional instances of organized text'. The three components of field, tenor and mode are closely related to the level of registers (Martin 1997). Further, Bhatia (1993: 13-14) defines and summarises genres as 'highly structured and conventionalised communication' which are 'identified and mutually understood' by participants in a discourse community.

Studies were conducted to reveal choices of registers in genres and its outcomes. Coupland (1983) studied fifty-one conversations in a travel agency in Cardiff, South Wales, between an assistant and local clients of various social backgrounds, and reported the relationship between the use of explicit/implicit expressions and social classes. It is interesting to note that the study showed that not only the social and physical setting but also social and cultural backgrounds of the participants have an influence on structure in discourse. The notion that the social and cultural backgrounds of participants impact upon their use of language was conceptualised in Halliday and Hasan (1985) with the term of contextual configuration. The two-way relationships between participants' language use and their

social and cultural backgrounds are assumed to affect the construction of a context. This point is examined in the current study.

Walter (1988) examines the language of a jury summation in a courtroom in order to investigate what is a successful summation. One of the findings of the research is that jurors' evaluations on summation are not directly reflected in the decision made in the trial. This finding emphasises the fact that a speech event can be more closely related to a speech genre than a speech performance in the case of a jury summation, which further emphasises the importance and idiosyncracies of context.

In the current study, genres are treated as structured conventions and registers as choices of participants' linguistic and paralinguistic actions in conversation. Context is viewed both as a pre-established frame which affects participants' actions and as renewable entity through interactions. I will consider to what extent and in what way participants in intercultural settings follow the conversational styles of the target language and their own native language, and how their adjustment to the target discourse community and the preservation of the conversation styles in the native language affect participants' construction of the context.

2.2.2 Social interactional context

Michael McCarthy (1998: 31-32) introduces the notion of 'goal-orientation in interaction' to conceptualise the relationship between text and social action by extracting characteristics of genre in everyday conversation. Genre is here defined as a 'social compact', which is social behaviour that 'the participant enters upon in the unfolding discourse process'. There are four dimensions of social compact: (1) *expectations* of use of 'generic resources' appropriate for the setting; (2) *recollections*, which refers to participants past experiences; (3) *formulations*, which is an action to

comment on and summarise ‘current, ongoing activity’; and (4) *instantiation* to initiate ‘a new set of goals’.

The current study attempts to analyse British tutor-British student conversation and British tutor-Japanese student conversation. In the case of the Japanese students, the first two elements of social compact are considered mainly in relation to the participants’ assumptions for academic tutorials and their past experiences in the target culture and their own culture. The last two are related to their recognition of the setting through the process and their representations to achieve transactional and interactional goals in the context.

McCarthy (2000) also investigates hairdressers’ talk and divided talk into four types:

> 1. Phatic exchanges (greetings, partings)
> 2. Transactional talk (requests, enquiries, instructions)
> 3. Transactional-plus-relational talk (non-obligatory task evaluations and other comments)
> 4. Relational talk (small talk, anecdotes, wider topics of mutual interest)
> (McCarthy 2000: 104)

Almut Koester (2006) focuses on the spoken workplace genre in relation to *interaction* and *transaction* of communication. The spoken data she used was collected from eight offices in the UK and the US between 1996 and 1997 and built up as the ABOT corpus, a small-scale corpus of American and British Office Talk.

Based on McCarthy (2000), five types of talks are defined in the research: (1) non-transactional conversation, (2) phatic communication, (3) relational episodes, (4) relational sequences and turns, and (5) interpersonal markers (Koester 2006: 56). Koester (2006) analysed the use of discourse features such as modality and hedges in spoken workplace genres with transactional goals and relational goals, and the relationship between these transactional and relational episodes. He concluded that relational talk functions to build 'a solidarity and common ground' at work (Koester 2006:161).

Furthermore, McCarthy (1998:10) introduced five *context types* : transactional, professional, pedagogical, socialising and intimate, and three *goal types* : provision of information, collaborative tasks and collaborative ideas:

Table 2.2.2-1 Social interactional context

Goal types / ***Context type***	*Information provision*	*Collaborative task*	*Collaborative idea*
Pedagogic	English lecture	Individual computer lesson	Small group tutorial
Transactional	Commentary by museum guide	Choosing and buying a television	Chatting with hairdresser
Professional	Oral report at group meeting	Colleagues window dressing	Planning meeting at place of work
Socialising	Telling jokes to friends	Friends cooking together	Reminiscing with friends
Intimate	Partner relating the story of a film seen	Couple decorating a room	Siblings discussing their childhood

(Adolphs 2001: 49, Carter 2004: 150)

In the table above, *context types* categorise social context into five types while *goal*

types divide *aims to achieve* through social action into three categories. Even in a particular context, specific actions to achieve both transactional goals and interactional goals are observed in discourse level. These factors are 'multi-functional' and affect all of the aspects of discourse in communication from phonological features to social activities. This multi-functional nature of transaction and interaction can be illustrated with a sample conversation between a customer and a seller below:

	Transaction		*Interaction*
	act	turn	interpersonal
Customer: Can I have three steak pies, please.	request	initiate	politeness
Seller : Oh no problem, love.	accept	respond	friendliness

(From British National Corpus, Davies 2004)

Figure 2.2-1 Multi-functional nature of discourse

This instance is categorised as a 'transactional-collaborative task' in context. At the social activity level, for example, buying three steak pies can be achieved through this transaction, and can establish a positive relationship between a customer and a seller. Additionally, at the level of discourse, a speech act to request or to initiate a turn to buy three steak pies can be interpreted as a transaction, and showing politeness using 'please' or showing a sort of familiarity by addressing the customer as 'love' can function in interaction to build a relationship between participants.

In the current study, the social interactional context mapping and the multi-functional nature of discourse will be taken into consideration. The conversation data

analysed in falls within the pedagogic – collaborative idea context, and the academic tutorials between a tutor and a student at university which constitute the data are analysed with reference to transactional and interactional aspects in discourse.

2.2.3 Contextualisation

Gumperz (1992: 230) introduces the term *contextualisation*, which refers to 'speakers' and listeners' use of verbal and nonverbal signs to relate what is said at any time and in any one place to knowledge acquired through past experience, in order to retrieve the presuppositions they must rely on to maintain conversational involvement and assess what is intended'. There are four cues for contextualisation in speech production according to Gumperz (1992): prosody, paralinguistic signs, code choice and choice of lexical forms:

1. Prosody, which I take to include intonation, stress or accenting and pitch register shift.
2. Paralinguistic signs of tempo, pausing and hesitation, conversational synchrony, including latching or overlapping of speaking turns, and other "tones of voice" expressive cues.
3. Code choice from among the options within a linguistic repertoire (Gumperz 1972), as for example in code or style switching or selection among phonetic, phonological or mophosyntactic options.
4. Choice of lexical forms or formulaic expressions, as for example opening or closing routines or metaphoric expressions […]

(Adapted from Gumperz 1992: 231)

In the current study, paralinguistic signs listed in the second, such as turn exchanges and pausing, and choice of expressions listed in the fourth, such as the use of response tokens, are focused on. Participants' verbal and non-verbal behaviours are assumed to contextualise the situation, and simultaneously the context where conversation occurs affects participants' choice of contextualisation cues. This two-way relationship will be considered in the current research.

2.3 Listenership and response tokens

2.3.1 Response tokens

As discussed in the earlier section, Yngve (1970: 567) broadens the concept of linguistics as 'the scientific study of how people use language to communicate'. The term *backchannels* was introduced as signals which 'the person who has the turn receives short messages such as *yeah* and *uh-huh* without relinquishing the turn' (Yngve 1970: 568), and how turn-taking occurs was analysed and described as follows:

> The turn-change signals, or the people using them, are not infallible, it seems. There are cases where mistakes occur and each subject appears to assume he has the turn, resulting in their both speaking at once.
> (Yngve 1970: 574)

Response tokens are defined as both verbal and non verbal responses which listeners make in conversations. There are several terms to refer to this function such as *response tokens* in Gardner (2002) and O'Keeffe et al. (2007), *listener response* in Maynard (1990) and *minimal response* in Fellegy (1995). I adopt the term response

tokens in this current research, and the objectives of the current study include both verbal and visual response tokens.

Duncan (1974) conducted research on spoken discourse in a clinical setting and divided response tokens into two types, vocal and visual signals. He regarded both as objects of research in spoken discourse and attempted to match these items with three signals used in spoken discourse based on the data from his observation: (1) *auditor back-channel signal* including verbalised signals such as *m-hm*, sentence completions, request for clarification, brief restatement, and head nods and shakes, (2) *speaker within-turn signal* referring to completion of grammatical clause and turning of head towards auditor, and (3) *speaker continuation signal* such as turn head away from auditor (ibid: 166-167). The current study adopts Duncan's broad definition of response. Both visual and verbal response tokens will be treated as objectives of the current research.

McCarthy (2002) reviewed Duncan's study and pointed out the importance of listenership in relation to the broad range of behaviours which are observed in conversation: 'Duncan's range of items is indicative both of the potential range of behaviour that may be considered relevant to the study of listenership and, once again, of the difficulty in establishing the boundary between backchannelling, turn-taking and floor-grabbing' (ibid: 52). Through this, the close relation between the use of response tokens and turn management strategies is indicated.

McCarthy (2002: 69) investigated the relationship between good listenership and the use of response tokens, and formed the notion of the 'multi-functioning' of response tokens as follows:

> […] they [listener response tokens] not only mark acknowledgement and confirm understanding, but may also express agreement, and in this way, social action is co-ordinated and fine-tuned on several levels simultaneously.
>
> (McCarthy 2002: 53)

As described in the previous chapter, O'Keeffe (2007) defines listenership as 'the active, responsive role that listeners have in conversation', and considers good listenership as something 'both natural and desirable for efficient spoken communication' (ibid: 142). How participants realise good listenership in conversation and whether any cultural differences can be observed in realisation and perception of good listenership will be considered in the current study.

2.3.2 Response tokens in spoken English

As described in Chapter One, a number of linguists have made attempts to investigate the use of response tokens in conversation from a range of aspects; from discourse analysis to intercultural communication (Adolphs 2008, Duncan 1974, Heritage 1997, LoCastro 1987, Maynard 1990, McCarthy 1998, O'Keeffe & Farr 2003, Sacks 1992, Yngve 1970)

In an early project, Edmondson (1981) reported four functions of response tokens: (1) *Go-Ons* to show the hearer's intention to let the speaker continue, (2) *Accepts* to show the hearer's understanding, (3) *Exclaims* to show the hearer's emotional reaction such as doubt, surprise, interest and sympathy, and (4) *Okays* to show the hearer's satisfaction of the information given (ibid: 125). Maynard (1989) redefined six main functions of response tokens based on Edmondson (1981):

1. Continuer – indicates simply that the listener is bypassing the chance to initiate a repair (similar to Edmondson's Go-Ons)
2. Display of Content Understanding – is used when there may be doubt on the part of the speaker as to the listener's understanding (similar to Edmondson's Accept)
3. Support Towards the Speaker's judgment – occurs as a response to a speaker's evaluative statement
4. Agreement – serves as a response to a question or question-like statement
5. Strong emotional Response – consists of a laugh or exclamation (similar to Edmondson's Exclaims)
6. Minor addition, Correction or Request for Information – includes listener comments that changes "the quality of the currently activated information"

(Maynard 1989: 171-172)

Rod Gardner (2002) also describes functions of common response tokens: *mm hm* and *uh huh* as continuers, *mm* as acknowledgement and terminator, *yeah* as acknowledgement, *oh* and *right* as newsmarker, and *okay* as closing. Although these definitions of response tokens are valuable to this current research, the multi-functional nature of response tokens has not been researched up until this point.

In recent study, the form and functions of response tokens have been

investigated in relation to listenership. The multi-functional nature of response tokens has been assumed as one of the important characteristics of response tokens (Knight et al. 2006, Maynard 1990, McCarthy 2002, O'Keeffe et al. 2007), and therefore the idea of four basic functions of response tokens defined by O'Keeffe and Adolphs (Knight et al. 2006, O'Keeffe & Adolphs 2008) has been incorporated into this investigation:

> (1) Continuers [CN]: Maintaining the flow of discourse.
> (2) Convergence tokens [CV]: Markers of agreement/convergence.
> (3) Engaged response tokens [EN]: Markers of high engagement where addressee(s) respond on an affective level to the content of the message.
> (4) Information receipt tokens [IR]: Markers of points in the discourse where adequate information has been received.
> (O'Keeffe & Adolphs 2008 : 84)

In terms of forms of response tokens, O'Keeffe et al. (2007) divide verbal response tokens into three forms; minimal response tokens, non-minimal response tokens and clusters of combinations of these response tokens:

> Usually minimal response tokens are defined as short utterances (for example *yeah*) or non-word vocalisations (such as *mm*, *umhum*) while non-minimal response tokens are mostly adverbs or adjectives (for example *good*, *really great*, *absolutely*') or short phrases/minimal clauses (such as *you're not serious, Is that so?, by all means, fair enough, that's true, not at all*).

(O'Keeffe et al. 2007)

Participants' use of response tokens are analysed in this project based on these three forms of response tokens, namely minimal response tokens, non-minimal response tokens and clusters. This analysis was achieved by counting numbers of these three forms of response tokens in participants' utterances in the global pattern analysis (see Chapter Three and Chapter Five). Both forms and placements of response tokens will be examined with a multimodal and time-based corpus linguistic approach in order to investigate similarities and differences in listenership behaviour between British-British conversation and British-Japanese conversation.

2.3.3 Japanese learners' response tokens in spoken English

Given the fact that this current study attempts to compare forms and placement of response tokens in the British participants with the Japanese participants in English conversations from perspectives of intercultural communication, previous research on the use of response tokens in Japanese learners of English will be reviewed in this section.

White (1989) studied English conversations with American and Japanese participants and reported that Japanese participants have a tendency to use verbal response tokens more often than American participants. Maynard (1997a) also conducted a piece of research on comparing Japanese students' use of response tokens in English conversation with American students. In her research, verbal response tokens and non-verbal response tokens such as head nod (vertical head movement) and laughter were investigated, although head shake (horizontal head movement) and gaze were excluded. She reported that Japanese students used response tokens more

often than American students.

> The total occurrence of back channel observed in 3 minute segments from each of the 20 Japanese and 20 American pairs was 871 for Japanese and 428 for American listeners. This means that while Japanese listeners send some sort of back channel once every 4.13 seconds, Americans do so once every 8.14 seconds.
> (Maynard 1997a: 51)

The reasons for the frequent use of response tokens by Japanese listeners were explained from a cultural point of view. By using frequent response tokens, Japanese listeners intend to show their interest to the speakers as they do in Japanese language conversation.

The relation between participants' cultural backgrounds and their use of response tokens is important since it is assumed that norms and cultural values of Japanese learners can be reflected in their use of language. There are some neglected areas within the research, however: response tokens were analysed only quantitatively, and a qualitative analysis on the use of the response tokens is needed to clarify the Japanese listeners' intention to use response tokens. Response tokens might be used as continuers or initiations of turn or function as both depending on the context. Another drawback can be the rather simple interpretation that Japanese listeners do use response tokens frequently because they use *aizuchi* (see Section 2.3.4) frequently in Japanese language conversation. This point also needs to be clarified with a further research with qualitative analysis.

2.3.4 Response tokens in spoken Japanese

Japanese language has response tokens referred to as *aizuchi*. Some contrastive studies were conducted on the use of response tokens by Japanese listeners in comparison with native speakers of English, and the use of aizuchi in comparison with English response tokens. There are studies which report that nods, gaze and silence can be used more often than verbal response tokens in Japanese conversation (Hayashi et al. 2002, Maynard 1990, Mori 2002).

Maynard (1990) observed video-recorded pair conversations in Japanese language and in American English, and conducted contrastive conversation analysis placing focus on differences in listeners' backchannels in casual conversations between Japanese and American English. In this research, both verbalised responses and non-verbal response tokens such as head movement and laughter were taken into consideration. There are many findings reported, including sharing the use of 'completion' in both languages, and the more frequent use of response tokens in Japanese conversations:

> The types of back-channel responses sent by Japanese and American listeners were similar, both Japanese and American speakers used brief utterances and head movements. The major difference was in the frequency and the discourse contexts in which back channels occurred. Among Japanese listeners there were more repetitious back-channel responses (as punctuated vertical head movements repeated up to four times consecutively) and these back-channel strategies frequently

> overlapped with the American speaker's utterance. Back channels sent by Americans were almost exclusively limited to occurrence during intra-turn pauses.
>
> (Maynard 1990: 410)

There are some differences observed in the research in the participants' strategies and choices in the use of response tokens between the Japanese participants and the American participants. The Japanese participants' frequent use of response tokens in intra-turn pauses were reported in comparison with the American participants. In terms of repetitions in Japanese conversation, frequent use of exact repetitions in talk among friends in Japanese is reported by Fujimura-Wilson (2007).

As for non-verbal response tokens, head movement creates harmonious rhythm in Japanese conversation. However, that seems not to be the case in intercultural conversations in English between the American and the Japanese students as Maynard (1997b) describes below:

> I found no case of 'rhythmic ensemble' in our intercultural data, although Japanese speakers continue to use head movement with American partners. Here we witness that the sense of 'disengagement' or lack of 'togetherness' is not found in language per se, but in strategies of interactional management.
>
> (Maynard 1997a: 51)

From these findings, the conclusion that different languages might have similar items

to organise turns can be drawn, but these functions might be different from language to language.

Ward (1998) studied dyad conversations in Japanese and American English language with a corpus, and reported that low pitch was used as a cue of backchannels in Japanese language, which was not so important in English.

Table 2.3.4-1 Various rules for predicting backchannel feedback (Japanese)

Predictions from	Coverage	Accuracy	Figure of Merit
low pitch regions	56% (496/873)	34% (496/1447)	.195
random	25% (222/873)	24% (222/915)	.062
utterance end	68% (593/873)	22% (593/2751)	.146
utterance end and low pitch region	36% (314/873)	32% (277/978)	.115
utterance end and no low pitch region	32% (279/873)	16% (279/1773)	.050
eavesdropping human judge (estimate)	95%	67%	.64

(Ward 1998: 63)

Table 2.3.4-2 Various rules for predicting backchannel feedback (English)

Predictions from	Coverage	Accuracy	Figure of Merit
low pitch regions	48% (172/359)	18% (172/936)	.088
random	22% (80/359)	13% (80/618)	.029
utterance end	46% (164/359)	10% (164/1698)	.044
utterance end and low pitch region	30% (109/359)	19% (109/578)	.057
utterance end and no low pitch region	15% (55/359)	5% (55/1120)	.008

(Ward 1998: 63)

The tables above, which are taken from Ward (1998), show various rules before response tokens occur in Japanese and English language. The coverage of low pitch regions in Japanese language is 56 % with 34 % accuracy while the converage of low pitch regions in American English is 8% lower than the Japanese figure with lower

accuracy. Ward (1998: 63) listed six factors which are assumed to function as 'sound symbolism' or 'synaesthesia' in Japanese conversation and American English conversation: (1) agreement for nasalization, (2) contemplation form, (3) deference for breathiness, (4) the willingness to listen for number of syllables, (5) coldness for sharpness of final energy drop, and (6) energy and pitch height and slope.

Mizutani (1983) examines a relationship between speed of speech and frequency of response tokens in Japanese conversation, and found that the faster the speaker talked, the more frequently the listener gave response tokens. Although prosodic features in relation to the use of response tokens in conversation will not be included in the scope of this current study, it will be worth quoting studies by Ward (1998) and Mizutani (1983) as some of the limited number of studies which attempt to investigate prosodic cues for backchannels in Japanese and American English conversations.

In terms of functions of response tokens, LoCastro (1987: 104) describes a function of response tokens in Japanese as 'passing the opportunity to take over the floor'. The function seems to be similar to continuers in English language; however, it might be slightly different if we think about cultural values such as *other-centred-ness* in Japanese culture:

> A good conversation partner tends to empathize with others, being aware enough jointly create a conversation; this contrasts with American discourse patterns where conversations seem to be displays of 'antagonistic behaviour'.
>
> (LoCastro 1987: 105)

Some Japanese linguists have also investigated forms and functions of *aizuchi* (Hayashi et al. 2002, Kogure 2005, Mizutani 1983, Mori 2006, Ohama 2006). Hayashi et al. (2002) video-recorded a Japanese language conversation with four female Japanese participants and conducted a detailed analysis on the participants' use of backchannels, gaze and gesture. I quote a part of the transcription of the Japanese conversation where the participants were talking about similarities in designs of ladies' clothes between the time the conversation was recorded and the past:

14 Mari : [*demo are*] *mukashi wa, aya*-(0.2) *a*[*nna no* =]
But, in the past, (0.2) those kinds (of clothes)…

15 Yoko : [◦a!◦
◦oh! ◦

16 Mari : =*hayatta desho?*=
...were popular, weren't they?

17 Yoko : = *hayatta*
Were popular

18 (.)

19 Yoko : *wata*[*shie ne::, atta no ne::*, =
I found (them)

20 Mari : [*onnaji yo ne::::.*
(They) are the same, aren't they?

* Transcription symbols

(.)	A short, untimed pause
(0.3)	A timed pause
[	The onset of overlap
she	Underscore indicates prominent stress
◦she◦	Degree signs indicate lower volume than surrounding talk
she:	Colon indicates sound stretch

(Hayashi et al. 2002: 88)

Hayashi et al. (2002: 89-93) noted Yoko's initiation of turns. After Emi's turn in line 14, which allows Yoko to lean 'toward in the direction of Mari […] with *a!* (oh!) produced in a soft voice'. A response token *a!* can be interpreted as a newsmaker *oh* in English. Then in line 17, Yoko immediately inserted 'a repetition of the verb that Mari has just used in her preceding turn'. In a short pause in line 18, Yoko and Mari achieved a 'mutual gaze' which allows Yoko to initiate the next turn. These complicated combinations of response tokens, gaze, gesture and silence organise a sequence of Japanese conversation.

Ohama (2006) divides two functions of aizuchi in Japanese: (1) emotional expressions and (2) concept expressions. Forms of emotional expressions are classified into nine types: *a* type (*a, aa:, aa: aa:* etc), *un* type (*un, unun, nn, u:n, nn:* etc), *e* type (*e, ee:, ee* etc), *o* type (*o, oo,* etc), *hai* type (*hai,haihai, haa: ha: ha:* etc), *fu:n* type (*fu:n, fun:, funfun* etc), *hee* type (*hee* etc), *hoo* type (*hoo, ho: ho:* etc), and *maa* type (*maa, maa-ne* etc). Concept expressions are categorised into seven types: *sugoi* type (*sugoi, sugoi-ne* etc, equivalent to 'excellent', 'great'), *honto* type (*honto, honma, honto-desu-ka* etc, equivalent to 'really?'), *iya* type (*ieie, iya, ya:* etc, equivalent to 'No no'), *ii* type (*ii-desu-ne* etc, equivalent to 'good'), *repetition* type (repetition, formulation, sentence completion etc), and *sou* + postposition type (*sou, sou-da-yo, sou-yo-ne, desho* etc, equivalent to 'yes it is', 'is it?') (ibid: 167-168, translated by me). There seems to be more variations in Japanese response tokens than English and these items are selected depending on the context and the other speaker within the conversation.

Kogure (2005) compared frequencies of response tokens in Japanese between genders, and found that response tokens were used more in female-female

conversation than in male-male conversation, and reported adjustment of frequency in the use of response tokens in a cross gender conversation. Although gender aspects are out of focus in the current research, multi-functional nature of response tokens and amount of response tokens used in conversation in British-British conversation and British-Japanese conversation will be investigated.

2.3.4.1 Response tokens in American learners of Japanese

Mori conducted a study of the use of aizuchi in American learners of Japanese. I quote a part of a transcription in Mori (2002) below, which is a Japanese language conversation held by a native speaker of Japanese, Sasaki, and two non-native speakers of Japanese, Miles and Oakland. Mori evaluates the task involved discussions in Japanese on the subject of learning Japanese as a second language, and it appears that the participants showed spontaneous response in talk interaction, which can be interpreted as a positive nature of having a target language discussion.

		Japanese original text		English translation
12	Sasaki	oto- otosan ga naiteru no o::	Sasaki	Have- have you seen your
13		mita koto ga arimasu ka?		father crying?
14	Miles	nai n desu, itsumo a- nete mashita.	Miles	No, I haven't. I was always sleeping.
15	Sasaki	uha [haha	Sasaki	uha [ha ha
16	Oakland	[HA HA HA ha ha	Oakland	[HA HA HA ha ha
17	Miles	watashi no haha no- ano:::	Miles	Cause that was my mother's- we::ll, it
18		shi[goto deshita kara		was her [job.
19	Oakland	[uh hh	Oakland	[uh hh
20	Sasaki	u::n [ha ha	Sasaki	u::n [ha ha
21	Miles	[u::n	Miles	[u::n
22	Oakland	ya boku mo nai desu.	Oakland	No I haven't either.
23	Sasaki	hu::n ∘huu:n hu::n hu:n hu:n∘	Sasaki	hu::n ∘huu:n hu::n hu:n hu:n∘

∗ ∘ ∘quieter than the surrounding talk
(Mori 2002: 333)

The issue which I would like to highlight from the transcription above is an adjustment of response tokens in the target language by non-native speakers of Japanese. In line 20, Sasaki, the native speaker of Japanese, inserts *u::n*, which is a

common response token in Japanese conversation (Ohama 2006, Ward 1998). Immediately, Miles assimilates the sound and gives back the same response token in line 21. Mori (2002) also reports Miles's use of *aa:* in the conversation, which is an equivalent of *oh* and functions as a change-of-status token in Japanese. Although we can not identify non-verbal response tokens from the transcription, it can be said that the participants show their attempts to change their way of use of response tokens in order to follow the norms in the target language in Mori's study.

2.3.5 A classification of response tokens

Based on Maynard (1990), Gardner (2002) and O'Keeffe et al. (2007), a classification of response tokens has been made as shown below. There are two broad categories namely vocal and visual response tokens, both of which have several sub components.

Table 2.3.5-1 A classification of response tokens

		Examples & descriptions
Vocal response tokens	minimal response	*Mm, Uh-uh, Yeah, Right*
	non-minimal response	*lovely, definitely, I see*
	clustering of response tokens	*Mm mm, yeah right*
	laughter	chuckles and laughs
	pause	silent pause
Visual response tokens	head nods	Any vertical head movement
	hand gestures	Any hand movements
	head turning	Head is moving towards speaker
	head shake	Any horizontal head movement
	Self comfort	Crossed arms
	foot movement	Crossed legs

There are six sub components under vocal response tokens, *minimal response tokens*, *non-minimal response tokens*, *clustering of response tokens*, *laughter* and *pause*. Six gestures are included in visual response tokens, namely *head nods, hand gestures*, *head shakes, head turning*, *self comfort* and *foot movements*. However, grammatical and syntactic items of response tokens, such as completion, clarification, restatement, overlap, and other body movements, such as facial expressions and gaze, have been eliminated since these items are out of the scope of the current study. The response tokens listed in the table above will be examined in the pilot study of the current investigation, and the focus will be narrowed down in the main body of the discussion.

2.3.6 Transition from listenership to speakership

Duncan and Niederehe (1974) conducted research on signals and cues in conversation in relation to turn-taking and response tokens. They used the term, *speaker-state*, to express a speaker's role in conversation. There are two types of speaker-state signals, namely *turn-yielding signal*, which is used to initiate an turn exchange, and *gesticulation signal*, which indicates the speaker's retaining of his role and 'inhibiting the auditor's attempts to take the turn' (Duncan & Niederehe 1974: 235). Although gesticulation signal was rarely seen in their interview data, two types of cues for the signal are defined: (1) 'hand movements' away from body, and (2) 'a tensed hand position' such as a fist. The four types of cues for turn-yielding signal are described as follows:

(1) Shift away in head direction;

(2) Audible inhalation (a sharp, audible breath);

(3) Initiation of a gesticulation;

(4) Paralinguistic overloudness.

(Duncan & Niederehe 1974: 240)

Duncan and Niederehe (1974: 235) also point out significant relationship between those speaker-state signals and possibility of auditor's attempts to take the turn.

Drummond and Hopper (1993) investigated three types of response tokens, *Mm hm, Uh huh* and *Yeah* in relation to initiation of speakership by analysing four sets of telephone conversation data in English collected from a conversation library at South Western University. These response tokens are referred to as *acknowledgement tokens* (AT). They found that AT are often followed by further speech by the speaker who produced the AT, and examined the length of the speech that followed. In terms of the initiation of speech by each AT, two of the cases, where speech follows the AT or does not follow, were counted for each AT and summarised in the table below.

Table 2.3.6-1 Speakership Incipiency of *Mm hm, Uh huh* and *Yeah*

Encounter									Totals	
Speakership	F1		D8		A21		CIS		(Weighted %)	
Incipiency	nSB	SB	nSB	SB	nSB	SB	nSB	SB	nSB	SB
Tokens										
Mm hm	**1**	**0**	**13**	**2**	**4**	**0**	**36**	**1**	**54 (5%)**	**3**
Uh huh	**0**	**0**	**4**	**0**	**9**	**1**	**11**	**0**	**24 (4%)**	**1**
Yeah	**5**	**7**	**17**	**14**	**41**	**34**	**18**	**15**	**81 (46%)**	**70**

Note. Encounter key: F1= Mother/Daughter; D8= Kips & Cara, college flirts; A21= Pam & Gloria, pals; and CIS= Cancer Information Service. Speakership incipiency key: nSB= No speakership bid follows AT [acknowledgement token] (freestanding); and SB= Speaking bid follows AT.
(Drummond & Hopper 1993: 168)

From the results shown in the table above, differences between *yeah* and the other two items can be seen. It was found that '*uh huh* or *mm hm* rarely initiated further speakership; *yeah* initiated further speakership on almost half of its uses' (ibid). As for turn length, *yeah* initiated a *minimal turn*, such as 'yeah, it is', more often than the other two tokens. On the other hand, *mm hm* and *uh huh* were uttered as *freestanding*, which means these items are uttered without any further speech, in most cases, as shown in the table below.

Table 2.3.6-2 Acknowledgement Tokens and Turn Length

	Freestanding	Minimal Turn	Full Turn	Row Total
Mm hm	43	2	4	49
Uh huh	35	1	2	38
Yeah	72	22	36	130

(Drummond & Hopper 1993: 171)

If we look at the findings from the listeners' point of view, a hypothesis can be made that freestanding occurence of *mm, uh huh* and *yeah* have a function related to listenership, while *yeah* with a full turn has a function of initiating turn exchanges. These aspects of the use of response tokens will be investigated in depth through the current study. Not only the relationship between the forms of response tokens but also the placements of response tokens in listener status before turn exchanges will be examined in relation to leadtime.

2.4 Intercultural communication

This section will review theories in intercultural communication, for this research attempts to investigate the use of response tokens in British-British conversation and

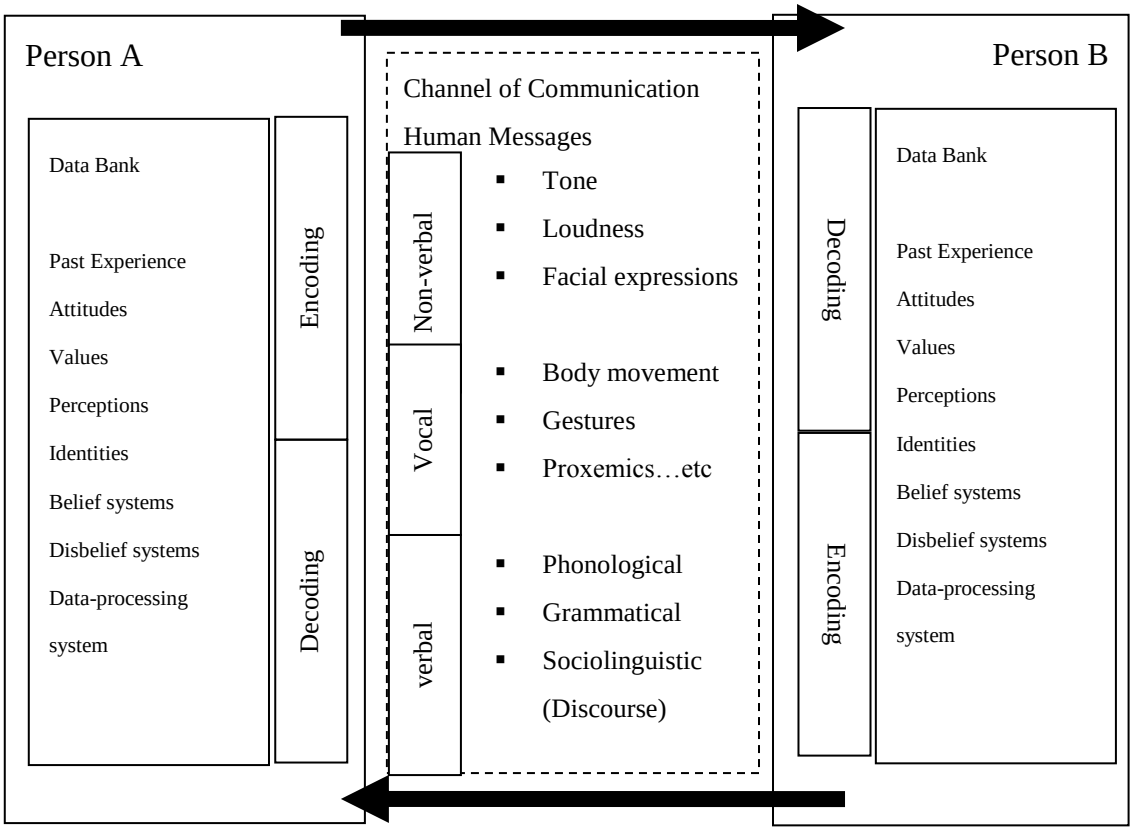
British-Japanese conversation from aspects of intercultural communication. This section starts with discussions on how language and culture are related, and moves on to theories about interlanguage pragmatics where concepts of *face* and *politeness* will be reviewed. This section ends by making links between pedagogic issues and the acquisition of interlanguage pragmatics.

2.4.1 Language and culture

A number of anthropologists have attempted to define the relationship between culture and language. In Whorf's view, on the one hand, language determines the way people think (Carroll 1956). In Ager (1994a, 1994b), on the other hand, the term, *languaculture* is introduced, which expresses the notion that language and culture will interact with each other and that there is no border between them. At the same time, the idea of *frames* (Ager 1994a, 1994b, Brown & Yule 1983), in other words *a structure of expectation* (Kramsch 1998) or *schemata* (Cook 1994), has been conceptualised to describe people's assumptions towards particular situations or other people's utterances. How language and culture interact with each other, and in what way cultural values and norms are reflected by the use of language, will be considered here.

2.4.1.1 The model of speech

Jandt (1998:26) illustrates the main components of communication such as *encoding* and *decoding* message. Singer (1998:228) added the elements in intercultural communication, such as past experience, values and identities, and described a model of communication.



(Adapted from Singer 1998:228, McKey, Davis and Fanning 1995, Clyne 2003)

Figure 2.4-1 A model of intercultural communication process

In addition to Singer's model, I have inserted three elements based on McKey, Davis and Fanning (1995), namely non-verbal, vocal and verbal layers of communication, into the channel of communication. The subsidiary components of these three elements, such as facial expressions and body movements in the non-verbal layer, tones and loudness in the vocal layer, and phonological, grammatical and sociolinguistic in the verbal layer, have also been added to Singer's model (Clyne 2003, McKey et al. 1995). The purpose for describing this model is to visualise a concept of discourse in communication. Furthermore, the use of visual response tokens, such as head nods and hand gestures, are categorised into 'non-verbal' in the

channel of communication. Conversely, the use of verbal response tokens can be placed both in vocal and verbal since a non-minimal response token *mm* can be categorised as vocal but the other response token *yeah* can be verbal. In addition, turn-taking management and the use of response tokens might be related to the sociolinguistic layer and discourse.

Clyne (2003) undertook several contrastive studies into the three levels of verbal communication: phonology, grammar and discourse, and showed that mistakes in decoding and encoding will happen in all of these levels, which can be obstacles to mutual understanding. If a sender encodes concept X in one way and the receiver decodes (or interprets) X in another way, the message being sent will be different from what the sender has intended, or will not be comprehensible. What I would like to consider in the current study is the cases where a receiver interprets messages in a different way from the sender's intention because of differences in cultural values represented in language use in an interlanguage setting. Ager (1994b) referred to the places where misunderstanding in intercultural communication happens as 'rich points', and suggests that discourse will be the place where cultural values tend to be reflected.

2.4.1.2 Discourse and cultural value

Tannen (1986: 152) reported a successful story where one of her students, a Chinese male student, overcame a problem between himself and his American female friend by changing his own conversation style. In this anecdote, the Chinese student first thought of his American friend as 'an intolerable person: a compulsive talker' since she kept talking without giving him any room to turn-take. After learning intercultural communication in class, however, he changed his conversational styles:

> Whenever she cut me off, I immediately cut her off in return […] I tried by all means to dominate the conversation. She has a tendency of ignoring the third person present when she talks to someone. So, I cut her off many times to drag George [a friend of both] into the conversation, to show that I controlled the conversation.
>
> (Tannen, 1986:151)

This story is an example of how people expect the use of turn-taking to be different in different cultures, and values behind them can be seen from the perspective of *collectivism* versus *individualism*.

In China, Japan and other Asian countries, on the one hand, people tend to think that harmony of the group is more important than individuals, and this view is referred to as collectivism. One of the main reasons why the Chinese boy in the former story felt uncomfortable with the American student's conversational style at first can be that he felt that she talked without thinking harmony of the group and did not invite him or others to join the conversation, which meant that she did not make any effort to develop harmonious two-way communication between them. In the US, on the other hand, being independent, expressing one's opinions explicitly and controlling the conversation can be regarded as more acceptable than reserving and maintaining harmony in groups.

McKey, Davis and Fanning (1995) describe rules for effective expression in American conversation style; such as '[m]essages should be direct' and 'immediate'. Maynard (1993) compares cultural differences between Japan and the US as follows:

> Differences between Japan and the United States become apparent in how one understands and comes to terms with the relationship between self and society. The starting point for the Japanese experience lies in society, while the starting point for Americans lies in the concept of self. By 'starting point' I mean the primary and deep-rooted self-concept one is encouraged to identify with early in life.
> (Maynard 1993: 6)

Although this instance might not be applied directly to the current study where a comparison has been made in conversation styles between British English and Japanese, this is useful as an exemplar of the dichotomy between Western and Eastern cultures.

The concept of *wakimae* is also introduced as one of the underlying and shared concepts in Japanese culture, which is translated as 'discernment' and refers to:

> […] sets of social norms of appropriate behaviour people must observe to be considered polite in society. The manipulation of politeness strategies is a concrete method for meeting the social rules of wakimae. Both American and Japanese speakers must behave according to the wakimae code. […] Although both Japanese and American speakers wish to make their interactions comfortable by meeting wakimae standards, how they arrive at their comfort zones differs.
> (Maynard 1997b: 57)

In addition to the concepts of *other-centred-ness* and wakimae, the concept of hierarchy can be recognised in Japanese conversation, as shown by the way that Japanese students initiate a discussion in Watanabe's study. Watanabe (1993) observed the differences in starting discussions between Japanese and American students. She observed that Japanese students negotiated and decided the order of the speakers before the discussion began, while American students spoke out spontaneously without any particular order. Conversely, Japanese participants are conscious of who has the right to speak first and of being reserved. Scollon and Scollon (1995: 81) state about Eastern cultures that 'it is certainly accurate to say that hierarchy in relationships is much more consciously observed than it is in the west' and linked Confucianism with the way people think.

The current study concerns these differences in cultural values reflected in conversational styles. In the case of the British-Japanese conversation in the current research, it is assumed that the Japanese participants represent their own cultural values in conversation in the interlanguage setting.

2.4.1.3 Confucianism in intercultural communication

Yamada (1997) exemplified a conversation between a Japanese person and an American in a business situation as a case where the participants follow 'different rules'.

> Mark: How are we going to get the tapes back here?
>
> Masa: I will go pick them up.
>
> Mark: Oh, I thought Amanda was going to go because you couldn't go until tomorrow.

Masa: Yes.

Mark: So, Amanda's going to pick them up.

Masa: Amanda is very busy. I will pick them up.

Mark: But you can't go until tomorrow, right?

Masa: Yes. [silence]

Mark: [Laughing] But they have to be picked up today - in fact they probably should have picked up yesterday. [silence]

Mark: So Amanda better go today.

Masa: OK.

(Yamada 1997:23)

Yamada interpreted the conversation as follows: Mark places value on 'equal opportunity' and insists that Amanda, one of their colleagues, has to go and pick the data up since she is at work and has the responsibility to do it. Masa, however, thinks about this issue in the way of *other-centred-ness*, in other words, thinking about things from others' point of view, and shows sympathy to Amanda since Masa knows that Amanda has been working for long time and must be tired. *Other-centred-ness* can also be based on the value of interdependence, where people in a family, a community or a work place, should depend on each other and think about what other people think and what can be done for other people. This thought can be rooted in Confucianism, which teaches that '[i]f one wants to establish oneself, one has to establish others' (Yamada 1997: 12). Japanese people regard a person who can think about things from others' points of view as thoughtful and virtuous.

If we focus on Masa's way of sending his message, it can be noticed that Masa

did not express why he wanted to go to pick up the data explicitly. This kind of implication, in other words *high context* or *uncertainty*, can be observed in Japanese and other Asian cultures. Jandt (1995:229) described China and Japan as high context cultures, which means that people try to send messages implicitly and let the receiver notice the meanings, in contrast to western cultures as low context cultures. Mead (1999: 238) believed that Japan has lower needs to avoid uncertainty, which means that Japanese conversation style is more indirect, while the US has higher needs. Part of the reason why people in Japan and other Asian countries avoid explicit utterances can also be caused by Confucianism since people in Asian countries expect that their conversation partners will know what they mean even though they do not say it explicitly. In other words, if s/he says their message directly, the receiver may interpret that as s/he thinks the receiver may not have ability to perceive his/her intention, which can be an underestimation of the receiver. Underestimating the receiver can be taken as insulting and lead to trouble in their relationship. These concepts of maintaining one's esteem have been discussed in the area of pragmatics under the terms of face and politeness, which will be reviewed in the next section.

2.4.2 Interlanguage pragmatics

2.4.2.1 Pragmatics and interlanguage pragmatics

Pragmatics has many definitions. According to Levinson (1983), pragmatics covers both 'context-dependent aspects of language structure' and 'the inter-relation between language structure and principles of language usage'. Based on the definition of Eggins and Slade (1997), McCarthy, Matthiessen and Slade (2002) place pragmatics as one of the disciplinary approaches to discourse analysis.

Pragmatics is treated as philosophy of language derived from Austin (1962)

and Searl (1969), which has ‘shed light on how people interpret particular utterances’ (McCarthy et al. 2002: 60). In pragmatics as philosophy, sentences or utterances for analysis are often *invented* by linguists while pragmatic studies with corpus analysis examine collections of *naturally occurring* conversations. This differentiates the current research with corpus data from the previous studies in pragmatics as philosophy.

Presenting linguistic action patterns, which are interpreted as appropriate in a particular context by the other participants, can be an issue raised not only in acquisition of pragmatic competence in the first language but also in interlanguage. As described in Chapter One, interlanguage pragmatics is defined as ‘the study of nonnative speakers' use and acquisition of linguistic action patterns in one second language (L2)’ according to Kasper (1993:3).

Johnstone (2002: 31) described *the inter-organism perspective* in language learning as ‘[…] acquiring a language means acquiring a world’ and Bakhtin also (1956) describes the relationship between language and a world as follows:

> […] when one begins to hear voices in languages, jargons, and styles, these cease to be potential means of expression and become actual, realized expression; the voice that has mastered them has entered into them. They are called upon to play their own unique and unrepeatable role in speech (creative) communication.
>
> (Bakhtin 1956: 121)

Vygotsky’s (1962 [1934]) words can be one of the precise expressions of *the intra-*

organism perspective, which is a two-way interaction between language and thought.

> The relationship of thought to word is not a thing but a process, a continual movement back and forth from thought to word and form word to thought.
> (Vygotsky 1962[1934]:125)

These two perspectives, namely the inter-organism perspective and the intra-organism perspective, are both taken into consideration in the current study. The relationship between a world and a language, and a language and a thought, which is reflected in participants' listenership behaviours, will be described not as fixed objects but as ongoing 'processes' in the discussion at a later stage (see Section 5.2.4).

2.4.2.2 Face and politeness

The concepts of politeness, or, *rapport management* (Spencer-Oatey 2000), are involved in the areas of interlanguage pragmatics. Such principles of politeness have been investigated by a number of researchers.

Grice (1975, 1989) introduces 'the cooperative principle' with four key elements, namely quantity, quality, relation and manner. These elements are assumed to be fundamental for better communication. Following Goffman (1955), Brown and Levinson (1987: 60-61) defined *face* as 'something that is emotionally invested, and that can be lost, maintained, or enhanced, and must be constantly attended to in interaction', and 'some acts intrinsically threaten face' is referred to as *face-threatening acts* (FTA). Three politeness strategies are introduced: *positive politeness*, which is 'the expression of solidarity', *negative politeness*, which is 'the expressions

of restraint' and *off-record*, which is 'the avoidance of unequivocal impositions' (Brown & Levinson 1987: 2).

However, there are some limitations in these principles. These principles are described only from the speaker's point of view and the notion of cultural differences in conversation styles is also missing. Leech (1983) categorised six areas of politeness, such as tact, generosity, approbation, modesty, agreement and sympathy. Maximisation of these six categories varies from culture to culture and these differences can be described as follows:

> So far, our knowledge of intercultural differences in this sphere is somewhat anecdotal: there is the observation for example, that some eastern cultures (*e.g.* China and Japan) tend to value the Modesty Maxim much more highly than western countries; that English-speaking cultures (particularly British?) gives prominence to the Maxim of Tact and the Irony Principle; that Mediterranean cultures place a higher value of the Generosity Maxim and a lower value of the Modesty Maxim.
>
> (Leech 1983: 150)

In his analysis, some Eastern cultures including Japan are categorised as a culture which has a tendency to value the Modesty Maxim more than English-speaking cultures. As with the natures of Japanese culture described as high context, hierarchical and other-centred-ness by Yamada (1996) and Scollon and Scollon (1995) in the previous section, the expectation of enhancing the Modesty Maxim can

also be rooted in the virtues from Confucianism, which are shared by people in Japan and other Eastern Asian cultures and affects people's conversation styles in these cultures. In this way, to describe people's natures within a culture is useful, as is analysis of the context where participants with different cultural backgrounds are involved. Therefore, how language users can accommodate in an intercultural and interpragmatic setting will be explored in the current study.

Spencer-Oatey (2000: 3) introduced the term *rapport management* to refer to 'the use of language to promote, maintain or threaten harmonious social relations' with raising awareness of interactional perspectives in intercultural settings. Spencer-Oatey distinguished the term *face* from *right*. The term *face* is defined based on Goffman (1972: 5) as 'the positive social value a person effectively claims for himself [sic] by the line others assume he has taken during a particular context', and 'social rights' are defined as 'fundamental personal/social entitlements that a person effectively claims for him/herself in his/her interactions with others' (2002: 540, emphasised by the author).

Table 2.4.2-1 Rapport management from Spencer-Oatey (2000)

	face management (personal/social value)	sociality rights management (personal/ social entitlements)
personal/independent perspective	quality face (cf. Brown and Levinson's positive face)	equity rights (cf. Brown and Levinson's negative face)
social/interdependent perspective	identity face	association rights

(Spencer-Oatey 2000:15)

Within these two concepts, rapport management is explained with four components:

(1) *Quality face*, which is related to 'our sense of self-esteem', (2) *Identity face*, which refers to 'our sense of public worth', (3) *Equity rights*, which are associate with 'the notion of cost-benefit' and 'autonomy-imposition', and (4)*Association rights*, which are divided into 'interactional association – dissociation' and 'affective association - dissociation' (Spencer-Oatey 2000: 14-15). Bearing these elements of face and social rights in mind, she defined five domains of politeness:

> (1) Illocutionary domain.
>
> It concerns the rapport-threatening/rapport-enhancing implications of performing speech acts, such as apologies, requests, compliments, and so on.
>
> (2) Discourse domain.
>
> This domain concerns the discourse content and discourse structure of an interchange. It includes issues such as topic choice and topic management […], and the organization and sequencing of information.
>
> (3) Participation domain.
>
> This domain concerns the procedural aspects of an interchange, such as turn-taking […], the inclusion/exclusion of people present, and the use/non-use of listener responses (verbal and non-verbal).
>
> (4) Stylistic domain.
>
> This domain concerns the stylistic aspects of an interchange, such as choice of tone […], choice of genre-appropriate lexis and syntax, and choice of genre-appropriate terms of address or use of honorific.
>
> (5) Non-verbal domain.

> This domain concerns the non-verbal aspects of an interchange, such as gestures and other body movements, eye contact, and proxemics.
> (Adapted from: Spencer-Oatey 2000:19-20)

In the current study, participation domain and non-verbal domain from Spencer-Oatey (2000) will be considered mainly in the analysis. Listenership behaviour observed in the participants in the conversation data will also be examined from the perspective of rapport management with these concepts of face and rights.

2.4.2.3 Language and social identities

In this section, previous research on social identities and multiple identities in language users will be explored since natures of language users' identities in conversation of their native language and in interlanguage settings is relevant here.

In their study on social interaction in school counselling, Erickson and Shults (1982) described the multiple nature of social identities of a high school student as follows:

> Social identity can be thought of as a package with diverse contents. Technically it can be defined as a set of whose components are various attributes of social status on many different dimensions.
> (Erickson and Shults 1982: 13-14)

Through this, Erickson and Shults (1982) point out that all attributes of people, such as appearance, occupations, family structure, hobbies, and educational background, construct their social identities.

Zimmerman (1998: 87-88) defined identity as 'an element of context for talk in interaction' and distinguishes *situational identity*, which is changeable according to the context, with *transportable identity*, which is the fundamental and solid nature a person carries. Instability of identity has also been highlighted by Blommaert (2005: 207, original author's emphasis), where he proposed that 'we see identity not as a property or a stable category of individuals or groups, but as "particular forms of semiotic potential, organised in a repertoire"'. Maynard (2007) also drew attention to the relationship between identity and language.

> Language is a source for our individual identity. Although and because language is stabilizing and conventionalized, it yearns for a decentralizing, momentary, and creative formation. By manipulating this tension, we are able to create, mark, and transform our identities through languaging. By echoing multiple voices in a creative way, an individual person finds his or her own voices.
>
> (Maynard 2007: 71)

Through the presentation of him/herself in language in talk, a person constructs and reconstructs his/her identities continuously – even in their native language. By referring to Goffman's (1959) expressions *given* and *given off*, Coupland (2007) argues controllability of projection of identities as follows.

> When we 'give' expressions or self-identities, we have reasonably

> strong strategic control. When we 'give off' expressions or self-identities, we have low control and they 'leak' from our behaviour and our verbal and non-verbal displays.
> (Coupland 2007: 111)

At the same time, people are influenced by the other people's discourse strategies and discourse frameworks. This point is also described by Maynard as follows:

> […] borrowing someone else's style is a strategy through which we invite different identities into our own. And by assuming multiple identities, we engage in a creatively expressive activity that is languaging.
> (Maynard 2007: 95)

The participants' projections of their identities and their mutual influence or mutual adjustment in their listenership behaviour in an intercultural and interlanguage setting will be discussed at a later stage (see Section 6.1).

2.4.3 Acquisition of interlanguage pragmatics

The importance of culture in language learning has also been underlined in recent decades. Atkinson (1999: 625) announced that ' "[c]ulture" is a central yet underexamined concept in TESOL [Teachers of English to Speakers of Other Languages]' and Dodd (1998) introduced the concept of *intercultural competence*. In the US, the American Council on the Teaching of Foreign Language, Inc. (1999) defines 'National Standards in Foreign Language Education' with 5 Cs in foreign

language learning, namely communication, communities, comparisons, connections and cultures. While in the UK, CILT (The Chartered Institute of Logistics and Transports) The National Centre for Languages (2004) organised the INCA project, which stands for the Intercultural Competence Assessment project, in order to provide 'a definitive record of progress in key attributes of Intercultural Competence'. Kubota (2006) described culture with 3P, *products*, *practises* and *perspectives*, and emphasised the importance of learning intercultural communication in the language classroom.

Thomas (1983) introduced the term *pragmatic failure* as a breakdown in communication between people from different cultural backgrounds. Pragmatic failure was divided into two types; (1) *pragmalinguistic failure*, which is related to a part of grammar, and (2) *sociopragmatic failure*, which is a part of discourse. The former is assumed much more easily to overcome than the latter, since grammatical mistakes can be recognised much more easily than sociolinguistic misunderstandings.

The current study will discuss whether there is any point where misunderstandings and pragmatic failure between the participants can occur because of the differences in their listenership behaviour.

2.5 Summary

Through the literature review in this chapter, four critical themes for the current study have been reviewed: namely CA, context, response tokens and intercultural communication. Some of the key concepts in CA, such as turn-taking organisation, speaker selection, turn and floor, and discourse framework were reviewed. The classification of conversational gestures was also made in this section. The notion of context, the concept of contextualisation, and social interactional context mapping

were studied, since the relationship between conversation style and the context where conversation occurs is one of the central issues here, previous research on the use of response tokens in native speakers of English and Japanese and learners of these two languages was reviewed. The categorisation of response token was also made. Theories in intercultural communication including cultural values, interlanguage pragmatics, politeness and social identities, were also studied for discussion in the current research. Based on these theories, the research methods for the current study were established and a pilot study was conducted. They will be discussed in the next chapter.

Chapter 3 Research: Pilot study

3.0 Introduction

There are two research methods applied to the current study: (1) observation with video-recorded data; and (2) a time-related multimodal corpus-based approach. In this chapter, a pilot study conducted with data from two 10-minute conversations will be reported, and a time-related multimodal corpus analysis developed through the pilot study will also be explained. In addition, the concept of leadtime introduced in the course of the pilot study will be outlined.

3.1 Research method

Four techniques of research in CA are defined by Heritage (1984b):

> (1) the use of interviewing techniques in which the verbal formulations of subjects are treated as an appropriate substitute for the observation of actual behaviour;
> (2) the use of observational methods in which data are recorded through field notes or with pre-coded schedules;
> (3) the use of native intuitions as a means of inventing examples of interactional behaviour;
> (4) the use of experimental methodologies involving the direction or manipulation of behaviour.
>
> (Heritage 1984b: 236)

Anderson (1998: 8) defined four categories of methods in educational research, descriptive, explanatory, generalization and theoretical. A multimedia corpus- based

approach can be classified as Heritage's second technique for research on CA, namely observation with recorded data, and as descriptive research in Anderson's definition, which is used for clarifying what is happening or has happened in real life conversations. This current research can be termed descriptive research since the aim is to capture and describe the features of spoken discourse in English as a first language and English as an interlanguage in an academic setting. A comparison between these two contexts will also be conducted based on a multimodal corpus analysis which has been applied as the investigation's central method.

Triangulation, the idea of combining several approaches in a piece of research, is supported by a number of writers on research methods (Bryman 1988, Cohen et al. 2000, Hopkins 2002). It can be classified into seven types: data triangulation, investigator triangulation, theory triangulation, methodological triangulation, interdisciplinary triangulation, time triangulation and location triangulation (Brown & Rodgers 2002: 244). Essentially, triangulation is effective because it can improve validity and reliability of research. Validity is defined as 'the degree to which the results can be accurately interpreted', and reliability as 'the degree to which the results of a study are consistent' (Brown & Rodgers 2002: 241).

Methodological triangulation and data triangulation are used into consideration in this project. Two methods, namely corpus analysis and ethnographic observation for methodological triangulation, have been implemented. For data triangulation, two 10-minute conversation data sets are analysed in the pilot study while four 39-minute conversation data sets are examined in the main study. The amount of data analysed is still small and generalisation cannot be made from the results here. However, for the scalability of the research, larger data sets are used in

the main study. The purpose is not to provide generalisable results but to describe occurrences of the use of response tokens in relation to turn-taking structure within a particular academic setting.

3.1.1 Video-recorded data

Due to the innovations of information technology, researchers can deal with both audio and visual data for CA. Heath (1997) noted the importance of video-recorded observation in CA as follows:

> The possibility of capturing aspects of the audible and visible elements of 'in situation' human conduct as it arises within its natural habitats provides researchers with unprecedented access to social actions and activities. With ethnomethodology and conversation analysis, the technology opens up the possibility of developing a sociology which takes the visual, material as well as vocal aspects of human interaction seriously, as a topic for investigation and analysis.
>
> (Heath 1997: 278-280)

Face to face dyad conversations in English were video-recorded for the current study. With the visual and audio data, the participants' listenership behaviour was analysed with a time-related multimodal corpus approach as stated in research questions in Section 1.3.

3.1.2 Corpora of spoken language

As described in Chapter One (see Section 1.3), corpora are used not only in studies on

grammar and lexis within written language, but also, more recently, with pragmatic research in conjunction with spoken language.

Aston and Burnard (1998) summarised a list of corpora of English which have been developed since the 1960s. Based on their study, several major corpora have been categorised into seven groups: (1) Geographical corpora (Garside & Leech 1982, Greenbaum 1991, Johansson & Jahr 1982, Kucera & Francis 1967, Peters 1986), (2) Spoken language corpora (Carter & McCarthy 2006, Chafe et al. 1991, Greenbaum & Svartvik 1990, Handford 2007, Nesi 2000), (3) Mixed corpora (Aston & Burnard 1998, Ide & Macleod 2001, Renouf 1984), (4) Historical corpora (Kyto & Rissanen 1988), (5) Child and learner corpora (Granger 1993, O'Donoghue 1991), (6) Genre and topic-specific corpora (Davison 1993, McPherson & Herxheimer 2001, Schonell et al. 1956), and (7) Multilingual corpora (Armstrong-Warwick et al. 1994).

In terms of spoken corpora, the London-Lund corpus can be recognised as the earliest spoken corpus and was established through two projects: the Survey of English (SEU) at University College London in 1959 by Randolph Quirk and the Survey of Spoken English (SSE) at Lund University in 1975 by Jan Svartvik (Greenbaum & Svartvik 1990). A hundred spoken text data is tagged and stored in the London-Lund Corpus with metadata such as *text category*, *year of recording* and *speaker category*. Based on the London-Lund corpus, Stenstrom (1990) categorises and analyses discourse items such as *pardon*, *sorry* as apologies and *kind of*, *sort of* as hedges.

In the 1990s in the US, the Corpus of Spoken American English (CSAE) was published at the University of California, Santa Barbara, which includes approximately 200,000 words. The corpus consists of transcription and sound data of

about thirty hours of conversations in Standard American English (Chafe et al. 1991: 69) , which are provided in both a book and CD-ROMs and suitable for analysing the text from a discourse point of view.

A few years after CSAE, the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) was launched through the CANCODE project organised by Ronald Carter and Michael McCarthy in the School of English Studies at Nottingham University together with Cambridge University Press (McCarthy 1998). The CANCODE is described as 'a unique collection of five million words of naturally-occurring, mainly British (with some Irish) spoken language, recorded in everyday situations' (Carter & McCarthy 2006: 11). What makes CANCODE unique is not only the collection of everyday spoken language, but also the fact that the corpus is based on a *genre approach* and stored with the conversations organised into various genres. McCarthy (1998) divided the genres into five categories as we have seen in the earlier chapter: transaction, professional, pedagogical, socialising and intimate.

A small spoken English corpus of about 20,000 words in total has been developed for the current study, and therefore the integration of corpus-based study in conjunction with conversation analysis is implemented.

3.1.3 Coding systems and annotation

There are three types of coding strategies to add information to the text of a corpus: mark-up, annotation and metadata (Adolphs 2006, McEnery et al. 2006). Corpus mark-up is a coding system to provide information 'about' the text. The two mark-up systems, namely TEI (the Text Encoding Initiative) and CES (the Corpus Encoding Standard), are widely used, and languages used for developing websites such as SGML (Standard Generalised Mark-up Language) and XML (Extensible Mark-up

Language) are based on TEI scheme (McEnery et al. 2006). Here is an example of TEI tags:

> <extent> Approximately 100 million words </extent>
>
> (McEnery et al. 2006: 24)

In the example above, the information surrounded by the extent tag expresses the size of the corpus as additional information about the text.

Annotation is analytical information, which is added to a text (Adolphs 2006) including POS (Part Of Speech) tags and parsing. POS tags are mainly used for analysing grammatical analysis as shown in the example from CANCODE below:

> And [Cand] the [Dthe] security [Nsg] guard [Nsg] was [VFpastBe] walking [VPpres] about [T] checking [VPpres] everything [Pind] was [VFpastBe] okay [Jbas] and [Cand] and [Canc] then…
>
> Key: [Jbas] adjective, base; [Nsg] noun, singular; [Cand] conjunction, coordinating; [Dthe] definite article; [VFpastBe] verb, finite, past; [VPpres] verb, particle, present; [Pind] pronoun, indefinite.
>
> (Adolphs 2006: 24)

The first annotation in the example above [Cand], for instance, expresses the 'conjunction *and*', and the second annotation [Dthe] identifies the 'definite article *the*'. Each word is grammatically annotated in the sample above in order to extract words by grammatical functions from the corpus.

The other type of information about corpora is *metadata*, which provides further information about the source of the text (Adolphs 2006), such as the aim of the

project where the text is used, or information about data sampling. In the current study, I have adopted the coding systems employed in the CANCODE (see Section 3.1.2), and added some new tags. Furthermore, based on the coding systems in the CANCODE, additional codes to refer to functions of response tokens and types of turn-taking structure are added to the transcripts in my research.

3.1.4 The multimodal annotation interfaces

Two multimodal annotation software packages were considered as research tools in the current research: (1) Transana, which is conversation analysis software developed by Chris Fassnacht at the University of Wisconsin, Madison, and (2) Digital Replay System (DRS), which is developed in the School of Computer Science and IT at the University of Nottingham.

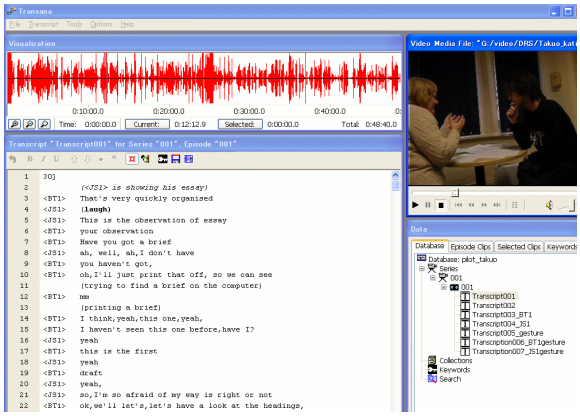


Figure 3.1-1 Transana

Transana consists of several functions such as a video viewer, a transcription area, a data control window and a sound wave bar as shown in Figure 3.1-1. Transana also enables transcribers to add time stamps on transcriptions. DRS in Figure 3.1-2 includes all of these functions. In addition, DRS has a function to combine separate

audio and visual data and align them on a track viewer. A drawback of DRS, however, is its operational environment. DRS requires users to install the Java Runtime Environment, which makes its operation heavier.

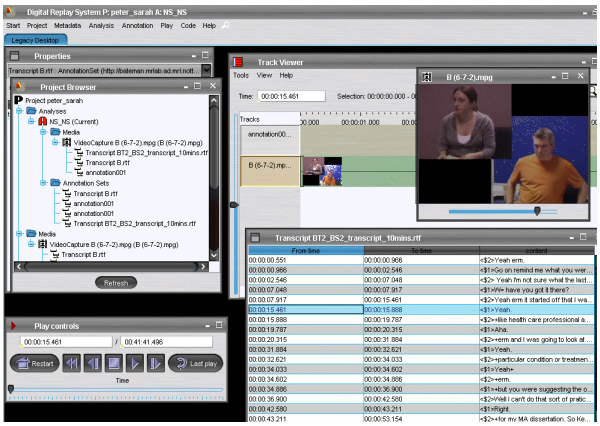


Figure 3.1-2 Digital Replay System

Although DRS has the advantages of being able to analyse more than one visual and audio data with time alignment, Transana was used mainly in the pilot study because of its usability and the fact that its functions are adequate for the analysis. Transana can be easily used since it functions without the Java Runtime Environment. In addition, this study does not involve the use of more complex functions such as combining data. DRS provides the time stamp function, which enables us to track the beginning and the end of each utterance. However, only the starting point of each utterance is time stamped in the current study since the distance between TTP and the point where a response token is uttered is measured for the analysis. Because of this, time stamps can be added and exported into text files easily by Transana.

3.1.5 Developing a mini corpus

Carter and McCarthy (1995) studied the features in spoken English conversation, and

concluded that spoken grammar is substantially different from written-based grammar. In their study, a mini-corpus, which was extracted from CANCODE, was examined. Carter and McCarthy say that even a small amount of spoken data can reveal features of spoken grammar.

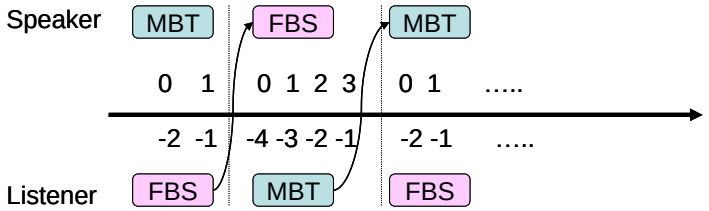
In the current research, a small amount of authentic conversation data was collected and stored in a self-developed mini corpus in order to establish a research method, and similarly a model for conversation analysis was developed to investigate linguistic features in the use of response tokens. Consequently, the mini corpus was developed by using multimodal annotation software, spread sheets and Microsoft Access.

3.1.6 Leadtime

3.1.6.1 Leadtime and floor transition point

As described in Section 1.2, leadtime is a new concept developed for the current study. Leadtime is applied to both listener status and speaker status in order to measure the length of time of speaker/listener status with turn transition point as a datum point. The datum point is described as 0 in leadtime. Leadtime is also used to describe the time distance between the point where a response token or a discourse marker is used and the floor transition point.

As illustrated below, participants are exchanging the floor in conversation. The time where either participant takes the floor is also defined as TTP, which is indicated as 0 in leadtime.



Keys: FBS= female British student, MBT= male British tutor

Figure 3.1-3 Floor-taking and leadtime

While a participant is in speaker status, her or his leadtime increases in positive numbers. In contrast, while a participant is in listener status, their leadtime is described in negative numbers. Two seconds before floor-taking, for example, is described '-2' in leadtime. The point where turn-taking has actually occurred is referred to as TTP, which can be differentiated from TRP, for TRP is defined as a point where there is the potential for speaker change (Schegloff 2007). Turn-taking may or may not occur at TRP.

3.1.6.2 Speaker turns and backchannel turns

I would like to distinguish *speaker turns* and *backchannel turns*. Speaker turns are similar in concept to floor-taking-turns in Furo (2001), whereas backchannel turns are a similar concept to non-floor-taking turns. When a participant has taken the floor, the participant has speakership in a conversation. This can be referred to as speaker turns. While a participant is listening to a speaker, the participant may take turns with response tokens. In this case, the participant still contributes to the conversation without taking the floor of the conversation. These are referred to as backchannel turns. To make it simple, I take any utterance with more than or equal to three words as a speaker turn. Any utterance less than three words is classified as a backchannel turn. I assume that there might be some exceptions, which can be determined by

listening to the conversation. The length of backchannel turns is defined based on observation of the conversation data.

Transcript 3.1-1 Speaker turns and backchannel turns

1 <FBT> so okay from this I have a picture of Japanese class working very quietly

2 <MJS> mm

3 (pause)

4 <FBT> is that right?

5 <MJS> yeah

6 <FBT> teacher says okay open your books and working on exercise three

7 <MJS> yeah

8 <FBT> and they are

9 <MJS> I have experienced like that

10 <FBT> so there is no speaking

11 <MJS> no especially

12 i= if students want to say something erm they can do but normally erm

13 just teacher says something

14 (pause)

Keys: FBT = female British teacher, MJS = male Japanese student.
(From the corpus in my research)

For example, in the transcription above, MJS (Male Japanese Student) said *mm* in line 2, where MJS takes a backchannel turn but not a speaker turn since MJS just gives a response token and FBT (Female British Tutor) keeps the floor of the conversation. In line 9, however, MJS has taken the floor of the conversation and in line 11 MJS secures the floor of the conversation.

Thus, essentially, speaker turns start with a participant's taking the floor and

last while the participant has retained the floor, whereas backchannel turns are short responses from a listener while they are listening to the speaker, which is related to listenership.

3.1.7 Definitions of head nods and hand gestures

Based on the classification of conversational gestures and types of vocal response tokens reviewed in Chapter Two, five items were selected and focused on in the pilot study: three verbal response tokens, *erm*, *yeah* and *mm*, and two visual response tokens, head nods and hand gestures as shown in the table below:

Table 3.1.7-1 Targeted items of response tokens

			Descriptions
Vocal response tokens	Erm		
	Yeah		
	Mm		.
Visual response tokens	Head nods	HN	Any vertical head movement
	hand gesture	HG	Any hand movements

Head nods are defined as any vertical head movement in the current study, and can be differentiated from head shakes which are defined as any horizontal head movements. In addition, hand gestures are defined as any hand movements, which are continuously delivered in conversation and sometimes a unit of hand gestures or a series of head nods is not clearly divided. When hand gestures and head nods were analysed here, time spaces between movements were counted. If there were several hand gestures or head nods within one second, they are counted as a hand gesture or a head nod since the time scale in seconds will be applied to time-related transcripts. If a hand gesture is continuously used for more than a second, this gesture is divided

into two gestures according to the time scale in second.

3.2 A pilot study

The pilot study was conducted for two reasons. The first reason was to narrow the research focus through the process of collecting and analysing actual data. The second reason was to establish a research method and to estimate what amount of data is required in order to address the research questions (see Section 1.3).

3.2.1 The data

3.2.1.1 Participants

In the pilot study, I conducted video-based observation on a British-British conversation and a British-Japanese conversation with the multimodal annotation tool, Transana. The British-British conversation data was recorded and initially transcribed by a research group in the School of English Studies at The University of Nottingham in August 2006. The British-Japanese conversation data was video-recorded by myself in February 2007. These conversations were recorded in the situation where only a video-camera was left on in the meeting rooms and the researcher left the room after setting the video camera in order to reduce the participants' consciousness for the recording:

Table 3.2.1-1 Pilot data and participants

	Participants		Type
	Tutor	Student	
Conversation Data 1 (C1)	Male British Tutor (C1_MBT)	Female British Student (C1_FBS)	British-British
Conversation Data 2 (C2)	Female British Tutor (C2_FBT)	Male Japanese Student (C2_MJS)	British-Japanese

As described in the table above, the British-British conversation involved a male

British tutor in his forties (MBT) and a female British student in her mid twenties (FBS) in a PhD supervision in English Studies at the University of Nottingham. I shall refer to this British-British conversation data as Conversation Data One (C1). The British-Japanese conversation was held by a female British tutor in her early fifties (FBT) and a male Japanese student (MJS) in his mid twenties, which was in a tutorial concerning assignments in the MA in English Language Teaching course at Nottingham Trent University. I shall refer to the second set of data as Conversation Data Two (C2) (see Table E.1.1-1).

3.2.1.2 Types of genres

Based on the mapping of social interactional contexts developed by Michael McCarthy (1998) and Ronald Carter (2004) (see the CANCODE in Section 3.1.2), the context type 'pedagogic-collaborative idea', especially face to face tutor-student supervision, was chosen as a focused context as illustrated in the table below:

Table 3.2.1-2 A targeted context

Goal types / ***Context type***	*Information provision*	*Collaborative task*	*Collaborative idea*
Pedagogic	English lecture	Individual computer lesson	Small group tutorial
Transactional	Commentary by museum guide	Choosing and buying a television	Chatting with hairdresser
Professional	Oral report at group meeting	Colleagues window dressing	Planning meeting at place of work
Socialising	Telling jokes to friends	Friends cooking together	Reminiscing with friends
Intimate	Partner relating the story of a film seen	Couple decorating a room	Siblings discussing their childhood

(Adolphs 2001: 49, Carter 2004: 150)

There are two reasons why this genre was chosen. The first reason is data availability; the dyad conversation data in supervisions at university is relatively accessible for me as a research student at university. The second reason is consideration of the needs of English. It is reported that the two areas where people often use English as second language or lingua franca are business and educational settings (Graddol 2006). English conversation routines and rules in these two areas can be highly prioritised in research in applied linguistics and English education.

3.2.1.3 Length of data

The total length of these conversation data is about 50 minutes each. As a pilot study, however, only the first 10-minutes of the two conversations were analysed, since the main aim of the pilot study was to develop a research method and implement the method with a small data set. In doing so, the research method and the model of conversation analysis applied to the current research was conceptualised.

3.2.2 The procedures

3.2.2.1 Data modification (1): Combining visual data and vocal data

With a multimodal annotation tool, Transana, participants' utterances and gestures were transcribed and annotated separately based on a time line. After timestamps were added to the transcriptions, four separated data sets, namely two participants' utterances and gestures, were combined with Microsoft Access.

Transcript 3.2-1 below is a final outcome of this modification process. Two participants' utterances and gestures are aligned with the timeline, and although there is no utterance at times 4 to 6 in the timeline in the sample transcript above, C1_FBS's utterance at time 3 lasts until time 6 in the timeline. When silent pauses occur in conversation, the symbol <$E>pause</$E> appears in the transcript. Since

the time scale is set in seconds, pauses which last more than 1 second are taken as silent pauses:

Transcript 3.2-1 Combined time-based transcription for C1

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1	MBT	0	-2		Go on remind me what you were doing.		Yeah erm.
2		1	-1				
3	FBS_F	-4	0	HB		SC/scratch forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4		-3	1	SC/wrist			
5		-2	2				
6		-1	3				
7	MBT_F	0	-1	HG	W= have you got it there?	HG	
8	FBS_F	-27	0				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+
9		-26	1				
10		-25	2	HF			
11		-24	3			HG	
12		-23	4				
13		-22	5			HT/HG	
14		-21	6				
15		-20	7	HN	Yeah.	HG	
16		-19	8	HB		HG	+like health care professional and lay person client patient+
17		-18	9			HG	
18		-17	10				
19		-16	11	HN			
20		-15	12		Aha.		+erm and I was going to look at how metaphor was used erm I suppose from a pragmatics perspective in trying to negotiate a kind of shared understanding of what patient symptoms and perhaps kind of explanations of+
21		-14	13			HG	
22		-13	14				
23		-12	15			HG	
24		-11	16	HN			
25		-10	17				
26		-9	18			HG	
27		-8	19			HG	
28		-7	20	HN			
29		-6	21			HG	
30		-5	22	HN			
31		-4	23				
32		-3	24		Yeah.	HG	
33		-2	25	HN			+particular condition or treatment options+
34		-1	26		Yeah+		
35	MBT_F	0	-1	HG	+but you were suggesting the other that you didn't want to do that <$G?>.		+erm.

Keys: C1= conversation 1 (NS-NS), C1_ MBT_transcript= Conversation 1 (NS-NS) male British tutor's transcription,
MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_gesture= male British tutor's gestures, FBS_gesture= female British student's gestures,
C1_FBS_transcript= Conversation 1 (NS-NNS) female British student's transcription,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
+= describe the continuous of the sentence, = unfinished sentence, <$G?>= inaudible sounds

The modification process can be divided into two steps. The first step is to combine visual and verbal data. The second step is to add floor and each participant's leadtime to a combined transcription. Procedures to combine participants' utterances and

gestures will be described first, and, in order to obtain this combined data, some modifications needed to be made.

Originally the data from Transana included a timeline. The timeline is labelled with 3 to 7 digit numbers, such as 551 in the first line in Transcript 3.2-2 below. The symbols <$1> <$2> indicate participants:

Transcript 3.2-2 Sample transcription from Transana

```
¤<551>      <$2>  Yeah erm.
¤<966>      <$1>  Go on remind me what you were doing.
¤<2546>     <$2>  Yeah I'm not sure what the last
                  proposal was that you saw erm it=
¤<7048>     <$1>  W= have you got it there?
¤<7917>     <$2>  Yeah erm it started off that I was
                  going to look at er the use of metaphor
                  in kind of health care interactions
                  between+
¤<15461>    <$1>  Yeah.
¤<15888>    <$2>  +like health care professional and lay
                  person client patient+
¤<19787>    <$1>  Aha.
```

Transcript 3.2-3 Timeline in seconds

Timeline_ original	Timeline (second)	Participant	Transcription
551	1	FBS	Yeah erm.
966	1	MBT	Go on remind me what you were doing.
2546	3	FBS	Yeah I'm not sure what the last proposal was that you saw erm it=
7048	7	MBT	W= have you got it there?
7917	8	FBS	Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+

Keys: FBS=Female British Student, MBT=Male British Tutor

Although the time stamps are tracked in milliseconds (ms) in Transana, the time scale is reduced and rounded into seconds. For example, 551 ms is rounded into 1 second as

shown in Transcript 3.2-3 above. By reducing the time scale, I am attempting to describe an overview of occurrences in turn-taking structure with the 10-minutes conversation data at the first stage. After the timeline was rounded, the verbal utterances of the two participants were transferred to two separated spread sheets. As shown in Transcript 3.2-4 and Transcript 3.2-5 below, verbal utterances of C1_FBS and C1_MBS were extracted to two spread sheets:

Transcript 3.2-4 Sample transcription of C1_FBS's utterances

Timeline	Part.	C1_FBS_Transcript
1	FBS	Yeah erm.
3	FBS	Yeah I'm not sure what the last proposal was that you saw erm it=
8	FBS	Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+
16	FBS	+like health care professional and lay person client patient+
20	FBS	+erm and I was going to look at how metaphor was used erm I suppose from a pragmatics perspective in trying to negotiate a kind of shared understanding of what patient symptoms and perhaps kind of explanations of+
33	FBS	+particular condition or treatment options+
35	FBS	+erm.
37	FBS	Well I can't do that sort of pratical data reasons in that I can't get hold of that kind of data in time+
43	FBS	+for my MA dissertation. So Kevin suggested that I can use some data which they have on it's like an independently run website and they do interviews with patients+
53	FBS	+and also I think there are interviews with doctors explaining certain er like conditions which are er described on the website.
63	FBS	Erm but it doesn't have like interactional+
66	FBS	+qualities.
68	FBS	Yeah.
69	FBS	And also it's erm well on the website at least I don't know if you can get hold of the data separately but erm it's edited as well so they only put on sections of the interview+
79	FBS	+which isn't ideal but erm maybe <$G?>.

Keys: C1_FBS= Conversation 1 (NS-NS) female British student, FBS= female British student

Transcript 3.2-5 Sample transcription from C1_MBT's utterances

Timeline	Part.	C1_MBT_Transcript
1	MBT	Go on remind me what you were doing.
7	MBT	W= have you got it there?
15	MBT	Yeah.
20	MBT	Aha.
32	MBT	Yeah.
34	MBT	Yeah+
35	MBT	+but you were suggesting the other that you didn't want to do that <$G?>.
43	MBT	Right.
53	MBT	Right.

Keys: C1_MBT= Conversation 1 (NS-NS) male British teacher, MBT= male British tutor

These data sets were combined with participants' gestures at a later stage. The data includes the timeline in seconds in the first column, identifications of participants in the second column and participant's verbal utterances in the third column. Response tokens, such as *yeah erm* at the first second and *erm* in 35 seconds in the transcript of

C1_FBS's utterances, are also transcribed and time-stamped.

Annotation symbols used in CANCODE (Adolphs 2006) were also applied to the transcriptions here. The plus symbol + indicates a continuous sentence and the equal symbol = signals an unfinished sentence. <$G?> indicates inaudible sounds and <$E>pause</$E> describes silence in conversation.

Conversational gestures of each participant were also transcribed and time-stamped separately. Several abbreviations were used to express conversational gestures:

HG	hand gesture (any hand gesture observed)
HN	head nods (any vertical head movement observed)
HS	head shake (any horizontal; movement observed)
HF	head forward
HB	head back
HT	head turn to the partner
SC	self-comfort/part of body or things used for SC

Conversational gestures of each participant were also transcribed and time stamped. Although these items in conversational gestures have been observed and transcribed, only HG and HNs were focused on in this analysis. In addition, the transcribed data of participants' gestures were exported into spread sheets separately as shown in Transcript 3.2-6 and Transcript 3.2-7 below. Definitions of hand gestures and head nods were described in the previous section (see 3.1.7). When hand gestures and head nods were annotated, the time spaces between movements were counted. In a case

where more than one hand gesture was observed in one second, they were reduced into one hand gesture. The same rule was applied to counting head nods.

Transcript 3.2-6 Sample gesture annotation of C1_FBS

Timeline	C1_FBS_gesture
3	SC/scratch forehead
7	HG
11	HG
13	HT/HG
15	HG
16	HG
17	HG
21	HG
23	HG
26	HG
27	HG
29	HG
32	HG
37	SC/sleeve
39	HG
41	HG
46	HG
49	HG
52	HG

Keys: C1_FBS= Conversation 1 (NS-NS) female British student, SC/scratch forehead= self comfort with scratching forehead, HG= hand gestures, HT/HG = head turns and hand gestures, SC/sleeve= self comfort with sleeve

Transcript 3.2-7 Sample gesture annotation of C1_MBT

Timeline	C1_MBT_gesture
3	HB
4	SC/wrist
7	HG
10	HF
15	HN
16	HB
19	HN
24	HN
28	HN
30	HN
33	HN
35	HG
40	HN
44	HN
51	HN
57	HN
61	HN
66	HN
67	HG

Keys: C1_MBT= Conversation 1 (NS-NS) male British tutor, HG= head back, SC/wrist = self comfort with wrist, HG= hand gestures, HF= head forward, HN= head nods

These four data sets, namely two participants' utterance data and two participants' gesture data, were imported into Microsoft Access database in order to be combined with the timeline as a primary key.

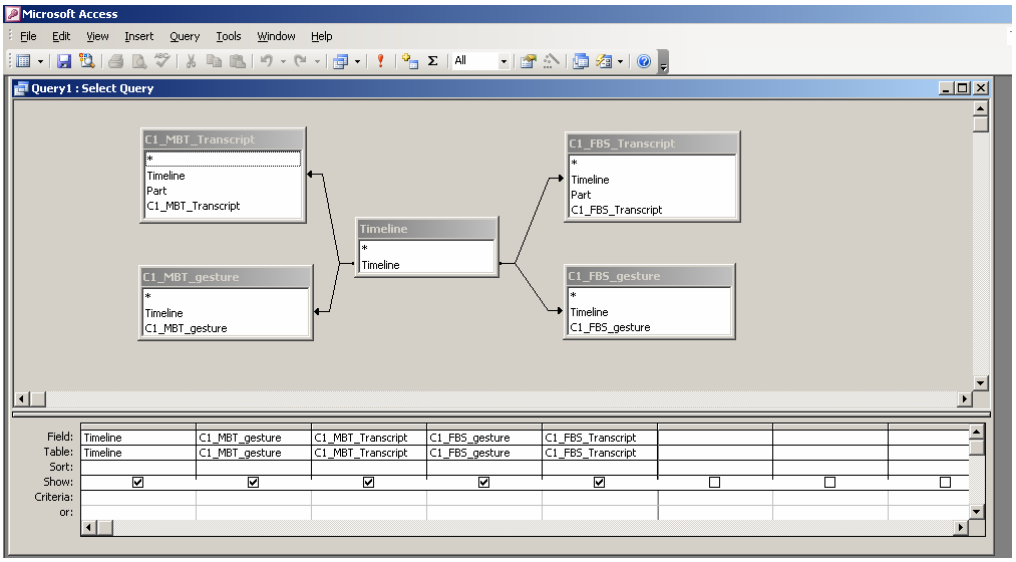


Figure 3.2-1 Relational database with Microsoft Access

These four sub-tables were combined with a timeline as shown in Figure 3.2-1 above. A main table including columns for timeline, C1_MBT_gesture, C1_MBT_transcript, C1_FBS_gesture and C1_FBS_transcript, was created. The main table was then exported into a spread sheet as shown in Transcript 3.2-8 below: the first column shows the timeline in seconds and the second and the third columns are C1_MBT's utterances and gestures. C1_FBS's utterances and gestures are allocated in the fourth and the fifth columns. The same modification procedures were taken with the British-Japanese conversation data (C2) in the pilot study. The two main tables of C1 and C2 were used for the analysis.

Transcript 3.2-8 A main table

Timeline	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1		Go on remind me what you were doing.		Yeah erm.
2				
3	HB		SC/scrach forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4	SC/wlist			
5				
6				
7	HG	W= have you got it there?	HG	
8				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+

Keys: C1_MBT_gesture= Conversation 1 (NS-NS) male British tutor's gestures,
C1_MBT_transcript= Conversation 1 (NS-NS) male British tutor's transcriptions,
C1_FBS_gesture= Conversation 1 (NS-NS) female British student's gestures,
C1_FBS_transcript= Conversation 1 (NS-NS) female British student's transcriptions,
HB= head back, SC/wrist = self comfort with wrist, HG= hand gestures,
SC/scratch forehead= self comfort with scratching forehead,

After obtaining the main tables, additional modifications were processed. The process of adding columns for floor and leadtime to a main table will be explained in the following section.

3.2.2.2 Data modification (2): Adding leadtime

Three columns were inserted into the main tables. One is a column for floor to indicate points where participants have taken the floor of the conversation, and another two columns were added to express participants' leadtime separately. In order to insert leadtime of the participants, each participant's TTPs were annotated as a datum time. As described in the previous chapter, TTP is a point where either of the participants has taken the floor of the conversation. A column for the floor of the conversation was then inserted as shown in Transcript 3.2-9 below. Further, since MBT has taken the floor at times 1 and 7 in the timeline, MBT_F is indicated in the column. At times 3 and 8 in the timeline, FBS_F is indicated to express FBS's taking the floor. Both of the participants' utterances are shown at the first second. C1_MBT

says ‘Go on remind me what you were doing’ in the first row and C1_FBS utters ‘Yeah erm’ within the same second.

Transcript 3.2-9 Sample 1

Timeline	Floor	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1	MBT_F		Go on remind me what you were doing.		Yeah erm.
2					
3	FBS_F	HB		SC/scrach forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4		SC/wlist			
5					
6					
7	MBT_F	HG	W= have you got it there?	HG	
8	FBS_F				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+

Keys: MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
FBS= female British student, MBT= male British tutor

The sequence of the utterances in the same second cannot be recognised in the table above since each participant’s transcription is allocated into two separate columns. This latter point is a drawback of the method since the order of utterances in the same second disappears in the transcript. However, at this stage I would like to adhere to this method with the timeline in seconds in order to have an *overview* of the use of response tokens in relation to the transition from listener to speaker.

The next process was to indicate TTP in the timeline, which is described as 0 in leadtime. While either of the participants is holding the floor, the leadtime increases. In sample 2 below, for example, MBT takes the floor at the first second and holds the floor until 2 seconds, so I put 0 in the first row and 1 in the second row in the third column. These numbers express C1_MBT’s leadtime. In turn, FBS holds the floor from 3 seconds to 6 seconds, so I put 0 at 3 seconds and increment the leadtime up to 3 in the column for FBS’s leadtime. After FBS’s taking the floor, MBT then takes the

floor again, so I put 0 at 7 seconds in MBT's leadtime. However, FBS takes it back soon after, and so I put 0 at 8 seconds in FBS's leadtime. As described above, participants' TTP is a datum time to define each participant's leadtime.

Transcript 3.2-10 Sample 2

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1	MBT_F	0			Go on remind me what you were doing.		Yeah erm.
2		1					
3	FBS_F		0	HB		SC/scratch forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4			1	SC/wrist			
5			2				
6			3				
7	MBT_F	0		HG	W= have you got it there?	HG	
8	FBS_F		0				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+

Keys: MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
FBS= female British student, MBT= male British tutor

As shown in sample 2 above, while participants are in *speaker* status, they have positive numbers in their leadtime. Negative numbers in leadtime indicate that participants are in *listener* status. For example, in sample 3 below, negative numbers such as -2 and -1 are added in the first two rows in FBS's leadtime, which means FBS is in listener status for 2 seconds before FBS takes the floor again. From 3 seconds to 6 seconds, negative numbers are seen in MBT's leadtime column, which also indicates that MBT is in listener status for 4 seconds:

Transcript 3.2-11 Sample 3

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1	MBT_F	0	-2		Go on remind me what you were doing.		Yeah erm.
2		1	-1				
3	FBS_F	-4	0	HB		SC/scratch forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4		-3	1	SC/wrist			
5		-2	2				
6		-1	3				
7	MBT_F	0	-1	HG	W= have you got it there?	HG	
8	FBS_F	-27	0				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+

Keys: MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
FBS= female British student, MBT= male British tutor

By adding the leadtime of each participant, the length of time of participants' speaker status and listener status can be measured. In addition, leadtime can be used to analyse the time distance between a point where particular response token items are used and TTP. At 15 seconds in Transcript 3.2-12 below, for instance, C1_MBT gives a head nod (HN), which occur in MBT's listener status since MBT_leadtime has a negative number at that time. In other words, it can be interpreted that this HN occurs 20 seconds prior to MBT's next floor-taking. At the same time, C1_MBT utters *yeah* as shown in C1_MBT's transcript in the sixth column. C1_MBT has delivered HNs several times before he takes the floor at 35 seconds in the timeline. The tendency can be seen that the more C1_MBT is close to his next TTP, the more he gives HNs. In terms of C1_FBS, she has taken the floor at 3 seconds in the timeline and used many hand gestures (HGs) while she is in speaker status.

Transcript 3.2-12 A time-related transcription for C1

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
1	MBT	0	-2		Go on remind me what you were doing.		Yeah erm.
2		1	-1				
3	FBS_F	-4	0	HB		SC/scratch forehead	Yeah I'm not sure what the last proposal was that you saw erm it=
4		-3	1	SC/wrist			
5		-2	2				
6		-1	3				
7	MBT_F	0	-1	HG	W= have you got it there?	HG	
8	FBS_F	-27	0				Yeah erm it started off that I was going to look at er the use of metaphor in kind of health care interactions between+
9		-26	1				
10		-25	2	HF			
11		-24	3			HG	
12		-23	4				
13		-22	5			HT/HG	
14		-21	6				
15		-20	7	HN	Yeah.	HG	
16		-19	8	HB		HG	+like health care professional and lay person client patient+
17		-18	9			HG	
18		-17	10				
19		-16	11	HN			
20		-15	12		Aha.		+erm and I was going to look at how metaphor was used erm I suppose from a pragmatics perspective in trying to negotiate a kind of shared understanding of what patient symptoms and perhaps kind of explanations of+
21		-14	13			HG	
22		-13	14				
23		-12	15			HG	
24		-11	16	HN			
25		-10	17				
26		-9	18			HG	
27		-8	19			HG	
28		-7	20	HN			
29		-6	21			HG	
30		-5	22	HN			
31		-4	23				
32		-3	24		Yeah.	HG	
33		-2	25	HN			+particular condition or treatment options+
34		-1	26		Yeah+		
35	MBT_F	0	-1	HG	+but you were suggesting the other that you didn't want to do that <$G?>.		+erm.

Keys: C1= conversation 1 (NS-NS), C1_ MBT_transcript= Conversation 1 (NS-NS) male British tutor's transcription, MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime, MBT_gesture= male British tutor's gestures, FBS_gestures= female British student's gestures, C1_FBS_transcript= Conversation 1 (NS-NNS) female British student's transcription, MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking, += describe the continuous of the sentence, = unfinished sentence, <$G?>= inaudible sounds

The same modification process was conducted on the British-Japanese conversation data (Conversation 2, C2), in order to compare these two data, C1 and C2, in the data analysis.

3.2.3 Data analysis

3.2.3.1 Global pattern analysis

The collected data was analysed both quantitatively and qualitatively. In the global pattern analysis, occurrences of response tokens and length of listener/speaker status were counted with leadtime. Some referential statistics, such as means, standard deviations, and variances, were also used in the global pattern analysis.

In addition, the three factors described below were considered:

1. The numbers of speaker turns and backchannel turns of each participant in the conversations,
2. The time distance between targeted response token items and TTP; and two visual response tokens, namely hand gestures (HGs), head nods (HNs), and their verbal response tokens, such as *erm*, *yeah* and *mm*,
3. The means, variances, and standard deviations (SD) of the targeted response tokens.

Furthermore, turn management with response tokens of each participant were analysed from these three aspects. The results were summarised in tables and graphs to visualise these features of conversation. The idea of counting response tokens comes from Drummond and Hopper's (1993) research, which was reviewed in Chapter Two. Based on Drummond and Hopper's (1993) research, the analysis here was done with the concept of leadtime, which can be applied to measure the time distance between particular response tokens and TTP in conversation.

3.2.3.2 Turn-structural analysis

Two previous qualitative studies on response tokens were taken into consideration. As reviewed in Chapter Three, Schegloff (1984) categorises conversation gestures, and I adapted the definition (the first three categories listed below) with one additional category (the fourth category):

> (1) a current nonspeaker is making a move for a turn at talk next
>
> (2) a current nonspeaker tries to communicate without interrupting
>
> (3) a current speaker is interrupted, and yields to the interrupter
>
> (4) a current speaker is describing what he is saying with hand gesture
>
> (Adapted from Schegloff 1984: 271)

The function (1) can be also described as *floor seekers*, which was introduced by Sacks (1992) (see Section 2.1.2).

As for response tokens, the idea of four basic functions of response tokens defined by O'Keeffe and Adolphs (2008) was taken into the current study:

> (1) Continuers [CN]: Maintaining the flow of discourse.
>
> (2) Convergence tokens [CV]: Markers of agreement/convergence.
>
> (3) Engaged response tokens [EN]: Markers of high engagement where addressee(s) respond on an affective level to the content of the message.
>
> (4) Information receipt tokens [IR]: Markers of points in the discourse where adequate information has been received.
>
> (O'Keeffe & Adolphs 2008 : 84)

Abbreviations of these four discourse functions were defined: continuers as CN, convergence tokens as CV, engaged response tokens as EN and information receipt tokens as IR. I shall refer to these functions collectively as *discourse function* (DF) in the current study.

Based on studies in conversation analysis (Sacks 1992, Sacks et al. 1974, Schegloff 2007) and discourse analysis (Carter & McCarthy 1997, Carter & McCarthy 2006, O'Keeffe et al. 2007), the functions of response tokens and utterances in relation to turn-taking structure and speaker change are also defined as follows:

(1) Floor-taking [FT]: Taking the floor of the conversation

(2) Floor seeker [FS]: Trying to take the floor of the conversation

(3) Listenership [LS]: Maintaining listener status

(4) Floor giving [FG]: Giving the floor of the conversation to other participants

These functions are referred to here as *conversation function* (CF). In the qualitative data analysis, participants' listener status was focused on. The targeted response tokens were analysed in terms of these two levels of functions of response tokens, namely conversation function (CF) and discourse function (DF), in relation to turn-taking structure.

Transcript 3.2-13 below is an example of C1_MBT's listener status. Based on the time-related transcripts obtained from the procedures reviewed in the previous section, two columns, namely C1_MBT_CF and C1_MBT_DF, were inserted in the fourth and the fifth columns.

Transcript 3.2-13 Conversation function and discourse function

Timeline	Floor	MBT_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
9		-26				
10		-25	LS	CN	HF	
11		-24				
12		-23				
13		-22				
14		-21				
15		-20	LS	CN	HN	Yeah.
16		-19	LS	CN	HB	
17		-18				
18		-17				
19		-16	LS	CN	HN	
20		-15	LS	CN		Aha.
21		-14				
22		-13				
23		-12				
24		-11	LS	CN	HN	
25		-10				
26		-9				
27		-8				
28		-7	LS	CN	HN	
29		-6				
30		-5	LS	CN	HN	
31		-4				
32		-3	FS	IR		Yeah.
33		-2	FS	IR	HN	
34		-1	FS	IR		Yeah+
35	MBT_F	0	FT		HG	+but you were suggesting the other that you didn't want to do that <$G?>.

Keys: MBT_leadtime= male British tutor's leadtime,
C1_ MBT_CF= Conversation 1 (NS-NS) male British tutor's conversation function,
C1_ MBT_DF= Conversation 1 (NS-NS) male British tutor's discourse function,
C1_ MBT_gesture= Conversation 1 (NS-NS) male British teacher's gestures,
C1_ MBT_transcript= Conversation 1 (NS-NS) male British tutor's transcriptions,
LS= listenership, FS= floor seeker, FT= floor-taking, CN= continuers, IR= information receipt tokens,
HF= head forward, HN= head nods, HB= head back, HG= hand gestures,
MBT_F= male British tutor floor-taking,
+= describe the continuous of the sentence, = unfinished sentence, <$G?>= inaudible sounds

At -20 in C1_MBT's leadtime, for instance, C1_MBT utters *yeah*, which can be interpreted as LS (listenership) in conversation function, and taken as CN (continuers) in discourse function at the same time. These two functions are indicated here in the fourth and fifth columns. At -3 in the leadtime, another *yeah* is observed. This response token in turn can function as floor seeker in conversation function. At the same time, this can send a message that C1_MBT has received enough information as information receipt tokens in discourse function. A few seconds later, C1_MBT takes

the floor of the conversation.

In this example, a strategy *before* floor-taking, which is a sequential process of continuers, information receipt tokens and the floor-taking, were observed. With these two layers of functions, features of listenership behaviour were analysed in the qualitative analysis.

3.3 Findings from the pilot study

In this section, findings from the pilot study will be described. In the global pattern analysis, the numbers of speaker turns and backchannel turns in each conversation were counted. Both participants' verbal and visual response tokens were also counted in terms of five-second intervals of leadtime in relation to TTP. In the turn structure analysis, floor exchanges with the use of verbal response tokens were identified first. Finally, collocation of visual and vocal response tokens was also analysed.

3.3.1 Salient findings from the global pattern analysis

Four salient findings were reported from the global pattern analysis:

1. In the British-British conversation, the numbers of speaker turns and backchannel turns in conversation were more balanced between the participants than in the British-Japanese conversation.
2. In the British-British conversation, HGs (hand gestures) were observed more often at TTP. The male Japanese student in the British-Japanese conversation rarely used HGs, whereas the female British tutor in the British-Japanese conversation used HGs all the time when she took the floor of the conversation.

3. The female British student in the British-British conversation used *erm* at TTP several times, while the male British tutor in the British-British conversation used *yeah* as a strategy used at TTP.
4. In the British-Japanese conversation, the male Japanese student used *mm* 50 times in the 10-minutes of conversation data. He also used *mm* constantly when in listener status. Conversely, the female British tutor in the British-Japanese conversation did not use *mm* at all. Further, the male British tutor and the female British student rarely used *mm* in the British-British conversation.

The details of findings will be discussed in the next sections with tables and graphs from the data.

3.3.1.1 Findings of the number of speaker turns and backchannel turns

Table 3.3.1-1 below summarises the numbers of turn exchanges in C1, the British-British conversation. The table shows how many times each participant takes speaker turns and backchannel turns, and how many times they fail to take speaker turns in C1. Some occurrences are observed where a listener utters some words to take the floor but the turn construction unit is not completed, and the current speaker retains the floor. These occurrences are defined as failures of floor-taking. In addition, the frequency of silent pauses which occur in the conversation and how long these pauses are in total is shown in the last column.

Table 3.3.1-1 Numbers of taking speaker turns and backchannel turns in C1

	Speaker turns	Speaker turn failed	Backchannel turns(vocalised	Pause
C1_MBT	20	1	122	2 (4sec)
C1_FBS	19	3	119	

Keys: C1= conversation 1 (NS-NS), C1_ MBT = Conversation 1 (British-British Conversation) male British tutor, C1_ FBS = Conversation 1 (British-British Conversation) female British student

C1_MBT takes speaker turns 20 times and vocalises backchannel turns 122 times. C1_FBS takes speaker turns 19 times and vocalises backchannel turns 119 times, which are almost the same amount as C1_MBT. There are only 2 silent pauses in the 10-minutes British-British conversation. The total length of silent pauses is 4 seconds.

Conversely, unequal participation is observed in C2, the British-Japanese conversation. As shown in Table 3.3.1-2 below, C2_FBT takes speaker turns 25 times, which is three times more than that of C2_MJS. C2_FBS takes vocalised backchannel turns 109 times, which is twice more than that of C2_MJS.

Table 3.3.1-2 Numbers of taking speaker turns and backchannel turns in C2

	Speaker turns	Speaker turn failed	Backchannel turns(vocalised	Pause
C2_FBT	25	0	109	18 (185sec)
C2_MJS	8	0	53	

Keys: C2= conversation 2 (NS-NNS), C2_ FBT = Conversation 2 (British-Japanese conversation) female British tutor, C2_ MJS = Conversation 2 (British-Japanese conversation) male Japanese student

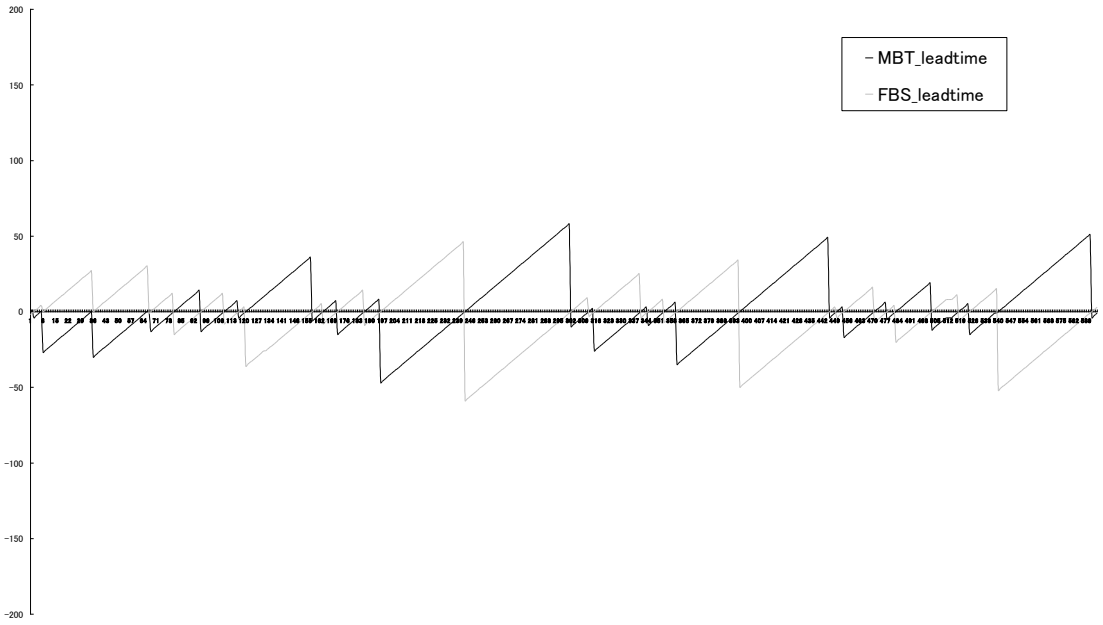
In terms of silent pauses, 18 pauses occur in the conversation and the total the length of time of the pauses is 185 seconds. This means that in the 10-minutes British-Japanese conversation, there are more than 3 minutes of pauses. The data indicates that in the British-Japanese conversation, C2_FBT talks most of the time while

C2_MJS listens, and there are many pauses between the utterances.

The existence of silence in conversation has been pointed out by several linguists and taken into consideration in conversation analysis (Sacks 1992, ten Have 2001). It is also reported that silent pauses have been observed more often in Japanese conversation (Hayashi et al. 2002, Maynard 1990, Mori 2002). Yamada (1997) refers to the particular kind of silence in Japanese conversation as *sasshi*, which is a silence used for taking time to guess what the partner is thinking, what should be said and with what timing. The silence *sasshi* is normally used with *gaze*, or, in other words *eye monitoring*. Whether the large amount of silence in C2, the British-Japanese conversation, is caused by the influence from this feature of Japanese conversation or not is an issue to be investigated.

3.3.1.2 Equality and inequality in turn-taking

Exchanges of the floor in conversation were described based on participants' leadtime as shown in Figure 3.3-1 below and Figure A.1-1 in the appendix A. The X axis is the timeline in seconds, and the Y axis is the length of time of leadtime in seconds, which has negative and positive values. When the line is drawn in the positive dimension in Y axis, the participant is in speaker status, in other words, holding the floor of the conversation. In approximately the first 80 seconds, for instance, C1_FBS is in speaker status. At the same time, C1_MBT is in listener status in the first 80 seconds, hence his leadtime is indicated in the black line below and is shown in the negative dimension on the Y axis. Additionally, many short floor exchanges of the floor between C1_FBS and C1_MBT occur in C1, and thus it can be said that the contributions of both participants in the conversation are balanced. As a result, the figure shows an 'equality' of turn-taking in conversation.



Keys: C1= conversation 1 (British-British conversation), y axis = leadtime (seconds), x axis = timeline (seconds), MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,

Figure 3.3-1 Numbers of taking speaker turns and backchannel turns in C1

In the British-Japanese conversation in Figure A.1-1 in the appendix, on the other hand, the numbers of floor exchanges are fewer than the British-British conversation. Relatively long floor exchanges by C2_FBT can be observed; C2_FBT with a black line is seen in the positive area most of the time while C2_MJS with a grey line is in the negative area in contrast. In addition, there are several points where both black and grey lines are shown in the negative dimension. That means both of the participants are not in speaker status; in other words, there are silent pauses when both participants have negative numbers in their leadtime. The figure showing from the exchanges of the floor in C2, the British-Japanese conversation, illustrates an inequality of turn-taking. Almost all the time C2_FBT holds the floor in this conversation.

In Chapter Two, Heritage's (1997) notion of symmetries in conversation was reviewed. He points out that equal participation in conversation is more often observed in daily conversations than in institutional and professional conversations. The pedagogic situation chosen for the current study, namely PhD or MA supervisions at university, can be categorised as an institutional conversation. It can be said that C2, the British-Japanese conversation, reflects this feature of institutional conversation more than C1, the British-British conversation.

3.3.1.3 Findings about hand gestures and head nods

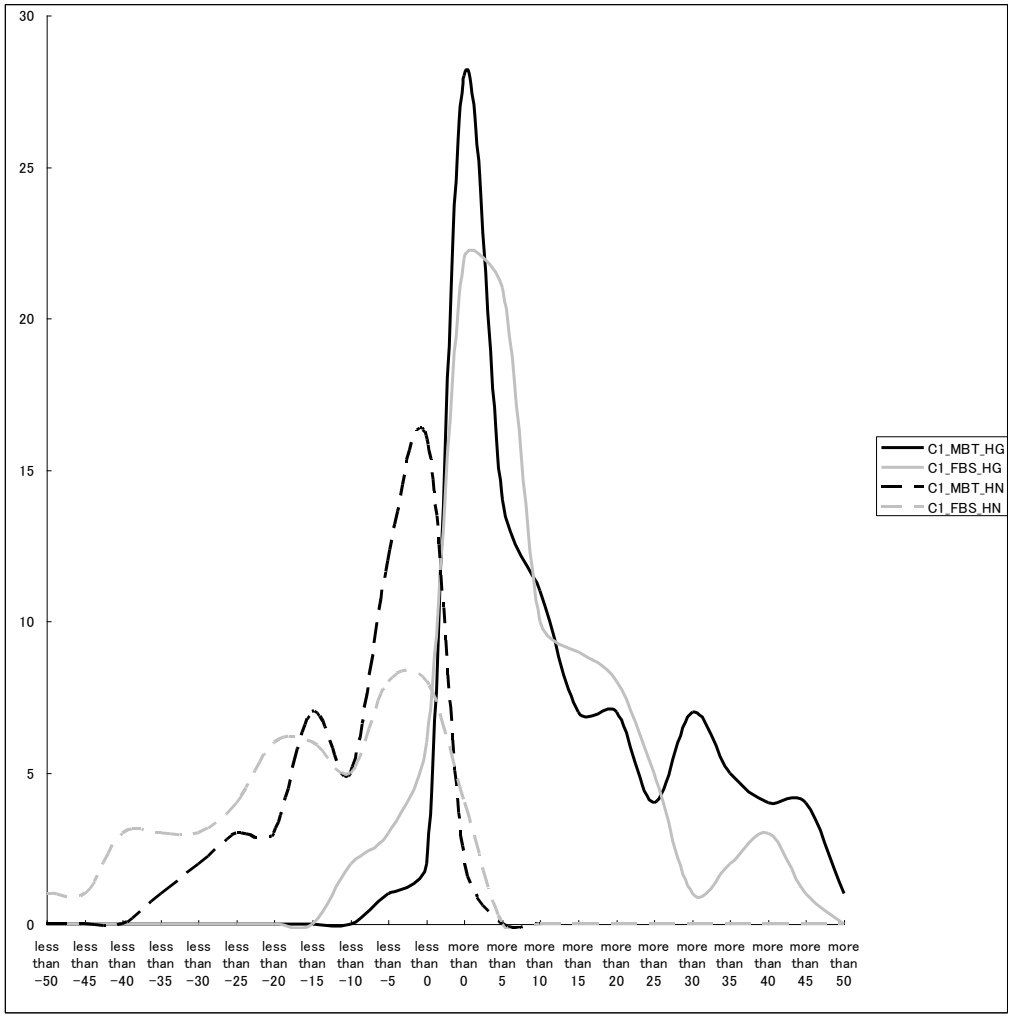
The numbers of HGs and HNs were counted based on each participant's leadtime. As explained previously in this paper, leadtime is defined as a time scale to describe time distance from TTP. Hence, leadtime 0 means a participant has just taken the floor of the conversation, leadtime 3 means that the speaker has been holding the floor for 3 seconds, and leadtime -3 means that the participant has been in listener status for 3 seconds before his or her next floor-taking.

Numbers of targeted conversational gestures were summarised in five-second intervals of leadtime in order to obtain an overview of the use of conversational gestures in relation to turn-taking. At 19 seconds in the timeline in Transcript 3.3-1 below, for instance, C1_MBT has a HN. This HN is put into a five-second time interval 'less than -15' since the MBT's leadtime at that moment is -16. That means C1_MBT uses a HN 16 seconds before he takes the next floor. In the same way, C1_FBS's HGs at 21 seconds is put into the leadtime scale 'more than 10' since FBS's leadtime at that moment is 13. This means C1_FBS uses HGs 13 seconds after she takes the floor of the conversation.

Transcript 3.3-1 HNs and HGs in C1

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
19		-16	11	HN			
20		-15	12		Aha.		+erm and I was going to look at how metaphor was used erm I suppose from a pragmatics perspective in trying to negotiate a kind of shared understanding of what patient symptoms and perhaps kind of explanations of+
21		-14	13			HG	
22		-13	14				
23		-12	15			HG	
24		-11	16	HN			

Keys: C1= conversation 1 (British-British Conversation),
MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_gesture= male British tutor's gestures, FBS_gestures= female British student's gestures,
C1_MBT_transcript=conversation 1 (British-British Conversation) male British tutor's transcription,
C1_FBS_transcript=conversation 1 (British-British Conversation) female British student's transcription,
HN= head nods, HG= hand gestures



Keys: HG= hand gestures, HN= head nods, C1= conversation 1 (British-British Conversation),
y axis = the numbers of HN and HG, x axis = ranges of leadtime (seconds),
C1_MBT_HG = C1 male British tutor's hand gestures, C1_FBS_HG = C1 female British student's hand gestures,
C1_MBT_HN = C1 male British tutor's head nods, C1_FBS_HN = C1 female British student's head nods

Figure 3.3-2 Numbers of HGs and HNs in C1

The summary of HGs and HNs in C1, the British-British conversation, scaled in five-second ranges of leadtime is shown in Figure 3.3-2 above and Table A.1.2-1 in the appendix A. From the table, it can be seen that the use of HGs is increasing at the moment when the participant takes the floor. This trend can be seen in both C1_MBT and C1_FBS, for the number of HGs in C1_MBT in the scale 'more than 0' is 28 times and that in C1_FBS is 22. The figure of HNs in C1_MBT around TTP is increased, and the number of C1_MBT's HNs in 'less than -5' is 12 and 'less than 0' is 16. C1_FBS's use of HNs also increases around TTP. However, in C1_FBS's case, HNs are used not only at the point of taking the floor but also when she is in listener status.

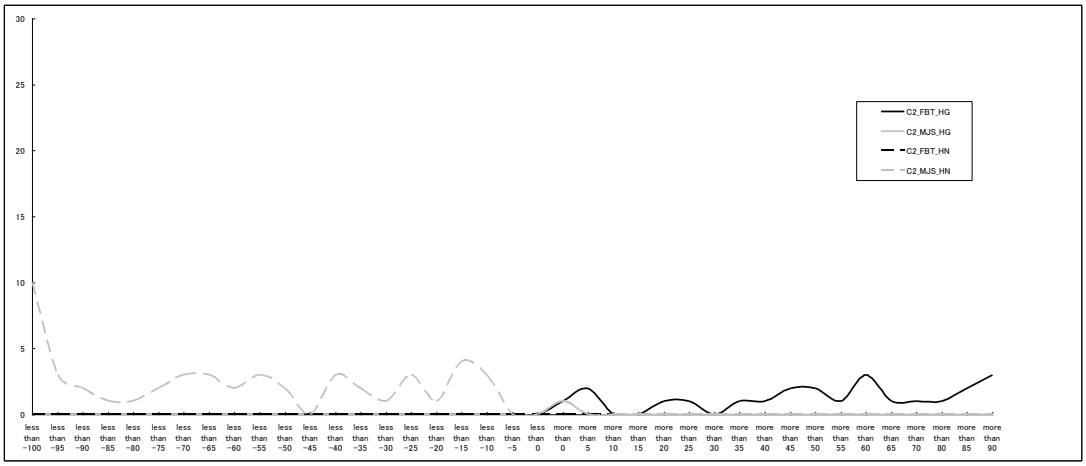
In Figure 3.3-2 above, we can also see the increasing trend of HGs and HNs at the floor boundary. The X axis is leadtime, which has negative and positive values to express listener and speaker status, and is scaled into five-second intervals. Alternately, the Y axis expresses how many times the gestures used at each five-second scale. Both HNs and HGs are used more frequently around the floor boundary.

However, the peaks of these two gestures use are slightly different. HNs are used before the participant takes the floor and HGs are used at the moment or soon after the participant takes the floor. This tendency is more emphasised in C1_MBT with black lines than C1_FBS with grey lines. C1_FBS's use of HNs (the broken grey line), on the one hand, is continuously observed even when she is in listener status and more than 35 to 40 seconds before she takes the next floor. C1_MBT's use of HNs (the black broken line), on the other hand, dramatically increases then from 10 to 5 seconds before he takes the floor.

For the comparison, HGs and HNs observed in C2, the British-Japanese

conversation, were analysed with the same method. The features about HGs and HNs from C2 data are different from that of C1. In Table A.1.2-2 in the appendix and Figure 3.3-3 below, HGs and HNs used by C2_FBT and C2_MJS are summarised. From the total numbers of these gestures at the bottom of the table, we can see that C2_FBT uses HGs 23 times in total, which is rather small amount if compared with C1 data where both of the participants use HGs more than 90 times in the same duration of conversation. C2_MJS also uses HGs only once in the 10-minutes conversation. The only one HG is used at the point when C2_MJS takes the floor.

In terms of HNs, C2_MJS uses HNs 50 times in total, which is almost the same amount that the participants in C1 have. At the same time, C2_FBT does not use HNs at all. C2_MJS's use of HNs is slightly increased 10 or 15 seconds before points where C2_MJS takes the floor, six times in 'less than -10' and 'less than -15'.



Keys: HG= hand gestures, HN= head nods, C2= conversation 2 (British-Japanese conversation),
y axis = the numbers of HN and HG, x axis = ranges of leadtime (seconds),
C2_FBT_HG = C2 female British tutor's hand gestures, C2_MJS_HG = C2 male Japanese student's hand gestures,
C2_FBT_HN = C2 female British tutor's head nods, C2_MJS_HN = C2 male Japanese student's head nods,

Figure 3.3-3 Numbers of HGs and HNs in C2

The graph of HGs and HNs in C2 is also described in Figure 3.3-3 above. The lines are rather moderate: C2_FBT's HGs with black solid line is continuously observed even more than 90 seconds after she takes the floor, while C2_MJS's HNs with grey broken line is seen constantly in the negative area. This means that he is in listener status even more than 100 seconds before he again takes the floor.

In Table A.1.2-3 in the appendix, I have summarised means, variances, and standard deviations of HGs and HNs in C1. The mean leadtime of C1_MBT's HG is 15.23 and C1_FBS is 10.45 although the figures of the variance and the standard deviation are large, such as 226.80 in variation and 14.98 in standard deviation for C1_MBT's HG, and 153.86 in variance and 12.34 in standard deviation for C1_FBS's HG. In terms of HNs, the mean leadtime of C1_MBT's HNs is -11.78 and C1_FBS is -18.10. Both of the mean figures are given in negative numbers, which means HNs are used in listener status. The variance and the standard deviation of HNs are also large numbers such as 96.25 in variance and 9.71 in standard deviation for C1_MBT, and 217.50 in variance and 14.61 in standard deviation for C1_FBS.

The result from this small pilot study cannot be generalised since the data is just a 10-minute conversation and qualitative aspects should be taken into consideration. However, a tendency can be observed that HG is used at the moment or soon after a speaker takes the floor, while a HN is often used in listener status.

In C2, however, it seems more difficult to have standardised figures about the use of HGs and HNs. As shown in Table A.1.2-4 in the appendix, the mean of leadtime for C2_FBT's HG is 58.65, which may not be significant since the variance is more than 1000 and the standard deviation is also a large number. The mean leadtime for C2_MJS'HG is 2, however, which also cannot be accurate since the

number of samples is only one. Furthermore, C2_FBT's HN has not been observed, hence no data is obtained. The mean leadtime for C2_MJS's HN is -65.54. However, this also cannot be a meaningful figure since the variance is more than 1000.

3.3.1.4 Findings about *erm, yeah* and *mm*

As for vocal response tokens, *erm, yeah* and *mm* were counted and summarised. Table A.1.3-1 in the appendix shows how many times these three items are used in C1, the British-British conversation, in each five-second leadtime interval, which is the same scale as that used in the analysis on conversational gestures.

In summary, C1_MBT uses *erm* 15 times while C1_FBS uses *erm* 28 times, and it can be noted that the use of *erm* by C1_FBS increases at the floor boundary. C1_MBT uses *yeah* 43 times, which is more than C1_FBS uses, and C1_MBT's use of *yeah* is observed often at the interval 'less than 0' where he attempts to take the floor. The tendency can be seen in C1_FBS since C1_FBS's use of *yeah* is also slightly increased at leadtime 'less than 0' and 'more than 0'. These findings imply that the use of *erm* and *yeah* might be related not only to their individual preference or habit of using these items but also to their social role, such as a tutor and a student, and gender.

Both of the participants in C1 rarely use *mm*. C1_MBT uses *mm* once at listener status while C1_FBS uses *mm* twice at speaker status, which is surprising since we assume *mm* as a response token is used only in listener status. In order to have more insight on what is going on with C1_FBS's use of *mm*, I would like to examine the details of the transcription for these two cases. The first case of C1_FBS occurs at 310 seconds as shown in Transcript 3.3-2 below, which is after a small pause and soon before C1_MBT takes the floor.

Transcript 3.3-2 *mm* in C1_FBS (1)

Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
301	FBS_F	-10	0			HN	Yeah I mean I was I was going to have to touch on intentionality behind metaphor anyway+
302		-9	1				
303		-8	2			HG	
304		-7	3			HG	
305		-6	4	HN		HG	
306		-5	5	SC/neck	Yeah yeah.		+whichever data I use.
307		-4	6			HG	
308		-3	7		Yeah.		
309		-2	8				
310		-1	9				<$E> pause <\$E> Mm.
311	MBT_F	0	-3		So have you looked at this stuff? I mean is there+		
312		1	-2			SC/nose	Yeah I+

Keys: C1_FBS= conversation 1 (British-British conversation) female British Student,
MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_gesture= male British tutor's gestures, FBS_gestures= female British student's gestures,
C1_ MBT_transcript= Conversation 1 (British-British conversation) male British tutor's transcription,
C1_FBS_transcript= Conversation 1 (British-British conversation) female British student's transcription,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
HN= head nods, SC/neck self comfort with neck, HG= hand gestures, SC/nose= self comfort with nose,
"+"= describe the continuous of the sentence, <$G?>= inaudible sounds,
<$E>pause<$E> = silence in conversation

Transcript 3.3-3 *mm* in C1_FBS (2)

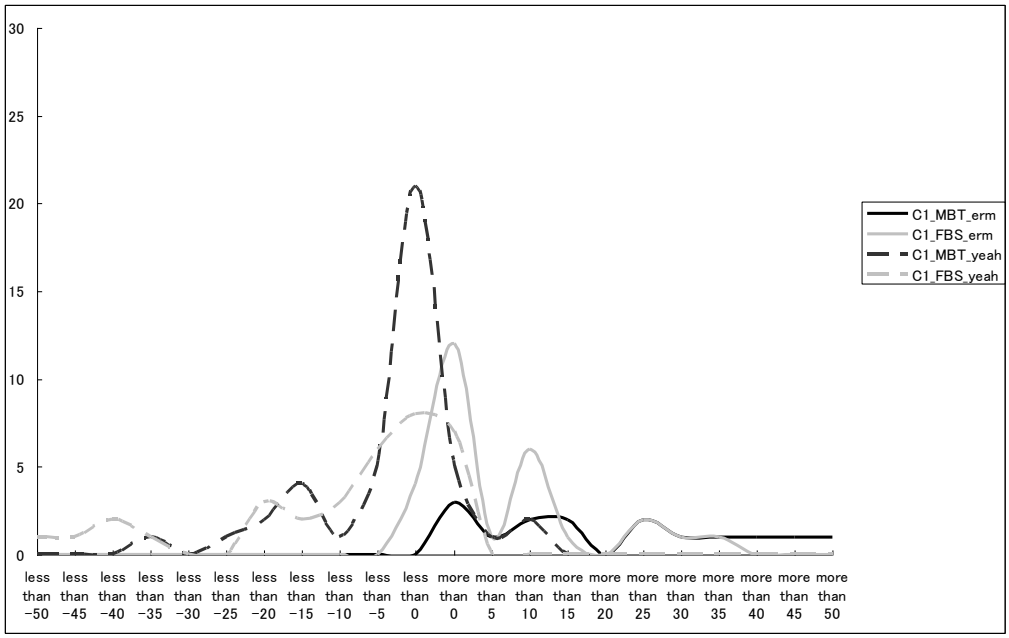
Timeline	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_gesture	C1_MBT_Transcript	C1_FBS_gesture	C1_FBS_Transcript
360	FBS_F	-35	0				Yeah well that's what cos all the conceptual metaphor theories talk about erm kind of this embodied experience thing and how you use your embodied experience+
361		-34	1				
362		-33	2			HG	
363		-32	3	HN			
364		-31	4				
365		-30	5			HG	
366		-29	6	HN			
367		-28	7	HN			
368		-27	8			HG	
369		-26	9		Yeah.		+to understand more abstract concepts
370		-25	10	HN		HG	
371		-24	11		Yeah.		but erm well this particular one that he's explaining here erm the chronic pain disorders which don't really have a medical explanation but are still embodied experience+
372		-23	12				
373		-22	13			HG	
374		-21	14				
375		-20	15			HG	
376		-19	16	HN			
377		-18	17			HG	
378		-17	18				
379		-16	19	HN		HG	
380		-15	20		Right.		
381		-14	21			HG	+but he's explaining them through I dunno perhaps more abstract things at times.
382		-13	22				
383		-12	23				
384		-11	24		Oh really? Oh right.		
385		-10	25				Mm. Well especially when er patients talk about it in terms of well the military metaphor erm+
386		-9	26	SC/arm			

Keys: C1_FBS= conversation 1 (British-British conversation) female British Student,
MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_gesture= male British tutor's gestures, FBS_gestures= female British student's gestures,
C1_ MBT_transcript= Conversation 1 (British-British conversation) male British tutor's transcription,
C1_FBS_transcript= Conversation 1 (British-British conversation) female British student's transcription,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
HN= head nods, SC/arm self comfort with arm, HG= hand gestures,"+"= describe the continuous of the sentence

The second case of C1_FBS's *mm* occurs at 385 seconds and 10 seconds before C1_MBT takes the floor as shown in Transcript 3.3-3 above. Prior to C1_FBS's *mm* at 385 seconds, C1_MBT gives a response token 'Oh really? oh right.', which I do not perceive as a speaker turn since the utterance is more like two two-word response tokens to the speaker and C1_FBS keeps the floor continuously.

C1_FBS's second *mm* can be taken as a response to her partner's comment or saying it to herself and taking time to think about what she is going to say next like 'let me see'. I will consider this issue more in Section 3.3.2 on qualitative analysis with the original video-recorded data.

A graph for the use of *erm* and *yeah* in C1 is illustrated in Figure 3.3-4 below. The broken black line, which is C1_MBT's *yeah*, increases soon before floor-taking.



Keys: C1= conversation 1 (British-British conversation),
y axis = the numbers of erm and yeah, x axis = ranges of leadtime (seconds),
C1_MBT_erm = C1 male British tutor's erm, C1_FBS_erm = C1 female British student's erm,
C1_MBT_yeah = C1 male British tutor's yeah, C1_FBS_yeah = C1 female British student's yeah

Figure 3.3-4 Number of *erm* and *yeah* in C1

The grey broken line, which is C1_FBS's *yeah*, also increases although the curve is rather moderate. Whereas the grey solid line, which is C1_FBS's *erm*, has a peak around the floor boundary while C1_MBT uses a few *erm* constantly even until 50 seconds after he takes the floor, which is different from C1_FBS.

The numbers of *erm, yeah* and *mm* in C2, the British-Japanese conversation, were also counted and are described in Table A.1.3-2 in the appendix. C2_FBT uses *erm* 10 times both in listener and speaker status, which is less than those of participants in C1. If compared with C1_FBS, this total is about a third of C1_FBS's use of *erm*. C2_MJS, however, has not used *erm* at all. In addition, *yeah* has been observed in both participants utterances, which is not a large amount either. Although C2_FBT has used *mm* only four times, C2_MJS's use of *mm* is outstanding, which occurs 33 times very frequently in the 10-minute conversation while he is in listener status.

From the table, two graphs are illustrated; one is for the use of *erm*, and *yeah*, another is for the use *yeah* and *mm* in C2. Figure A.1-2 in the appendix shows the numbers of *erm* and *yeah* in C2 visually. The black and grey broken lines, which represent the numbers of *yeah* in C2_FBT and C2_MJS, increase at floor boundaries; in other words, the interval 'more than 0', although these are not a large number. The black solid line, which is C2_FBT's *erm,* is also a small amount. However, they are seen in both negative and positive areas. Figure A.1-3 in the appendix shows the number of *yeah* and *mm* in C2. C2_MJS's use of *mm* is expressed in the grey broken line, which is spread out and covers the whole negative area. This means that he uses *mm* all the time when he is in listener status. It seems that there is not a close relationship between the timing of the floor-taking and his use of *mm.*

Means, variances, standard deviations of these three verbal response tokens, *erm*, *yeah* and *mm*, in C1 and C2 were also examined. The mean leadtime for C1_MBT's *erm* is 22.27 and C2_FBS's *erm* is 7.43 as shown in Table A.1.3-3 in the appendix. These figures, however, cannot be meaningful since the variances are large, such as 287.35 for C1_MBT and 123.37 for C1_FBS.

The same can be said in the use of *yeah.* The mean leadtime for C1_MBT's *yeah* is -6.35 and C1_FBS is -12.37 while the variance for C1_MBT is 146.42 and C1_FBS is 253.59. Although these figures might not be able to capture the features of these response tokens precisely, it can be said that *erm* tends to be used in speaker status while *yeah* tends to be used in listener status. Furthermore, in terms of *mm* in C1, there are rather small samples, such as one example for C1_MBT's *mm* and two for C2_FBS'*mm*. Hence, it seems difficult to figure out valid statistics from these data.

The same statistics data from C2 is shown in Table A.1.3-4 in the appendix, however, there are some differences in these figures from C1. The mean leadtime for C2_FBT is 41.80. Again the variance cannot be significant since the figure is more than 2000. There is no sample for C2_MJS's *erm*. The mean leadtime for C2_FBT's *yeah* is 2.00 and C2_FBT's *mm* is -8.50. The figures seem more relevant if compared with other items although the figures come from only three of four samples. Moreover, the comparison of the mean leadtime of *yeah* in C2_FBS with C2_MJS might be interpreted that *yeah* is often used in C2_FBS's speaker status and C2_MJS's listener status, since the mean of *yeah* in C2_FBS is a positive figure, while the mean of C2_MJS and the means of *yeah* in C1 are also negative figures.

The sample transcription in Transcript 3.3-4 below shows that C2_FBT uses *yeah* at 68 seconds in the utterance 'I think, yeah, this one, yeah', and this is the point

where she takes the floor:

Transcript 3.3-4 *yeah* in C2_FBT (1)

Timeline	Floor	FBT_leadtime	MJS_leadtime	C2_FBT_gesture	C2_FBT_Transcript	C2_MJS_gesture	C2_MJS_Transcript
68	FBT_F	0	-7		I think,yeah,this one,yeah,	SC/chin	
69		1	-6				
70		2	-5				
71		3	-4	HT	I haven't seen this one before,have I?		
72		4	-3				yeah

Keys: C2_FBT= conversation 2 (British-Japanese conversation) female British tutor,
FBT_leadtime= female British tutor's leadtime, MJS_leadtime= male Japanese student's leadtime,
C2_FBT_gesture= C2 female British tutor's gestures, C2_MJS_gestures= C2 male Japanese student's gestures,
C2_ FBT_transcript= C2 male British tutor's transcription, C2_MJS_transcript= C2 male Japanese student's transcription,
FBT_F= female British tutor floor-taking, HT= head turns, SC/chin = self comfort with chin,

Another example in Transcript 3.3-5 below also shows that C2_FBT uses *yeah* at 334 seconds in the timeline, which is 2 seconds after she takes the floor.

Transcript 3.3-5 *yeah* in C2_FBT (2)

Timeline	Floor	FBT_leadtime	MJS_leadtime	C2_FBT_gesture	C2_FBT_Transcript	C2_MJS_gesture	C2_MJS_Transcript
332	FBT_F	0	-134		somewhere but not in oh here		
333		1	-133				
334		2	-132		yeah		
335		3	-131				
336		4	-130		Wajnryb states observation as a part of raw material, which helps teachers professional grow. Therefore it is indispensable that teachers use observation as a tool for learning about teaching.		

Keys: C2_FBT= conversation 2 (British-Japanese conversation) female British tutor,
FBT_leadtime= female British tutor's leadtime, MJS_leadtime= male Japanese student's leadtime,
C2_FBT_gesture= C2 female British tutor's gestures, C2_MJS_gestures= C2 male Japanese student's gestures,
C2_ FBT_transcript= C2 male British tutor's transcription, C2_MJS_transcript= C2 male Japanese student's transcription,
FBT_F= female British tutor floor-taking,

In these particular cases, C2_FBT uses *yeah* not as a response token to the listener, but as a discourse marker or an information token after her holding the floor in order to express her understanding. The mean leadtime of C2_MJS's *yeah* is -16.57. However, the variance is more than 1000. That means the figure cannot be meaningful. In addition, the meantime of C2_MJS's *mm* is not significant either for the same reason.

3.3.1.5 Findings about pauses

The use of silent pauses in C1 and C2 will be reported in this section. Particularly in C2, the British-Japanese conversation, silent pauses were observed often, and the amount and placement of silent pauses seem to be related to turn-taking structure.

I describe pauses as negative numbers in leadtime as shown in Transcript 3.3-6 below. There is a long pause from 452 seconds to 466 seconds, where C2_MJS takes the floor. As C2_FBS's leadtime, on the one hand, I put the leadtime -17 at 453 seconds since the point the pause starts is 17 seconds before C2_FBT takes the floor, which includes both pauses and listener status.

Transcript 3.3-6 Leadtime with pauses in C2

Timeline	Floor	FBT_leadtime	MJS_leadtime	C2_FBT_gesture	C2_FBT_Transcript	C2_MJS_gesture	C2_MJS_Transcript
449		7	-17		ah sorry,		mm
450		8	-16		I=, I might still little bit,for this reason		
451		9	-15				
452	pause	-17	-14		<$E> pause <\$E>		
453		-16	-13				
454		-15	-12				
455		-14	-11				
456		-13	-10		mm		
457		-12	-9				
458		-11	-8		<$E> pause <\$E>		
459		-10	-7				
460		-9	-6				
461		-8	-5				
462		-7	-4				
463		-6	-3				
464		-5	-2				
465		-4	-1				
466	MJS_F	-3	0				un
467		-2	1				probably I mention two kinds of observation
468		-1	2			HG	
469	FBT_F	0	-90		I think, it shouldn't be, it shouldn't be marbled this		

Keys: C2 = conversation 2 (British-Japanese conversation),
FBT_leadtime= female British tutor's leadtime, MJS_leadtime= male Japanese student's leadtime,
C2_FBT_gesture= C2 female British tutor's gestures, C2_MJS_gestures= C2 male Japanese student's gestures,
C2_ FBT_transcript= C2 male British tutor's transcription, C2_MJS_transcript= C2 male Japanese student's transcription,
MJS_F= male Japanese student floor-taking, FBT_F= female British tutor floor-taking, HG= hand gestures
"="= unfinished sentence, <$E>pause<$E> = silence in conversation

As C2_MJS's leadtime, on the other hand, I put the leadtime -13 at 453 seconds, which is continued from the previous leadtime since C2_MJS has been in listener

status before the pause, and the point the pause starts is 14 seconds before C2_MJS takes the floor. During the silent pause, C2_FBT are reading the essay and C2_MJS is also looking down the paper.

By implementing the alignment of leadtime in this way, silent pauses between the end of the previous speaker's utterance to the next floor-taking are able to be described. However, this way of describing silent pauses has a drawback that negative leadtime is used to express both listener status and pauses, which could cause some confusion. Especially in a conversation including a significant amount of pauses like C2, the data needs to be analysed carefully since negative leadtime can be interpreted either as listener status or pauses according to the other participant's leadtime.

The use of silent pauses seems to significantly affect the frameworks and functions in conversation. In many cases the floor of the conversation was exchanged between participants, however, there are some places where pauses were involved in turn-taking management. Hence one factor we can easily remove because of the alignment of the data will be the transition orders with pauses. In Table 3.3.1-3 below, four cases where pauses are taken part in turn exchanges in C2 are listed.

Table 3.3.1-3 Transition cases with pauses in C2

Transaction cases	Count
1) FBT > pause > FBT	17
2) FBT > pause > MJS	1
3) MJS > pause > MJS	0
4) MJS > pause > FBT	0
Total	18

Keys: C2 = conversation 2 (British-Japanese conversation), MBT = female British tutor, MJS = male Japanese student

In the first case, C2_FBT takes the floor first, and after a pause C2_FBT takes back

the floor again. In case two C2_FBT takes the floor first, followed by a pause, and then C2_MJS takes the floor after the pause. Case three is a similar instance to case one but MJS takes the floor *before* and *after* a pause. Further, case four is similar to case three, where the different participant has taken the floor before and after a pause. C2_MJS has lost the floor after a pause and C2_FBT takes the floor in turn in case four. There are 18 pauses in C2 and 17 of them are categorised as the case one, and only one pause is categorised as case two. Alternatively, if we consider the fact that C2_FBT takes the floor 28 times and C2_MJS takes the floor only eight times, the feature that more than half of C2_FBT's floor-taking occurs after pauses could be understandable. Because of the limitation in the numbers of C2_MJS's speaker turns, most of C2_FBT's speaker turns follow not C2_MJS's speaker turns but silent pauses. One of the eight turn initiations in C1_MJS also occurs after a silent pause.

3.3.2 Salient findings from the turn structure analysis

Based on Drummond and Hopper's research methods (1993) reviewed in Chapter Two, an attempt was made to describe floor exchanges placing focus on listener response tokens in C1 and C2. Although I have analysed both response tokens in listener status and discourse markers in speaker status in the quantitative analysis, the focus was narrowed down into listener status in the qualitative analysis.

In the turn structure analysis, only verbal response tokens such as *yeah*, *right*, *erm* and *mm* were analysed in the first stage with two categorisations of functions of response tokens. The first categorisation of response tokens is defined by O'Keeffe and Adolphs in Knight et al (2006). There are four functions of response tokens. I shall refer to this categorisation as *discourse function*, namely continuers (CN), convergence tokens (CV), engaged response tokens (EN), and information receipt

tokens (IR). The other categorisation was developed based on Schegloff (1984), which expressed functions of response tokens and utterances in relation to floor exchange. I shall refer to this categorisation as a *conversation function*. There are also four functions in conversation function, namely floor-taking (FT), floor seeker (FS), listenership (LS) and floor giving (FG) (see Section 3.2.3.2). Conversation functions are also closely related to turn-taking structure, whereas discourse functions are used to describe listeners' intentions expressed by response tokens. There is an overlap between conversation functions and discourse functions. Conversation functions based on Knight et al. (2006) are instances of floor seekers and listenership in discourse function. The targeted response tokens were analysed with regard to these two levels in functions of response tokens, namely conversation function (CF) and discourse function (DF), in relation to turn-taking structure.

By using a combination of these two categorisations, namely discourse function and conversation function, listener response tokens were analysed. I extracted data of participants' listener status and examined floor-taking patterns with response tokens. As a result, five floor-taking patterns were observed in the conversation data C1 and C2 as shown in Table 3.3.2-1. Each case of listener status is categorised into these five turn-taking patterns:

Table 3.3.2-1 Floor-taking patterns

	Pattern A (SS)	Pattern B (SS)	Pattern C (SS)	Pattern D (OS)	Pattern E (SS)
LS	Continuers	-	Continuers	(Continuers)	-
FG	-	-	-	Partner's floor giving	-
FS	Information receipt tokens	Engaged tokens	Information receipt tokens/ Engaged tokens	Convergence tokens	Pause/ Discourse markers
LS	-	-	Continuers	-	-
FT	Floor taking	Floor taking	Floor taking	Floor taking	Floor taking

Keys: LS=Listenership, FG=Floor giving, FS=Floor seeker, FT=Floor-taking, SS=Self-selection, OS=Other-selection

The first column in the table above describes conversation function such as LS

for listenership, FG for floor giving, FS for floor seeker, and FT for floor-taking. In the second column to the sixth column, descriptions of the discourse function at each conversation function phase are shown. Speaker selection type is also indicated with SS (self-selection) or OS (other-selection) on the first row.

Pattern A consists of continuers (CN) as listenership (LS), and information receipt tokens (IR) as floor seeker (FS), which is often observed in C1_MBT's listenership behaviour. Transcript 3.3-7 below, for instance, shows a typical example of pattern A, which is quoted from C1_MBT's listener status in C1. The first column of the transcription describes the timeline; the participants' TTPs are indicated in the second column; the third column is C1_MBT's leadtime:

Transcript 3.3-7 C1_MBT listenership behaviour (1) Pattern A

Timeline	Floor	MBT_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
171	FBS_F	-15	LS	CN		Right.
172		-14				
173		-13				
174		-12				
175		-11				
176		-10	LS	CN	HN	
177		-9	LS	CN		Yeah.
178		-8				
179		-7				
180		-6	LS	CN	HN	
181		-5				
182		-4	LS	CN	HN	
183		-3	FS	IR		That's right.
184		-2	FS	IR		Yeah yeah.
185		-1				
186	MBT_F	0	FT		HG	Well it's shared but+

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
FBS_F=female British student floor-taking, MBT_F=male British tutor floor-taking
LS=listenership, FS=floor seeker, FT=floor-taking, HN=head nods, HG=hand gestures

Conversation function and discourse function of response tokens that

C1_MBT has used are described in the fourth and the fifth columns. C1_MBT's gestures are indicated in the sixth column. The last column shows C1_MBT's utterances, which is labelled as C1_MBT_Transcript.

In C1_MBT_transcript column, a freestanding *right* is observed at -15 in C1_MBT's leadtime. There is another freestanding *yeah* at -9 in C1_MBT's leadtime, which are items for showing listenership in conversation function and continuers in discourse function. At 3 to 2 seconds before his taking the floor, C1_MBT has non-minimal response tokens, such as *That's right* and *yeah yeah*, which can be interpreted as floor seekers in conversation function and information receipt tokens in discourse function since he does take the floor within 5 seconds after using these response tokens. He then takes the floor with an utterance beginning with the discourse marker *well*.

In Pattern B, on the one hand, engagement tokens (EN) are used as floor seeker. This pattern is often observed in relatively short listener status. A typical case of Pattern B can be seen in Transcript 3.3-8 below:

Transcript 3.3-8 C1_FBS listenership behaviour (1) Pattern B

Timeline	Floor	FBS_leadtime	C1_FBS_CF	C1_FBS_DF	C1_FBS_gesture	C1_FBS_Transcript
110		-7				
111		-6				
112		-5				
113		-4	FS	EN		No.
114		-3	FS		HS/HG	
115		-2	FS	EN	HS	No.
116		-1				
117	FBS_F	0	FT		HG	There have been some linguistic studies of them but no+

Keys: C1_FBS=Conversation 1 (British-British conversation) female British student,
C1_FBS_CF=C1_FBS conversation function, C1_FBS_DF=C1_FBS discourse function,
FBS_F=female British student floor-taking, FS=floor seeker, FT=floor-taking, EN=engaged tokens,
HS=head shakes, HG=hand gestures, HS/HG=head shakes and hand gestures

In this example, she uses *no* at 4 and 2 seconds before taking the floor and starts her

floor without any discourse marker.

Pattern C, on the other hand, often occurs in the case of a longer listener status, which starts with continuers as listenership:

Transcript 3.3-9 C1_FBS listenership behaviour (2) Pattern C

Time line	Floor	MBT_leadtime	FBS_leadtime	C1_MBT_Transcript	C1_FBS_CF	C1_FBS_DF	C1_FBS_gesture	C1_FBS_Transcript
276		34	-25	+even where it's not made explicit.	LS	CN	HN	Yeah.
277		35	-24					
278		36	-23					
279		37	-22	Erm so you know they don't you know it's this is all about the text+	LS	CN	HN	Right.
280		38	-21					
281		39	-20					
282		40	-19	+this is all about reception but it's also all about choice+	FS	EN	HN	Yeah uh-huh
283		41	-18					
284		42	-17					
285		43	-16					
286	FBS_F_failed	44	-15	+well all that really means is authorial choice. Yeah that's right. So like Andrew Goatly's stuff on the language of metaphor+	FS	IR/EN	HN/HG	Yeah I mean you can't not suggest+
287		45	-14					
288		46	-13					
289		47	-12					
290		48	-11					
291		49	-10	+makes it look very structuralist but actually he's talking about why an author chose x and not y.	LS	CN	HN/HG	Yeah.
292		50	-9					
293		51	-8					
294		52	-7					
295		53	-6					
296		54	-5		LS	CN	HN	Okay.
297		55	-4	So I wouldn't be too worried about that cos it's implicit in most of the metaphor stuff anyway.				
298		56	-3					
299		57	-2					
300		58	-1					
301	FBS_F	-10	0		FT		HN	Yeah I mean I was I was going to have to touch on intentionality behind metaphor anyway+

Keys: C1_FBS= conversation 1 (British-British conversation) female British Student,
MBT_leadtime= male British tutor's leadtime, FBS_leadtime= female British student's leadtime,
MBT_gesture= male British tutor's gestures, FBS_gestures= female British student's gestures,
C1_ MBT_transcript= Conversation 1 (British-British conversation) male British tutor's transcription,
C1_FBS_transcript= Conversation 1 (British-British conversation) female British student's transcription,
MBT_F= male British tutor floor-taking, FBS_F= female British student floor-taking,
HN= head nods, SC/arm self comfort with arm, HG= hand gestures,"+"= describe the continuous of the sentence,

In some occurrences, information receipt tokens or engaged response tokens as

floor seeker were observed in the midst of listener status. However the listener was going back to listenership without taking the floor until the next TTP. Transcript 3.3-9 above is an example of Pattern C.

In the transcription, C1_FBS is in listener status for about 60 seconds. A freestanding *yeah* is shown at -25 in C1_FBS's leadtime and another freestanding *right* is at -21 in C1_FBS's leadtime. At 19 seconds before she takes the floor, she produces a minimal response with *yeah.* At 15 seconds before her taking the floor, she attempts to take the floor with *yeah* and the beginning of the utterance, such as 'yeah, I mean you can't not suggest', which can be interpreted as a floor seeker in a conversation function, and an information receipt token, or engaged response token acting as a discourse function. However, the utterance is not completed and she goes back to listener status and showing continuers, such as *yeah* at -10 seconds and *okay* at -5 seconds in leadtime. Then without using response tokens as engaged response tokens, she is getting into her speaker status by catching the next more accessible TRP, where C1_MBT has also given away the speaker status to C1_FBS naturally. Hence, response tokens for floor seeker cannot be observed soon before the floor-taking in this occurrence.

Pattern D was observed only in other-selection (OS). As reviewed in Chapter Two, there are two types of speaker selections: *self-selection* (SS) and *other-selection* (OS). In Pattern D, the listener is given the floor by the previous speaker. Hence this pattern includes the previous speaker's floor giving (FG). Convergence tokens (CV), which mark agreement or disagreement with the question or comments, follow as a floor seeker.

As shown in Transcript 3.3-10 below, C1_MBT gives the floor to C1_FBS by

asking a question ‘So this is what? This is doctors trying to explain technicalities’ at timeline 340. There is no listenership before floor seeker in this sample transcription below since C1_FBS has a rather short listener status this time:

Transcript 3.3-10 C1_FBS listenership behaviour (3) Pattern D

Time line	Floor	FBS_leadtime	C1_MBT_Transcript	C1_FBS_CF	C1_FBS_DF	C1_FBS_gesture	C1_FBS_Transcript
340	MBT_F	-4	So this is what? This is doctors trying to explain+				
341		-3					
342		-2		FS	CV		Yeah erm.
343		-1	+technicalities.				
344	FBS_F	0	Right.	FT		HN/HG	Yeah.So I suppose as they would in a+

Keys: C1_FBS=Conversation 1 (British-British conversation) female British student,
C1_FBS_CF=C1_FBS conversation function, C1_FBS_DF=C1_FBS discourse function,
MBT_F=male British tutor taking the floor, FBS_F=female British student floor-taking
FS=floor seeker, FT=floor-taking, CV=convergence tokens,
HS/HG=head shakes and hand gestures

At 2 seconds before taking the floor, she produces *Yeah erm*, which can be interpreted as a floor seeker within the conversation function and a convergence token with hesitation within the discourse function. C1_FBS then takes the floor with the discourse marker *so*.

The last pattern was only observed in C2, the British-Japanese conversation, which has a number of pauses. In this pattern, a listener can take the floor after pauses. The typical pattern can be seen in Transcript 3.3-11 below. After the last utterance of C2_FBT’s previous floor at time 368, there is a long pause from time 373 to 395. During this pause, C2_FBT utters only one freestanding *erm* while she has been reading through C2_MJS’s assignment essay, and C2_MJS is silent and looking down at his writing with C2_FBT. In FBT_leadtime and MJS_leadtime where the discourse marker *erm* occurs, it can be noticed that both of them are negative numbers, which

means the participants are sharing the pause in silence. At 396 in leadtime in the transcript above, C2_FBT takes the floor without any discourse marker this time:

Transcript 3.3-11 C2_FBT listenership behaviour (1) Pattern E

Timeline	Floor	Pattern	FBT_leadtime	MJS_leadtime	C2_FBT_CF	C2_FBT_DF	C2_FBT_gesture	C2_FBT_Transcript
368			36	-98				I think this should start, start with this sentence
369			37	-97				
370			38	-96				
371			39	-95				
372			40	-94				
373	pause		-23	-93				<$E> pause <\$E>
374			-22	-92				
375			-21	-91				
376			-20	-90				
377			-19	-89				
378			-18	-88				
379			-17	-87				
380			-16	-86				
381			-15	-85				
382			-14	-84				
383			-13	-83				
384			-12	-82				
385			-11	-81	FS			erm
386			-10	-80				
387			-9	-79				
388			-8	-78				
389			-7	-77				
390			-6	-76				
391			-5	-75				
392			-4	-74				
393			-3	-73				
394			-2	-72				
395			-1	-71				
396	FBT_F	E	0	-70	FT			it is difficult for teachers to judge whether the way, the way they are teaching+

Keys: C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_FBT_CF=C2_FBT conversation function, C2_FBT_DF=C2_FBT discourse function,
FBT_F=female British tutor floor-taking, FS=floor seeker, FT=floor-taking,

On occasion, the next speaker shows intention to take a speaker turn with discourse markers, such as *erm* and *mm*. Although these items might not be referred to as response tokens since these utterances are made during pauses, I include this pattern as one of the floor-taking patterns. These discourse markers during pauses can be

identified as floor seekers in conversation function.

I examined each participant's listenership behaviour and attempted to match them to these five patterns in the pilot study:

Table 3.3.2-2 Participants' floor-taking patterns

	C1_MBT	C1_FBS	C2_FBT	C2_MJS
Pattern A	●	--	●	--
Pattern B	●	●	--	--
Pattern C	●	●	--	●
Pattern D	--	●	●	●
Pattern E	--	--	●	--
Unclassified	●	--	●	--

Keys: C1_MBT= Conversation 1 (NS-NS conversation) male British tutor,
C1_FBS = Conversation 1 (NS-NS conversation) female British students,
C2_FBT= Conversation 2 (NS-NNS conversation) female British tutor,
C2_MJS= Conversation 2 (NS-NNS conversation) male Japanese student.

In Table 3.3.2-2 above, patterns observed in each participant's listener status in C1 and C2 are summarised. The first column expresses floor-taking patterns and the first row shows the participants' name, such as C1_MBT for Conversation 1 (British-British conversation) male British tutor.

The black circles indicate whether the participant has the pattern or not. For instance, C1_MBT uses Patterns A, B, C and unclassified. A pattern which cannot be categorised in any of these five patterns from A to E is considered unclassified, and C1_FBS does not use any Pattern A and Pattern E. In C2 alternately, the British-Japanese conversation, C2_FBT uses Pattern A, D, E and unclassified, and C2_MJS uses Pattern C and Pattern D. The details of examination will be described in the following sections. At a later stage of the qualitative analysis, collocations of visual response tokens, such as head nods (HNs) and hand gestures (HGs), with verbal response tokens were also examined.

3.3.2.1 C1_MBT's use of response tokens

The pattern which is most frequently used in C1_MBT is Pattern A. In Pattern A, several types of freestanding response tokens, such as *yeah, right* and *uh-huh*, are used more than 5 seconds before taking the floor. These response tokens can be interpreted as listenership, which means that the participant shows acknowledgement and maintains his listener status. Between -5 and -1 seconds in leadtime, there is a tendency where the listener produces a minimal response with *right*, which can be interpreted as floor seeker in conversation function and information receipt tokens in discourse. The very moment when the listener becomes a speaker, he often uses discourse markers, such as *yeah, well* or *so*, at the beginning of his utterances. This latter pattern is summarised in Table 3.3.2-3 below:

Table 3.3.2-3 C1_MBT's floor-taking pattern: Pattern A (SS)

	Leadtime	Verbal response tokens	Functions
LS (Listenership)	x < approx. -5	freestanding *yeah* or *right* or *aha*	Continuers
FS (Floor seeker)	approx .-5 < x	non-minimal response or clusters with *right*	Information receipt tokens
FT (Floor-taking)	0 = x	yeah or well or so or but + full turn	-

Keys: SS=Self-selection, C1_MBT=Conversation 1 (British-British) male British tutor

In the table above, the first column shows conversation function adapted from Schegloff's analysis on floor-taking including listenership (LS), floor seeker (FS) and floor-taking (FT), which are aligned as time order from the top to the bottom. In C1_MBT's Pattern A, for example, the freestanding occurrences of *yeah, right* or *aha* are uttered more than 5 seconds before floor-taking. Non-minimal response token or clusters with *right*, such as *that's right* are then uttered less than 5 seconds before

floor-taking.

At TTP, C1_MBT takes the floor of the conversation with discourse markers, such as *yeah* and *well*. The second column shows leadtime when these particular statuses are observed. The third column describes occurrences of verbal response tokens in each category, and discourse function of these response tokens are indicated in the last column. The term *full turn*, which is shown in the third column in the bottom row in the figure above, is used to describe an utterance which has over three words and enables a listener to become a speaker. A typical example of C1_MBT's pattern A was described in Transcript 3.3-7 above.

There are several variations in this pattern. He sometimes uses non-minimal response tokens or the freestanding *right* when he is seeking the floor of the conversation. In Transcript 3.3-12 below, for instance, C1_MBT uses a freestanding *right* one second before his taking the floor. This response token can be interpreted as floor seeker although the instance is different from the typical floor seeker in Patten A:

Transcript 3.3-12 C1_MBT listenership behaviour (2)

Timeline	Floor	MBT_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
468		-1	FS	IR		Right.
469	MBT_F	0	FT			So this is fighting your illness as an+

Keys: C1_MBT=Conversation 1 (NS-NS) male British tutor, MBT_F=male British tutor floor-taking
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
FS=floor seeker, FT=floor-taking, IR=information receipt tokens

Floor seekers in pattern A tend to be non-minimal responses with *right*. The pattern of the floor-taking in this transcription follows the typical pattern of his listenership

behaviour, where he takes the floor with the discourse marker *so*.

The other version of Pattern A can also be observed where C1_MBT uses non-minimal responses with *yeah* instead of *right*:

Transcript 3.3-13 C1_MBT listenership behaviour (3)

Timeline	Floor	MBT_leadtime	C1_MBT_FA	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
508		-5	FS	IR	HN	
509		-4	FS	IR		Yeah yeah.Yeah yeah.
510		-3	FS	IR		Yeah.
511		-2				
512		-1	FS	IR		Yeah.
513	MBT_F	0	FT		HG	So are you interested in the conceptual content or the sort of stylistic realisation?

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor, MBT_F=male British tutor floor-taking
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
FS=floor seeker, FT=floor-taking, IR=information receipt tokens, HN=head nods, HG=hand gestures

In Transcript 3.3-13, four seconds before his taking the floor, C1_MBT uses *yeah* several times, which can be interpreted as floor seekers within the conversation function and information receipt tokens within the discourse function. These response tokens lead a next floor-taking with the discourse marker *so,* which is a typical feature of the floor-taking in Pattern A.

There is also a variant of Pattern A, I shall refer it as Pattern A+, where C1_MBT does not show any listenership before floor seeker because of a short listener status.

Transcript 3.3-14 C1_MBT listenership behaviour (4)

Timeline	Floor	MBT_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
118		-3				
119		-2	FS	IR	HN	Right. Okay.
120		-1				
121	MBT_F	0	FT			Ah you might be alright then you see the only worry is if if you're dealing with data that's been set up in an experimental situation by a linguist+

Keys: C1_MBT=Conversation 1 (NS-NS) male British tutor,
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
FS=floor seeker, FT=floor-taking, IR=information receipt tokens, HN=head nods

In Transcript 3.3-14 above, C1_MBT is in listener status for 3 seconds. Although C1_MBT does not show any listenership, he still uses a non-minimal response, *right okay*. This can be interpreted as a floor seeker in the conversation function and information receipt tokens within the discourse. Thus, this occurrence can be categorised as Pattern A+.

3.3.2.2 C1_FBS's use of response tokens

During the course of analysis on C1_FBS's floor exchanges, the concept of speaker selection was considered. There are two types of speaker selection: *self-selection* and *other-selection* respectively (Sacks 1992). If a speaker asks a question to a particular listener, and the listener who has been nominated answers the question, this is other-selection. Alternatively, when a speaker asks some questions to more than one listener without pointing out a particular listener, everyone in the conversation can self-select as the next speaker, and this is referred as self-selection.

In the 10-minutes conversation data of C1, C1_MBT's floor-taking is always led only by self-selection. However, in FBS's taking the floor, both types of speaker selections can be observed. 6 of 19 in FBS's taking the floor are self-selection (SS) and the other 13 are other-selection (OS) as shown in Table 3.3.2-4 below:

Table 3.3.2-4 Self-selection and other-selection in C1

	Floor-taking
C1_MBT	20 (SS: OS = 20: 0)
C1_FBS	19 (SS: OS = 6: 13)

Keys: C1=Conversation 1, C1_MBT=Conversation 1 (British-British conversation) male British tutor, C1_FBS=Conversation 1 (British-British conversation) female British student, SS=self-selection, OS=other-selection

FBS's floor-taking patterns were analysed in relation to speaker selections. There are two types of floor exchanges in the case of SS in FBS's taking the floor. When C1_FBS is in a relatively short listener status, she uses a freestanding *yeah* or *no* as a floor seeker within the conversation function, within 5 seconds before taking the floor. This can be interpreted as EN within the discourse since these response tokens are followed by the floor-taking. Further, when she has taken the floor, discourse markers, such as *and* or *well* are observed:

Table 3.3.2-5 C1_FBS's floor-taking pattern: Pattern B (SS)

	Leadtime	Verbal response tokens	Functions
FS (Floor seeker)	approx .-5 < x	freestanding with *yeah* or *no*	Engaged tokens
FT (Floor-taking)	0 = x	*and* or *well* or no discourse marker + full turn	-

Keys: C1_FBS=Conversation 1 (British-British conversation) female British student, SS=Self-selection,

I have summarised this first instance of C1_FBS's listenership behaviour in Table 3.3.2-5 and referred to this as Pattern B (SS). The typical case of Pattern B was described in Transcript 3.3-8.

The second instance in the case of SS is observed when C1_FBS has a longer listener status. While she is in listener status, she uses response tokens such as freestanding *yeah*, *right* or *okay*, which can be interpreted as listenership within the conversation function and continuers within the discourse function. In some cases, she seems to seek for the floor in the middle of her listener status by starting an utterance

with *yeah*. However, she fails to take the floor and goes back to listenership. Then she maintains listenership until the next TRP. I summarise this instance in Table 3.3.2-6 below as Pattern C (SS):

Table 3.3.2-6 C1_FBS's floor-taking pattern: Pattern C (SS)

	Leadtime	Verbal response tokens	Functions
LS (Listenership)	-	freestanding *yeah* or *right* or *okay*	Continuers
FS (Floor seeker)	-	Yeah + some words	Information receipt tokens Engaged tokens
LS (Listenership)	-	freestanding *yeah* or *right* or *okay*	Continuers
FT (Floor-taking)	0 = x	Yeah or and + full turn	-

Keys: C1_FBS=Conversation 1 (NS-NS) female British student, SS=Self-selection,

In this pattern, the timing when the floor seekers occur and how long her listenership is maintained vary from case to case. Hence any particular leadtime is not indicated in this figure apart from the TTP, which is defined as leadtime 0. When C1_FBS is seeking the floor during her listener status, *yeah* is used with some words, which can be interpreted as information receipt tokens or engaged response tokens within the discourse function according to the situation. A longer example is shown in Transcript 3.3-9 above.

The last feature observed in C1_FBS was a case of other-selection (OS). Floor giving (FG) is a distinctive feature seen in OS. As shown in Transcript 3.3-10 above, C1_MBT gives the floor to C1_FBS by asking a question. This instance is referred to as pattern D and summarised in Table 3.3.2-7 as follows:

Table 3.3.2-7 C1_FBS's floor-taking pattern: Pattern D (OS)

	Leadtime	Verbal response tokens	Functions
LS (Listenership)	-	freestanding *right* or *okay* or *yeah*	Continuers
FG (Floor giving) from the partner	approx. -1 > x	questions or comments from the participant	
FS (Floor seeker)	-1 < x or 0 = x	freestanding *yeah* or *no*	convergence tokens
FT (Floor-taking)	0 = x	yeah or no or well + full turn , some times	-

Keys: C1_FBS=Conversation 1 (British-British conversation) female British student, OS=other-selection,

Concerning the leadtime in the table above, 'less than -1' is indicated in FG since in most cases the previous speaker is giving a question a few seconds before the next speaker's floor-taking. The next speaker answers the question with *yes* or *no* a second before s/he starts their utterance and takes the floor. Pattern D is the only case of OS among the five floor-taking patterns. All of the other four patterns are cases of SS.

3.3.2.3 C2_FBT's use of response tokens

As stated above, C2 has a female British tutor (FBT) and a male Japanese student (MJS) as participants. There were some differences in features in their listenership behaviours from C1, the British-British conversation. It is important to note that it is more difficult to analyse the C2 data if compared with C1 since more pauses and unequal contribution between participants are involved in C2. Most of the time in the 10-minute C2 conversation data, C2_FBT was in a speaker status and C2_MJS was in listener status, or both shared pauses in conversation. Even when C2_MJS took the floor, his speakership did not last long, which caused the situation where C2_MBT was rarely in listener status. This fact itself might be recognised as one of the features of the British-Japanese conversation in this specific case.

Both of the participants in C2 had two types of floor-taking: self-selection (SS) and other-selection (OS). Most of C2_FBT's floor-taking was led by herself and only one case out of 25 did she use OS. In contrast, most of C1_MJS's floor-taking

was OS. In the conversation, C1_MJS asked questions or received some comments which required his responses, and only 1 out of 7 cases of floor-taking was led by himself as shown in Table 3.3.2-8 below.

Table 3.3.2-8 Self-selection and other-selection in C2

	Floor-taking
C2_FBT	25 (SS: OS = 24: 1)
C2_MJS	8 (SS: OS = 1 : 7)

Keys: C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
SS=self-selection, OS=other-selection,

There is a distinctive instance, which was observed in C2_FBT's floor-taking and led by SS. This instance is different from the other four patterns we reviewed in the previous section. In this instance, which I name as Pattern E, C2_FBT's taking of the floor comes directly after pauses, and during the previous pause she sometimes utters freestanding *erm* or *mm* to show her intention. These items can be interpreted as discourse markers for acknowledgement or change of state token from non-known to now known as described in the study on *oh* by Heritage (1984a) and Schiffrin (1987) (see chapter 2).

In addition, in some cases of silent pauses, C2_FBT looks at C2_MJS's essay and takes time to understand what he has written and conceptualise what she is going to say next. Silent pauses here are differentiated from pauses between speakers since C1_FBT is engaged in the task of reading his essay. Hence I recognise this *erm* and *mm* as not response tokens but discourse markers, which are related to speakership more than listenership. C2_FBT utters these response tokens in order to show C2_MJS that she understands the essay, whereas C2_FBT takes the floor with

discourse markers, such as *okay* or *yeah*, in about half of the cases. In the other half of the cases, C2_FBT's taking the floor occurs without these discourse markers. I have summarised this pattern in Table 3.3.2-9 below:

Table 3.3.2-9 C2_FBT's floor-taking pattern: Pattern E (SS)

	Leadtime	Verbal response tokens	Functions
Pause or FS (Floor seeker)	-	Pause, or pause with freestanding *erm* or *mm*	Pause or discourse marker
FT (Floor-taking)	0 = x	*okay* or *yeah* or no discourse marker+ full turn	-

Keys: C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor, SS=self-selection,

There are no listenerships seen in this pattern. Pauses or some discourse markers are used for floor seeking. Although this pattern does not describe C2_FBS's listenership behaviour since she has not been in listener status before the floor-taking, it does describe a pattern observed in relation to turn-taking, and it can be a fruitful finding in qualitative analysis that the freestanding *erm* and *mm* observed in this pattern are not listener response tokens but discourse markers – although the leadtime where these discourse markers occur is described as negative numbers because of the pauses. It is not possible to derive these conclusions from the previous quantitative analysis. The typical pattern was described in Transcript 3.3-11 above.

Another occurrence observed in C2_FBS is a variant of Pattern A as shown in Transcript 3.3-15 below:

Transcript 3.3-15 C2_FBT listenership behaviour (2)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtime	MJS_leadtime	C2_FBT_CF	C2_FBT_DF	C2_FBT_Transcript	C2_MJS_CF	C2_MJS_DF	C2_MJS_Transcript
38				1	-1			Have you got a brief			
39	MJS_F	D	OS	-2	0				FT		ah, well, ah,I don't have
40				-1	1						
41	FBT_F	A+	SS	0	-34	FT		you haven't got,			

Keys: C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_FBT_CF=C2_FBT conversation function, C2_FBT_DF=C2_FBT discourse function,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_FBT discourse function,
MJS_F=male Japanese student taking the floor, FBT_F=female British tutor floor-taking,
D=pattern D, A+= a variant of pattern A, OS=other-selection, SS=self-selection, FT=floor-taking,

Occurrences of Pattern A were observed often in C1_MBT's listenership behaviour, which include continuers as markers of listenership, and information receipt tokens for floor seeking. Because of C2_FBT's short listener status and her less frequent use of response tokens as continuers, the occurrence is different from the typical pattern shown in Pattern A. In the transcript above, C2_FBT asks about an assignment brief at time 38, C2_MJS answers the question at time 39. Then, C2_FBS takes the floor at time 41 without discourse markers, such as *right* or *okay*. Although there is no response token for continuers and information receipt tokens observed in this instance, I would like to categorise this instance as a variant of Pattern A since C2_FBT shows that she has received her partner's utterance when she takes the floor. By doing so, we can clearly see the differences in the use of response tokens and listenership behaviours in the same floor-taking patterns between C2_FBT and C1_MBT.

The other case of a variant of Pattern A in C2_FBT's floor exchanges is related to listenership in a unique way. There is a case where C2_MJS speaks out after a pause and C2_FBS responds to him immediately. At this point, the way C2_FBS treats her partner's utterance is interesting. In Transcript 3.3-16 below, both participants are in silence from time 468 to 464. C2_MJS leads his turn by himself at

466 in the timeline. Soon after he has completed the single sentence, however, C2_FBT takes back the floor and C2_MJS goes back to listener status. C2_FBT has not shown any response tokens before taking her turn; rather, she produces *yeah, yeah* to show her receiving C2_MJS's previous utterance at 472:

Transcript 3.3-16 C2_FBT listenership behaviour (3)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtime	MJS_leadtime	C2_FBT_CF	C2_FBT_DF	C2_FBT_gesture	C2_FBT_Transcript	C2_MJS_CF	C2_MJS_DF	C2_MJS_gesture	C2_MJS_Transcript
458		C+		-11	-8				<$E> pause <\$E>				
459		C+		-10	-7								
460		C+		-9	-6								
461		C+		-8	-5								
462		C+		-7	-4								
463		C+		-6	-3								
464		C+		-5	-2								
465		C+		-4	-1					FS	EN		un
466	MJS_F	C+	SS	-3	0					FT		HG	probably I mention two kinds of observation
467		A+		-2	1								
468		A+		-1	2								
469	FBT_F	A+	SS	0	-90	FT			I think, it shouldn't be, it shouldn't be marbled this				
470		E		2	-89								
471		E		3	-88	IR?		HT					
472		E		4	-87	IR?			yeah, yeah				

Keys: C2_FBT=Conversation 2 (NS-NNS) female British tutor,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_FBT_CF=C2_FBT conversation function, C2_FBT_DF=C2_FBT discourse function,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_MJS discourse function,
MJS_F= male Japanese student floor-taking, FBT_F=female British tutor floor-taking,
SS=self-selection, FS=floor seeker, FT=floor-taking,
IR=information receipt tokens, EN=engaged tokens, HT=head turn, HG=hand gestures

This *yeah, yeah* utterance cannot be considered as a response token since C2_FBT has already secured the floor but still has the same function as a response tokens for information receipt tokens in discourse. Further, it is worth noting that this *yeah* is uttered in C2_FBT's speakership in C2.

The last case I would like to highlight in C2_FBT's floor exchanges is a variation of Pattern D, which is the pattern of other-selection (OS). There is only one

case where C2_FBT's taking the floor is led by C2_MJS. In this case, C2_MJS indirectly asks C2_FBT to review his writing and to tell whether his writing is on the right track or not. As shown in Transcript 3.3-17 below, C2_MJS says 'so I'm so afraid of my way is right or not' at time 75 to ask for her advice on his writing. C2_MJS then gives away the floor of the conversation in order to wait for C2_FBT's response:

Transcript 3.3-17 C2_FBT listenership behaviour (4)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtime	MJS_leadtime	C2_FBT_CF	C2_FBT_DF	C2_FBT_gesture	C2_FBT_Transcript	C2_MJS_gesture	C2_MJS_Transcript
75	MJS_F	D	OS	-4	0						'so,I'm so afraid of my way is right or not
76				-3	1					SC/chin	
77				-2	2						
78				-1	3						
79	FBT_F	D+	OS	0	-61	FT			okay,we'll let's,let's have a look at the headings,		

Keys: C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_FBT_CF=C2_FBT conversation function, C2_FBT_DF=C2_FBT discourse function,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_MJS discourse function,
MJS_F= male Japanese student floor-taking, FBT_F=female British tutor floor-taking,
OS=other-selection, FT=floor-taking,
IR=information receipt tokens, EN=engaged tokens, HT=head turn, HG=hand gestures

At time 79, C2_FBS shows understanding of C2_MJS's request although we do not observe any floor seeking before her taking the floor. C2_FBS starts her utterance with the discourse marker *okay*.

In this section, an attempt has been made to capture C2_FBT's floor exchanges. It is recognised that C2_FBT does use listenership behaviours but her strategies are different from the participants in C1. For instance, C2_FBT tends to take the floor of the conversation directly. Although the instances do not quite match with those cases observed in the British-British conversation, there are some

similarities between these two when discourse functions in relation to floor exchanges are considered. The main differences are centred on C2_FBT's expression of listenership, and the less frequent use of listener response tokens is characteristic in C2_FBT. In the case of the variant of Pattern D, which we have just reviewed, for instance, there is no response token in C2_FBT's listener status. The same thing can be said in Pattern E. Although a freestanding *erm* or *mm* is observed during pauses before floor-taking, these items can be recognised as discourse markers related to speakership rather than listener response tokens. Several factors can be raised as characteristics in C2_FBS's listener behaviour, such as her short listener status, the situation where she has reviewed a paper, the participants' power relationship and cultural differences. Further analysis on a longer recording of data in the main study will be crucial in order to draw some conclusions from C2_FBT's listener behaviour.

3.3.2.4 C2_MJS's use of response tokens

C2_MJS used two types of floor exchanges: Pattern C and Pattern D. C2_MJS took the floor eight times, and 7 cases out of 8 are other-selection (OS). Transcript 3.3-18 below shows C2_MJS's listenership behaviour. This case is categorised in Pattern D. At leadtime 71, C2_FBT asks 'I haven't seen this one before, have I?' and C2_MJS gives the freestanding *yeah* twice as convergence tokens at 72 and 74 in leadtime. Then he takes the floor with the discourse marker *so*, which can be recognised as a typical case of Pattern D. This proves that C2_MJS, a Japanese learner of English as Second language, can use the same listenership strategy as the British student in C1, even though we could not see this similarity between C2_MJS and C2_FBT's listenership behaviour in the quantitative analysis:

Transcript 3.3-18 C2_MJS listenership behaviour (1)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtime	MJS_leadtime	C2_FBT_Transcript	C2_MJS_CF	C2_MJS_DF	C2_MJS_gesture	C2_MJS_Transcript
68	FBT_F	E	SS	0	-7	I think,yeah,this one,yeah,			SC/chin	
69				1	-6					
70				2	-5					
71				3	-4	I haven't seen this one before,have I?				
72				4	-3		FS	CV		yeah
73				5	-2	this is the first+			SC/chin	
74				6	-1	+draft	FS	CV		yeah
75	MJS_F	D	OS	-4	0		FT			so,I'm so afraid of my way is right or not

Keys: C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_MJS discourse function,
MJS_F= male Japanese student floor-taking, FBT_F=female British tutor floor-taking,
SS=self-selection, OS=other-selection, E=pattern E, D= pattern D, FS=floor seeker, FT=floor-taking,
CV=convergence tokens, SC/chin=self comfort with chin

There is also a variant of Pattern D observed in C2_MJS, where C2_MJS does not use response tokens as floor seeker and directly starts speaker turn with or without discourse markers such as *yeah* or *well*, which is similar to the variant observed in C2_FBT's floor exchanges:

Transcript 3.3-19 C2_MJS listenership behaviour (2)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtim	MJS_leadtime	C2_FBT_Transcript	C2_MJS_CF	C2_MJS_DF	C2_MJS_gesture	C2_MJS_Transcript
315		D+		143	-8	What else have you read about observation?				
316		D+		144	-7					
317		D+		145	-6					
318		D+		146	-5	<$E> pause <\$E>				
319		D+		147	-4					
320		D+		148	-3					
321		D+		149	-2					
322		D+		150	-1	Oh,you've got Wajnryb,				
323	MJS_F	D+	OS	-2	0		FT			yeah, I've got

Keys: C2_MJS=Conversation 2 (NS-NNS) male Japanese student,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_MJS discourse function,
MJS_F= male Japanese student floor-taking, OS=other-selection, D+= a variant of pattern D, FT=floor-taking,
CV=convergence tokens

At time 322 in Transcript 3.3-19 above, C2_FBT comments on his work and leaves the next floor to C2_MJS. On the other hand, C2_MJS does not produce any response tokens as a floor seeker in the conversation function and goes straight into his taking the floor with the discourse marker *yeah*, which can be interpreted as his agreement to C2_FBT's previous utterance. Less frequent use of response tokens for floor seeker is characterised in C2_MJS, which is also observed in C2_FBT.

Transcript 3.3-20 C2_MJS listenership behaviour (3)

Timeline	Floor	Pattern	Speaker selection	FBT_leadtime	MJS_leadtime	C2_FBT_Transcript	C2_MJS_CF	C2_MJS_DF	C2_MJS_gesture	C2_MJS_Transcript
442	FBT_F	E	SS	0	-24	so for this reason, it is helpful to have another person in a room				
443				1	-23					
444				2	-22					
445				3	-21					
446				4	-20					
447				5	-19					
448				6	-18		LS	CN	HN	mm
449				7	-17	ah sorry,	LS	CN		mm
450				8	-16	I=, I might still a little bit, for this reason				
451				9	-15					
452	pause			-17	-14	<$E> pause <\$E>				
453				-16	-13					
454				-15	-12					
455				-14	-11					
456				-13	-10	mm				
457				-12	-9					
458				-11	-8	<$E> pause <\$E>				
459				-10	-7					
460				-9	-6					
461				-8	-5					
462				-7	-4					
463				-6	-3					
464				-5	-2					
465				-4	-1		FS			un
466	MJS_F	C+	SS	-3	0		FT			probably I mention two kinds of observation

Keys: C2_MJS=Conversation 2 (NS-NNS) male Japanese student,
FBT_leadtime=female British tutor leadtime, MJS_leadtime=male Japanese students leadtime,
C2_MJS_CF=C2_MJS conversation function, C2_MJS_DF=C2_MJS discourse function,
MJS_F= male Japanese student floor-taking, SS=self-selection, E =pattern E, C+= a variant of pattern D,
LS=listenership, FS=floor seeker, FT=floor-taking, CN=continuers, HN=head nods

There is only one case where C2_MJS's speaker turn was led by himself. This case can be seen as a variant of Pattern C. In Transcript 3.3-20 above, C2_FBT takes the floor from 442 to 451, and C2_MJS gives a freestanding *mm* twice at 448 and 449, which can be interpreted as listenership within the conversation function and continuers within the discourse function.

In addition, in Pattern C, a listener sometimes attempts to take the floor in the middle of listener status, but fails to take the floor then goes back to listenership. In C2_MJS's case, however, no floor seeker is observed in the middle of his listener status. One second before the floor-taking, C2_MJS uses a Japanese response token *un*, which can be taken as a discourse marker for floor seeking since there is a pause before C2_MJS's floor-taking and both of the participants are in silence. Even though this C2_MJS's listener status includes a pause, I would construe this instance as a variant of Pattern C rather than Pattern E since some similarities can be seen in the listenership behaviour towards taking the floor in this instance with Pattern C. This instance, for example, has rather long listenership status and C2_MJS keeps producing response tokens for listenership during C2_FBT's speakership and waiting for a more accessible turn relevant point.

The discourse marker *un* is rarely seen in English conversation, however, *un* or *u:n* is often observed in Japanese conversation (see Chapter Two, Mori 2002), which can be interpreted as *yeah* or *well* for agreement or acknowledgment with hesitation depending on the situation. In the current research, I will not examine the features of response tokens in Japanese conversation in depth, and C2_MJS's use of this response token might be treated as an example of learners' L1 (first language) transfer. This can be thought of as a case where features of spoken discourse in

learners' native languages are reflected in their use of the target language in an interlanguage setting.

3.3.2.5 Summary of floor-taking patterns

In order to summarise the floor-taking patterns observed in C1 and C2, the numbers of occurrences of each of these five patterns in each participant were counted. As shown in Table 3.3.2-10 below, the first column describes the five turn-taking patterns: Pattern A to Pattern E, and Unclassified. The first row is the participants' name, such as C1_MBT. Each participant's floor exchanges are divided into the two speaker selection types, self-selection (SS) and other-selection (OS), which are indicated in the second row.

Table 3.3.2-10 No. of floor-taking patterns in C1 and C2 by participants

	C1_MBT		C1_FBS		C2_FBT		C2_MJS	
	SS	OS	SS	OS	SS	OS	SS	OS
	20	0	6	13	24	1	1	7
Pattern A	13 (A+, 3)	--	0	0	5 (A+, 5)	0	0	0
Pattern B	1	--	3	0	0	0	0	0
Pattern C	3	--	3	0	0	0	1 (C+ , 1)	0
Pattern D	--	--	--	13 (D+, 3)	--	1 (D+, 1)	--	7 (D+ , 5)
Pattern E	0	--	0	0	17	0	0	0
Unclassfied	3	--	0	0	2	0	0	0

Keys: C1=conversation 1 (British-British conversation), C2=conversation 2 (British-Japanese conversation),
C1_MBT=C1 male British tutor, C1_FBS=C1 female British student.
C2_FBT=C2 female British tutor, C2_MJS=C2 male Japanese student, SS=self-selection, OS=other-selection,
A+ = variants of Pattern A, C+= variants of Pattern C, D+= variants of Pattern D
13 (A+, 3) means three cases out of thirteen are variants of Pattern A

In the third row, the total number of times each participant took the floor is shown, which is grouped by speaker selection types. C1_MBT took the floor 20 times and 13 of these cases are Pattern A. C1_MBT also has one Pattern B, 3 Pattern C, and 3 Unclassified. However, no other-selection floor-taking is observed in C1_MBT. In C1_FBS, there are 6 floor exchanges led by herself, and a half of them are Pattern B,

the case of short listenership, and the other half is Pattern C, the case of longer listenership. All of C1_FBS's other-selection floor-taking follows Pattern D.

As for C2, the British-Japanese conversation, some variants of the patterns are observed, which I indicate with a symbol +. A variant of Pattern A, for instance, is described as Pattern A+. In C2_FBT, self-selection floor-taking occurs 24 times. The 17 cases out of 24 are Pattern E, which has pauses before the floor-taking. Five cases in C2_FBT's self-selection floor-taking are categorised as Pattern A. All of them are variants of Pattern A, which is also indicated as A+ in brackets, and the remaining two cases are put into Unclassified. C2_FBT's also has one case of other-selection, which is categorised as a variant of Pattern D. As for C2_MJS, only one case of self-selection is observed, which is categorised as a variant of Pattern C. C2_MJS has 7 other-selection floor exchanges, all of which are identified as Pattern D. Five cases of them are labelled as variants of Pattern D+.

From the overview of the summary of floor-taking patterns, it can be observed that there are some similarities observed in the strategies of floor exchanges between the two tutors, C1_MBT and C2_FBT. Both of them have Pattern A, although C2_FBT has more variants of the pattern. Similarities in floor exchanges between the students, C1_FBS and C2_MJS, were also observed, for they shared Pattern C and Pattern D, which was not recognised in the quantitative analysis.

3.3.2.6 Collocations of verbal response tokens with visual response tokens

In the previous section, the floor-taking patterns placing focus on verbal response tokens were examined. In this section, collocations of visual response tokens, such as head nods (HNs) and hand gestures (HGs), with verbal response tokens in relation to

turn-taking will be analysed.

In the process of analysis, each participant's verbal response tokens were extracted according to their conversation function and floor-taking pattern. Transcript 3.3-21 below, for example, shows C1_MBT's verbal response tokens, which are used as listenership in Pattern A:

Transcript 3.3-21 C1_MBT Pattern A: Response tokens in listenership

Timeline	Floor	Pattern	MBT_leadtime	FBS_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
15		A	-20	7	LS	CN	HN	Yeah.
20		A	-15	12	LS	CN		Aha.
43		A	-23	7	LS	CN		Right.
53		A	-13	17	LS	CN		Right.
177		A	-9	6	LS	CN		Yeah.
346		A	-7	2	LS	CN	SC/arm	Yeah.
348		A	-5	4	LS	CN	HN	Yeah.
506		A	-7	5	LS	CN	HN	Mm.

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
MBT_leadtime=male British tutor leadtime, FBS_leadtime= female British student leadtime,
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
A=pattern A, LS=listenership, CN=continuers, HN=head nods, SC/arm=self comfort with arm

The third column of the transcript above indicates that these response tokens are used in Pattern A, and the sixth column shows that these verbal response tokens function as listenership (LS). There are 8 verbal response tokens, which are uttered as listenership in Pattern A in C1_MBT's listener status, and 3 cases out of the 8 verbal response tokens co-occur with head nods (HNs) as shown in the column of C1_MBT_gesture in the transcript. At time 346, SC/arm (self comfort with arm) is observed; however, I do not take this as a visual response token. In order to narrow down the focus, two visual response tokens, head nods (HNs) and hand gestures (HGs), were targeted in the current study.

The same operation was conducted to examine C1_MBT's response tokens which were used as FS in Pattern A as described in Transcript A.2-1 in the appendix,

and the variants of this latter pattern were included in this analysis. Hence both Pattern A and Pattern A+ can be observed in the third column of the transcript. Further, a list of 24 verbal response tokens including several kinds of minimal responses were extracted, two cases of which are accompanied by HNs.

Although we can see SC/neck (self comfort with neck) at time 306, again I do not take this body movement for self comfort as a visual response token. Whether any gesture can be observed at C1_MBT's TTP in Pattern A or not was also examined. It can be said that the gestures used at TTP can function as both response tokens and discourse markers depending on the context. Additionally, it was worth analysing the use of gestures at TTP in relation to the floor-taking patterns. C1_MBT's utterances at TTP in Pattern A are selected in Transcript A.2-2 in the appendix. In C1_MBT's floor exchanges in Pattern A, 6 out of 13 are with gestures. In detail, C1_MBT has taken the floor four times with HGs, once with HNs and once with HSs and HGs. Indications of these gestures can be seen in the last second column, C1_MBT_gesture, in Transcript A-2-2 in the appendix. The use of HGs is also observed at times 35, 109, 186 and 513. A HN with discourse marker *yeah* is observed at time 66. At 481, he uses both a HG and a HS when he takes the floor.

After investigation of the collocations of visual response tokens with verbal response tokens in Pattern A, I have summarised the results as shown in Table 3.3.2-11 below. The first column shows the floor-taking pattern, Pattern A, and the second column shows conversation function, LS (listenership), FS (floor seeker) and FT (floor-taking):

Table 3.3.2-11 Pattern A: collocation of verbal and visual response tokens

				C1_MBT With gestture/ All (details)	C1_FBS With gestture/ All (details)	C2_FBT With gestture/ All (details)	C2_MJS With gestture/ All (details)
Pattern A (SS)	LS	HN	Continuers	3 / 8 (HN3)	--	0/0	--
	FS	HN	Information receipt token	2 / 24 (HN2)	--	0/0	--
	FT	HG	--	6 / 13 (HG4, HN1, HG+HS1)	--	0/ 4	--

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_FBS=Conversation 1 (British-British conversation) female British student
C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
with gestures/ all = the number of verbal response tokens with gestures/ the total number of verbal response tokens,
details= the details of visual response tokens, SS= Self-selection, LS= listenership, FS= floor seeker, FT= floor-taking,
HN= head nods, HG= hand gestures, HG+HS= hand gestures + head shake

Besides these conversation functions, frequently observed visual response tokens are indicated. In the third column, the discourse functions of response tokens are described, such as continuers and information receipt tokens. Then the numbers of each participant's verbal and visual response tokens according to the conversation function in Pattern A are indicated. In C1_MBT's use of response tokens as listenership in Pattern A, for instance, is described as '3/8', which means 3 out of 8 verbal response tokens are accompanied by visual response tokens. Under this value, there are details of visual response tokens in brackets. In the case of C1_MBT's listenership, these 3 verbal response tokens are all head nods, which is shown as '(HN3)' under the value. In the case of C1_MBT's floor seeker, two out of 24 verbal response tokens occurred with visual response tokens, which are also HNs. Alternately, in C1_MBT's floor-taking, 6 out of 13 cases are with visual response tokens. The details show that 4 of these 6 visual response tokens are HGs, and the remaining 2 are a HN and a HS with HGs. C2_FBT has 4 floor exchanges in Pattern A although these 4 instances are variants of Pattern A. As shown in the last two

column of the table, no verbal response tokens are observed in listenership and floor seeker. Although there are 4 cases where verbal response tokens have been uttered by C2_FBT at the floor-taking, no collocation of verbal and visual response tokens was observed. This means that, even at TTPs, C2_FBT has not used any hand gestures in Pattern A.

In the same way, I have examined the collocations of verbal and visual response tokens by participants in relation to the floor-taking patterns. In Pattern B as described in Table A.2.1-1 in the appendix, C1_MBT uses verbal response tokens for floor seeker with HNs, and when he takes the floor, he uses HGs although the case occurs only once in the 10-minutes pilot data. There are three times where C1_FBS takes the floor following Pattern B. Five verbal response tokens are used for floor seeker and one of them occurs with HSs (head shakes). There is one case where C1_FBS uses HGs when she takes the floor in Pattern B.

Three participants, C1_MBT, C1_FBS and C2_MJS, had Pattern C in their floor exchanges as shown in Table A.2.1-2 in the appendix. C1_MBT takes the floor three times with Pattern C, and always uses HGs at these points. Fifteen verbal response tokens for listenership are also observed in the cases of C1_MBT's Pattern C and he uses a HN once. For floor seeking, C1_MBT used verbal response tokens five times, one of which occurs with a HN. In C1_FBS's cases of Pattern C, verbal response tokens for listenership are observed 15 times, 11 of which are with visual response tokens. Most of them are HNs and there is only one case where a HG is also used with a HN. Also, in the last column, C2_MJS's use of verbal and visual response tokens in Pattern C is summarised. There are two cases where C2_MJS has used verbal response tokens for listenership and one of them is accompanied by a HN. For

floor seeking, he used a verbal response token once without any gestures. At the floor-taking, however, C2_MJS used HGs although this case occurred only once in the pilot data.

Although Patten D was observed often in students' floor-taking in the two conversations, there was one case where C2_FBT followed Pattern D. C2_FBT, however, did not use any verbal response token for floor seeker and took the floor without gestures as described in Table A.2.1-3 in the appendix. In contrast, frequent use of visual response tokens was observed in C1_FBS's cases of Pattern D. C1_FBS produced 5 response tokens for listenership and 3 of them were with HNs. There were 9 verbal response tokens used for floor seeker in C1_FBS's cases, and 4 cases out of 9 occurred with visual response tokens such as HNs and HNs with HGs. At the floor-taking, C1_FBS used visual response tokens 6 out of 13 times. In detail, she used 3 HGs, 1 HN, 1 HN with HGs and 1 HS with HGs. This same tendency was also seen in C2_MJS's listenership behaviour for listenership in Pattern D. C2_MJS used 21 verbal response tokens for listenership, and 12 times of them were accompanied by HNs. For floor seeking, however, C2_MJS has used 3 verbal response tokens without gestures and he did not use gestures at floor-taking either.

Pattern E, as shown in Table A.2.1-4 in the appendix, is only observed in C2_FBT's floor exchanges. She uses discourse markers for floor seeker in Pattern E seven times, none of which occur with gestures. At the floor-taking, whereas, HT (head turns) are observed twice out of 18 C2_FBT's floor-taking in Pattern E.

For the summary of the turn structure analysis, I established a proposed model of the five floor-taking patterns with verbal and visual response tokens based on the native British speakers' typical floor-taking patterns:

Table 3.3.2-12 Pattern A (SS) with visual response tokens

	Leadtime	Verbal response tokens	Functions	Visual response tokens
LS (Listenership)	x < approx. -5	freestanding *yeah* or *right* or *aha*	Continuers	HN
FS (Floor seeker)	approx .-5 < x	minimal response with *right*	Information receipt tokens	HN
FT (Floor-taking)	0 = x	yeah or well or so or but + full turn	-	HG

Keys: SS=Self-selection, HN= head nods, HG= hand gestures

Table 3.3.2-13 Pattern B (SS) with visual response tokens

	Leadtime	Verbal response tokens	Functions	Visual Response tokens
FS (Floor seeker)	approx .-5 < x	freestanding *yeah* or *no*	Engaged tokens	HN or HS
FT (Floor-taking)	0 = x	*and* or *well* or no discourse marker + full turn	-	HG

Keys: SS=Self-selection, HN= head nods, HS= head shakes, HG= hand gestures

Table 3.3.2-14 Pattern C (SS) with visual response tokens

	Leadtime	Verbal response tokens	Functions	Visual Response tokens
LS (Listenership)	-	freestanding *yeah* or *right* or *okay*	Continuers	HN
FS (Floor seeker)	-	*yeah* or *oh* + some words	Information receipt tokens Engaged tokens	HN or HN+HG
LS (Listenership)	-	freestanding *yeah* or *right* or *okay*	Continuers	HN
FT (Floor-taking)	0 = x	yeah or and + full turn	-	HG

Keys: SS=Self-selection, HN= head nods, HG= hand gestures

Table 3.3.2-15 Pattern D (OS) with visual response tokens

	Leadtime	Verbal response tokens	Functions	Visual Response tokens
LS (Listenership)	-	freestanding *right* or *okay* or *yeah*	Continuers	HN
FG (Floor giving) from the partner	approx. -1 > x	questions or comments from the participant	-	-
FS (Floor seeker)	-1 < x or 0 = x	freestanding *yeah* or *no*	convergence tokens	HN or HN + HG
FT (Floor-taking)	0 = x	(erm +) yeah or no or well + full turn , some times	-	HG

Keys: OS=other-selection, HN= head nods, HS= head shakes, HG= hand gestures

Table 3.3.2-16 Pattern E (SS) with visual response tokens

	Leadtime	Verbal response tokens	Functions	Visual response tokens
Pause or FS (Floor seeker)	-	Pause, of pause with freestanding *erm* or *mm*	Pause or discourse marker	-
FT (Floor-taking)	0 = x	*okay* or *yeah* or no discourse marker+ full turn	-	HT

Keys: SS=Self-selection, HT= head turns

In terms of Pattern A, HNs were used as listenership and floor seeker with verbal response tokens, which are indicated in the last column in Table 3.3.2-12 above. HGs were also accompanied with verbal response tokens at the floor-taking in Pattern A. As for Pattern B, HNs or HSs were observed with verbal response tokens for floor seeker, and HGs were used at the floor-taking in Pattern B as same as Pattern A. These instances of Pattern B with visual response tokens are summarised in Table 3.3.2-13 above. Again, visual response tokens are described in the last column.

Alternately, in Pattern C as described in Table 3.3.2-14, HNs were used for listenership. HNs or HNs with HGs were also observed with verbal response tokens for floor seeker. As with the previous two patterns, HGs were used at the floor-taking in pattern C. The same tendency as Pattern C was observed in Pattern D, as shown in Table 3.3.2-15, although Pattern D is the only case of other-selection. In Pattern D, HNs occurred with verbal response tokens for listenership and floor seeking. There were some cases where HNs were used with HGs when the participants were seeking floor. At the floor-taking, the use of HGs was identified. Pattern E has distinguished characteristics as shown in Table 3.3.2-16. Neither verbal nor visual response tokens are used in Pattern E, although at the floor-taking, however, a particular gesture, HT

(head turns) is observed.

At this stage, I will leave this proposed model for the turn-taking patterns with verbal and visual response tokens since it might not be accurate without more precise descriptions about the occurrences of each collocation of verbal and visual response tokens. These models will be re-examined at a later stage in the main study.

3.4 Summary

In this chapter, a research method of a time-related corpus-based approach with the key concept, leadtime, was introduced in the pilot study. This specifically addresses one of the research principles: that is, to establish a new model for conversation analysis with a time-related transcript and multi-modal annotations by introducing the concept of leadtime (see Section 1.2). The preliminary findings from the pilot study have also been reported, and with categorizations of discourse function of response tokens and conversation function related to turn-taking structure, five turn-taking patterns are recognised. Pattern A (LS → FS → FT) is frequently used by the tutors. Pattern D (FG → FT) is often used by the students. Collocations of verbal response tokens with visual response tokens, namely HNs and HGs, were also examined. These findings were later compared with the results from the main study for scalability of the research. In the next chapter, a model for analysing listenership behaviour in relation to turn-taking structure will be explained based on the pilot study. This model was used in the main study.

Chapter 4 Research: Main study

4.0 Introduction

The aim of this chapter is to explain a model for analysing listenership in conversation based on the research method developed through the course of the pilot study, with reference to categorisations of response tokens established by O'Keeffe and Adolphs (Knight et al. 2006, O'Keeffe & Adolphs 2008, O'Keeffe et al. 2007). As referred to in the earlier chapter (Section 2.1), most of the previous research in conversation analysis focuses on the speaker's role in conversation (Gumperz 1978, Heritage 1997, Sacks 1992, Schegloff 2007, Tannen 1984). In recent study, more and more researchers have become aware of listenership behaviour in conversation and a number of research studies on the listener's role have been conducted (Drummond & Hopper 1993, Duncan & Niederehe 1974, Fellegy 1995, Gardner 2002, McCarthy 2002, O'Keeffe et al. 2007). However, many areas still remain to be unveiled in particularly regarding listenership research.

A research method for conversation analysis with a time-related corpus was suggested based on the existing studies and the pilot study of this project. The concept of leadtime was applied to the new research method in order to make it possible to measure the time distance between the point where particular response tokens are uttered and the point where floor exchanges occur.

Two approaches for analysing listenership applied to the main study will also be described in this chapter: global pattern analysis and turn structure analysis. In the global pattern analysis, the targeted response tokens were counted and summarised in order to visualise the frequency of the targeted response tokens in relation to turn-taking structure. In the turn structure analysis, the use of verbal and visual response

tokens was investigated qualitatively in reference to turn structure. Seven turn-structural episodes were evidenced in the turn structure analysis. Findings from these two approaches with a time-related corpus will be reported in the next chapter.

4.1 A bridge from the pilot study to the main study

4.1.1 Scope of the main study

As stated in Chapter One, the main study question of this study is:

> What are the differences and similarities between British-British conversation and British-Japanese conversation in English in the context of academic supervision sessions?

In order to answer this, a research method for conducting linguistic research with time-related multimodal corpus needs to be established and this research method for analysing turn-taking structure will be treated as one of the original contributions made by the current study as described in Chapter One (see Section 1.2).

The first priority for my research was to investigate differences in strategies of the use of response tokens comparing British-British conversation with British-Japanese conversation in English. The second priority was placed on establishing a methodological framework for conversation analysis implementing the concept of leadtime. Another unique aspect of the current project is that visual response tokens were analysed in addition to verbal response tokens with a time-related corpus. As described in Chapter One, the current study attempts to pursue three issues: (1) to establish a method for analysing turn structure; (2) to indicate preliminary results from the turn structure analysis; and (3) to highlight areas for future research. The

practical procedures in the current study have also been indicated in Chapter One (see Section 1.3).

In this chapter, these procedures will be improved and described in detail based on the pilot study to the extent which is adequate to conduct the main study.

4.1.2 How does the pilot study and the main study relate?

In the course of the pilot study, a methodological framework integrating leadtime for conversation analysis was developed, which was then applied and extended in the main study. As described in the previous chapter, a new concept 'leadtime' was introduced in the current research (see Section 3.1.6). In the existing studies of response tokens at the early stage (Duncan 1974, Gardner 2002, LoCastro 1987, Maynard 1990), functions of response tokens were analysed through observations which were unaware of a concept of time. By implementing the model with leadtime, the time relationship between the use of particular response tokens and turn-taking structure can be analysed. In addition to leadtime, the concepts of TTP, speaker turns and backchannel turns were also applied in the main study. A *speaker turn* outlines a turn where a participant is holding the floor of the conversation, whereas a *backchannel turn* is a turn where a participant signals with response tokens without holding the floor of the conversation.

For scalability of the research, four 39-minute length conversation data were analysed in the main study while two 10-minute conversation data were examined in the pilot study. The former four sets of conversation data used were reduced to 39-minute length for the sake of comparability. Two of the four sets of conversation data used in the main study are the same conversation data examined in the pilot study.

4.1.3 Two approaches: global pattern analysis and turn structure analysis

The main study consists of two parts: namely *global pattern analysis* and *turn-taking structure analysis*. In the global pattern analysis, several types of quantitative analysis were conducted with the 39-minute length data of the four conversations. This data analysis was designed to show an overview of the turn management in conversation and the trends in the use of response tokens in relation to leadtime. In the turn structure analysis, the four conversations were analysed qualitatively, and although the amount of pauses in the conversations and the numbers of participants' floor exchanges vary from conversation to conversation, features in listenership of each participant were extracted and categorised according to turn-structural episodes. The items investigated with these two approaches are listed as follows:

- Global pattern analysis – quantitative analysis
  - Objectives
    1. To summarise the length of speaker status of each participant,
    2. To summarise the number of speaker turns of each participant and the average length of speaker turns,
    3. To summarise the number of verbal response tokens, such as *erm*, *yeah*, *mm* and *mhm*, and visual response tokens, such as head nods and hand gestures.
- Turn-taking structure analysis – qualitative analysis
  - Objectives
    1. To analyse forms and placements of verbal and visual response

tokens in reference to turn-structural episodes and discourse frameworks by integrating the concept of leadtime,

2. To conduct descriptive analysis on the use of hand gestures with time-related multimodal transcripts and image captures in reference to multi-functional nature of hand gestures.

In the pilot study, collocations of verbal response tokens with visual response tokens were examined from a quantitative approach. In the main study, however, this aspect was excluded from the research objective in order to narrow down the focus, although co-occurrences of visual response tokens with verbal response tokens were analysed descriptively.

In the analysis of the turn length and the placement of particular response tokens, some referential statistics, namely means, standard deviations and variance, were used as supplemental data to provide additional descriptions when response tokens were used in reference to turn taking structure.

4.1.4 Targeted items of response tokens

Based on the classification applied in the pilot study in Section 3.1.7, six items were selected for the analysis in the main study; four verbal response tokens, *erm*, *yeah*, *mm*, and *mhm* and two visual response tokens, head nods and hand gestures. In the pilot study, a minimal response *mm* included the two vocalised sounds *mm* and *mhm*. Through the process of annotation of the data, I recognised and observed the differences between these two sounds. In some existing studies, these sounds are described separately as *mm* and *mhm*, while other transcripts combine and transcribe these two sounds as an expression *mm*. The decision was made to treat these two

sounds as two separate response tokens in the current research, adapting to Carter and McCarthy (2006). I also assumed that these two minimal response tokens, *mm* and *mhm*, might have some differences and similarities in their functions in conversation.

Table 4.1.4-1 Targeted items of response tokens

			Notes
Vocal response tokens	Erm		
	Yeah		
	Mm		
	Mhm		
Visual response tokens	Head nods	HN	Any vertical head movement
	Hand gesture	HG	Any hand gestures

Conversational gestures have been categorised and analysed in detail by researchers such as Kendon (1972) and Knight et al (2006) as reviewed in Section 2.1.5, whereas relatively simple definitions are given to hand gestures and head nods in the current study. Hand gestures (HGs) are defined as 'any hand movements' and head nods (HNs) are defined as 'any vertical head movements' as described previously in Section 3.1.7. When HGs and HNs are counted, the spaces between movements are taken into account and in the case where more than one HG is observed within one second period, these HGs are treated as one HG on the time-related transcripts with the timeline noted in seconds. The same rule is applied to counting HNs.

There are many variations in the use of these selected items, such as *yes* instead of *yeah*, or non-minimal response *right okay* instead of a minimal response *right*. Based on the study by Drummond and Hopper (1993), these selected items were collected and summarised systematically in the quantitative analysis in the main

study. In other words, variations in forms of these selected items were ignored in the quantitative analysis. In the *qualitative* analysis, however, variations in forms of these selected items were examined descriptively.

4.1.5 Leadtime – A review

As described in Section 1.3 and Section 3.1.6, leadtime is defined as a time scale to measure the length of time of speaker/listener status with the turn transition point as a datum point, which is described as 0 in leadtime. Leadtime is used to describe the time distance between the point where a response token or a discourse marker is used and the floor transition point.

A variety of corpus-based approaches have been taken in recent linguistic research (Aston & Burnard 1998, McEnery et al. 2006, Tono 2004) and some research in conversation analysis has also been conducted with video-recorded data (Carroll 2004, Heath 1997). However, as far as I have studied, time-corpus approach with multimodal data has not yet been implemented in linguistic study. In the current research, the use of verbal and visual response tokens can be recognised and analysed in reference to leadtime. In Transcript 4.1-1 below, for instance, a minimal response *right* is uttered 7 seconds before BBC1_MBT's floor-taking at time 00:17:18. As shown below, visual response tokens, such as head nods (HNs) and hand gestures (HGs), are also annotated in transcripts. Here, a HG is observed in BBC1_FBS_gesture in the second to last column at 00:17:05 in the timeline, which indicates that BBC1_FBS used a HG when she has taken the floor of the conversation at 00:17:05:

Transcript 4.1-1 Sample extract for leadtime

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 17: 05	FBS_F	-13	0			HG	It's just too difficult for it.
00: 17: 06		-12	1				
00: 17: 07		-11	2		+literally erm.	HN	
00: 17: 08		-10	3			HG	Yeah the stuff I've seen it's just two word erm+
00: 17: 09		-9	4			HG	
00: 17: 10		-8	5				
00: 17: 11		-7	6	HN	Right.		
00: 17: 12		-6	7		Right yeah yeah.		+metaphors really.
00: 17: 13		-5	8				
00: 17: 14		-4	9				So I mean you get like grammatical metaphors with+
00: 17: 15		-3	10			HG	
00: 17: 16		-2	11	HN	Yeah yeah yeah.		+verbs and stuff.
00: 17: 17		-1	12				
00: 17: 18	MBT_F	0	-3		Yeah that's still clever if you can do that.	HG	

Keys: BBC1_MBT_leadtime =British-British Conversation 1 (British-British conversation) male British tutor's leadtime,
BBC1_FBS_leadtime =British-British Conversation 1 (British-British conversation) female British student's leadtime,
BBC1_MBT_gesture=BBC1_MBT's gesture, BBC1_MBT_Transcript=BBC1_MBT's transcription,
BBC1_FBS_gesture=BBC1_ FBS's gesture, BBC1_ FBS _Transcript=BBC1_ FBS's transcription,
MBT_F=MBT floor-taking, FBS_F=FBS floor-taking,
HN= head nods, HG= hand gestures

Leadtime is displayed as negative numbers when participants are in listener status and as positive numbers when they are in speaker status. Leadtime continues to be incremented until the next floor-taking instance.

4.1.6 Turn-structural episodes

Levison (1992) treats the term *episode* as a synonym of *speech event*, which refers to 'socially constituted, bounded, events with "constraints" on participants, setting, and so on' (ibid: 69, original author's emphasis). In turn, Adolphs (2008) defines an episode as 'the negotiation of a particular discourse function' (ibid: 95-96). In this study, episodes refer to recognised patterns in turn management strategies in conversation.

Based on the pilot study in Chapter Three and Ohama (2006), an attempt was made here to establish turn-structural episodes to categorise turn-taking patterns. In the pilot study in Chapter Three, five turn-taking patterns were recognised as shown in Table 4.1.6-1:

Table 4.1.6-1 Turn-taking patterns in the pilot study

	Pattern A (SS)	Pattern B (SS)	Pattern C (SS)	Pattern D (OS)	Pattern E (SS)
LS	Continuers	-	Continuers	(Continuers)	-
FG	-	-	-	Partner's floor giving	-
FS	Information receipt tokens	Engaged tokens	Information receipt tokens/ Engaged tokens	Convergence tokens	Pause/ Discourse markers
LS	-	-	Continuers	-	-
FT	Floor taking	Floor taking	Floor taking	Floor taking	Floor taking

Keys: SS=Self-selection, OS=Other-selection, LS=Listenership, FG=Floor giving, FS=Floor seeker, FT=Floor-taking,

These patterns have been defined based on speaker selections and functions of response tokens. In Sacks (1992), two types of speaker selections were noted, namely self-selection (SS) and other-selection (OS) (see Section 2.1.2.1). When a participant chooses himself as a speaker, the case is defined as self-selection. In a case of other-selection, a current speaker nominates the next speaker. In the pilot study, Patterns A, B, C and E are cases of self-selection and only pattern D is other-selection.

Ohama (2006) studied the relationship between turn-structural patterns and response tokens in Japanese conversation, and distinguished seven turn-taking patterns with five variants based on Sacks (1974) and van Lier (1988) as described in Table 4.1.6-2 below:

Table 4.1.6-2 Turn-taking patterns in Ohama (2006)

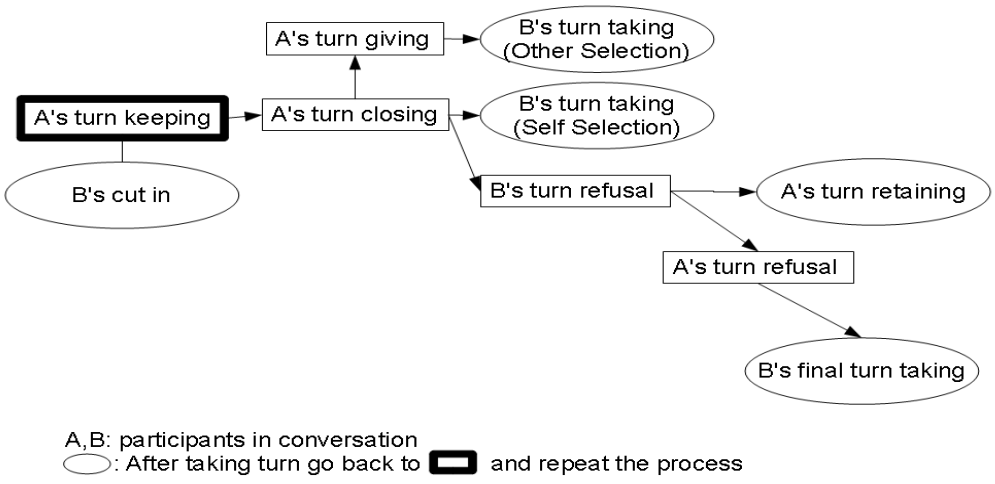
Self-selection	([+taking], [+direct], [+partner], [+closing], [+self])
Other-selection	([+taking], [+direct], [+partner], [+closing], [-self])
Turn refusal	([-taking], [+direct], [+partner], [+closing])
Turn retaining	([+taking], [-direct], [-partner], [+closing], [+self])
Turn re-refusal	([-taking], [-direct], [+/-partner], [+closing])
Final turn-taking	([+taking], [-direct], [+/-partner], [+closing], [+self])
Cut-in	([+taking], [+direct], [+partner], [-closing], [+self])

(Ohama 2006:46-47, translated by me)

The five variables shown above comprise the terms: ‘taking’ for turn-taking, ‘direct’

for whether the previous turn is a speaker turn or a response token, 'partner' for whether the previous turn belongs to the partner or not, 'closing' for whether turn-taking occurs at TRP, and 'self' for whether the turn has been self-selection.

From these variables, Ohama identifies seven turn-structural patterns which have been illustrated in Figure 4.1-1 below. It starts with A's turn keeping and one of the possibilities is the occurrence of B's turn-taking after A's turn closing, which is categorised as self-selection. Another possibility is B's turn-taking after A's turn giving, which is other-selection. The other pattern can be B's turn refusal after A's turn closing, which is described as *turn refusal*.



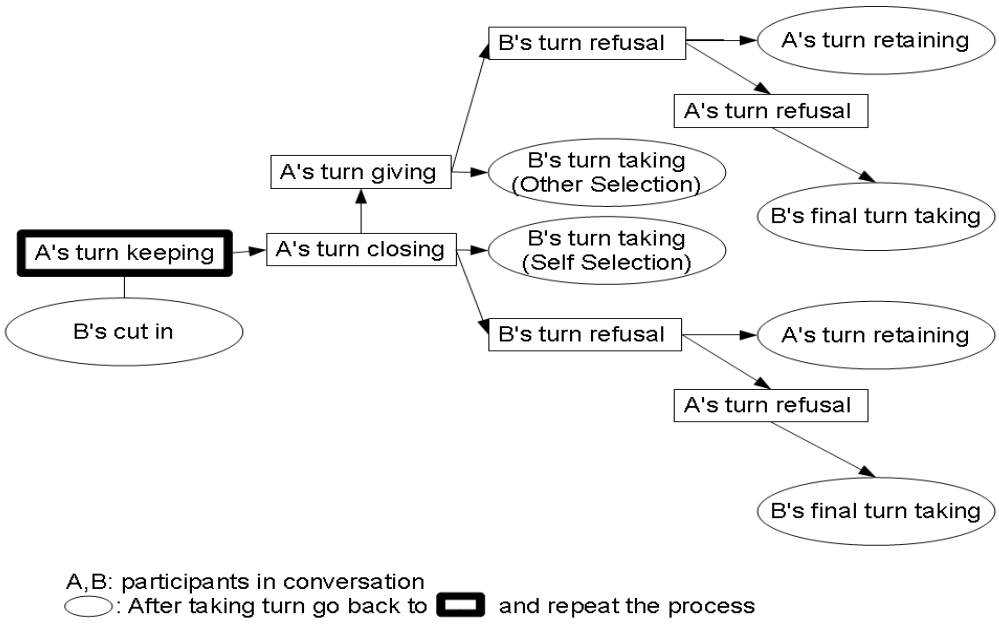
(Ohama 2006:48, translated by me)

Figure 4.1-1 A turn shifting mechanism with seven turn-taking types

When A retains a speaker turn after B's turn refusal, this is categorised as *turn retaining*; when A also disclaims the turn after B's turn refusal, this is also labelled as

turn-refusal; and when B takes a speaker turn after A's turn refusal, this is described as *final turn-taking*. The last type is *cut-in*, which is described as B's cut-in during A's turn keeping in the figure above.

I have revised the turn-taking mechanism and added two other possibilities as illustrated in Figure 4.1-2 below. One of the additional cases is A's turn retaining after B's turn refusal of A's turn giving. The other case is B's final turn-taking after A's turn giving. These two patterns were derived from the analysis of the current data.



(Adapted from Ohama 2006:48, translated by me)

Figure 4.1-2 A revised version of a turn shifting mechanism

As mentioned above, based on Ohama's (2006) model of turn shifting mechanism, seven turn-structural episodes have been established, placing focus on turn exchanges as shown in Table 4.1.6-3 below:

Table 4.1.6-3 Turn-structural episodes

Episode 1	A's turn closing → B's turn-taking
Episode 2	A's turn keeping → B's cut-in
Episode 3	A's turn closing → B's turn refusal → A's turn retaining
Episode 4	A's turn closing → B's turn refusal → A's turn refusal → B's final turn-taking
Episode 5	A's turn giving → B's turn-taking
Episode 6	A's turn giving → B's turn refusal → A's turn retaining
Episode 7	A's turn giving → B's turn refusal → A's turn refusal → B's final turn-taking

(Adapted from Ohama 2006)

These turn-structural episodes were used in the analysis in the main study, in relation to listenership behaviour recognised in the pilot study.

4.2 Data for the main study

4.2.1 Research data

The data required for the main study was collected at the University of Nottingham and Nottingham Trent University from 2005 to 2007. Two sets of British-British conversation data and two sets of British-Japanese conversation data were video-recorded for the main study. The two British-British conversations were recorded and initially transcribed by the research project members in School of English Studies at the University of Nottingham although annotations including time stamps on these transcriptions were added by me in order to make the transcripts accurate in relation to the use of response tokens. The two British-Japanese conversation data were recorded and transcribed by myself.

There were eight participants, four British tutors, two British students and two Japanese students. The first conversation is referred to as British-British Conversation 1 (BBC1), where a male British tutor (BBC1_MBT) and a female British student (BBC1_FBS) are having a supervision session on her MA dissertation about doctor-patient interactions. The second conversation data is referred to as British-British

Conversation 2 (BBC2), where a male British tutor (BBC2_MBT) and a male British student (BBC2_MBS) are having a supervision session on his PhD thesis concerning healthcare language. These two British-British conversations were recorded in the School of English Studies at the University of Nottingham. The third conversation is labelled as British-Japanese Conversation 1 (BJC1), where a female British tutor (BJC1_FBT) and a male Japanese student (BJC1_MJS) are having a tutorial on his assignment essay concerning classroom observation and his MA dissertation on teacher-student interactions in English classes in Japan. BJC1 was recorded during the MA in English Language Teaching course at Nottingham Trent University. The fourth conversation is labelled as British- Japanese conversation 2 (BJC2), where a male British tutor and a male Japanese student (MJS) are having a supervision session on his MA dissertation about English literature in the School of English Studies at the University of Nottingham. The four sets of conversation data are listed in Table 4.2.1-1 below (also see Table E.1.2-1):

Table 4.2.1-1 Participants for the main study

	Participants		Supervisions
	Tutor	Student	
British-British Conversation 1 (BBC1)	BBC1_MBT	BBC1_FBS	MA dissertation
British-British Conversation 2 (BBC2)	BBC2_MBT	BBC2_MBS	PhD thesis
British-Japanese Conversation 1 (BJC1)	BJC1_FBT	BJC1_MJS	MA dissertation
British-Japanese Conversation 1 (BJC2)	BJC2_MBT	BJC2_MJS	MA dissertation

The naming rules described here are applied to the report of the analysis and findings in the main study.

Four sets of face to face dyad conversation data in the context of 'pedagogic-collaborative idea' (Carter 2004), namely MA and PhD supervision at university, were collected for the main study in the same way as the pilot study.

As shown in Table 4.2.1-2 below, the length of each conversation is about 40 to 60 minutes long. BBC1 and BJC2 are about 40 minutes, and BBC2 is the longest data, which is about 60 minutes. The second longest data is BJC1, which is about 50 minutes. For standardisation of the analysis, the first 39 minutes of data was extracted from each conversation data:

Table 4.2.1-2 The length of the four conversations

	The length of time (HH:MM:SS)	
	Original data	Extracted data
British-British Conversation 1 (BBC1)	00:41:37	00:39:00
British-British Conversation 2 (BBC2)	01:00:27	00:39:00
British-Japanese Conversation 1 (BJC1)	00:48:01	00:39:00
British-Japanese Conversation 1 (BJC2)	00:39:01	00:39:00

Limberg (2007) categorised five phases in academic talks: prefacing, opening, outlining, negotiation and closing, and a study can be designed to analyse the use of response tokens depending on these phases. However, the current study does not focus on these conversation phases since all the five phases may not be present in the data since the data comprises excerpts of longer conversations in order to equalize the length. It is therefore not possible to take those into account here. More data of a particular phase in academic talks will be required if the research design is set to examine listenership behaviour in reference to the conversation phases.

4.2.2 Data modifications – A review

Data modifications in the main study have been conducted based on the procedures developed in the pilot study in Section 3.2.2. Firstly, each participant's utterances and gestures were transcribed and time stamped with a multimodal annotation tool, Transana. The annotated data was exported from Transana and combined with the timeline as a primary key by using Microsoft Excel and Microsoft Access.

A main table from each conversation with timeline was developed through these processes. A sample of the main table from BBC1 is shown in Transcript 4.2-1 below. Each response token used by participants in the conversation is tagged as either listenership (LS) or floor seeker (FS). Although these terms are from studies by O'Keeffe, Carter, McCarthy (McCarthy 2002, O'Keeffe et al. 2007) and Sacks (Sacks 1992), in order to make these definitions simple, I annotated any response tokens used more than 5 seconds before floor-taking as a *listenership* and less than 5 seconds as a *floor seeker*.

Transcript 4.2-1 A sample of the main table from BBC1

Timeline: indicates the time from the beginning of the conversation (HH:MM:SS)

Floor: indicates the timing and who takes the floor

BBC1_MBT_leadtime: BBC1_MBT's leadtime

BBC1_FBS_leadtime: BBC1_FBS's leadtime

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 29: 52		12	-2		There's a journal called metaphor and symbol erm+		
00: 29: 53		13	-1				
00: 29: 54	FBS_F	-8	0			HN	Yeah I found that online I do= I don't +
00: 29: 55		-7	1				
00: 29: 56		-6	2				
00: 29: 57		-5	3		Right yeah.	HG	+ you can actually get hold of it online but
00: 29: 58		-4	4				
00: 29: 59		-3	5				
00: 30: 00		-2	6		Okay.		you have to subscribe to it or something
00: 30: 01		-1	7	SC/hair			
00: 30: 02	MBT_F	0	-9		How annoying. I wonder who owns it. I wonder if Vernon might take it.		
00: 30: 03		1	-8				
00: 30: 04		2	-7				
00: 30: 05		3	-6				
00: 30: 06	Pause	-7	-5		<$E> pause </$E>		
00: 30: 07		-6	-4				
00: 30: 08		-5	-3				
00: 30: 09		-4	-2				
00: 30: 10		-3	-1				
00: 30: 11	FBS_F	-2	0		Yeah.		I'll check again cos it'll say where it's <$G?>.
00: 30: 12		-1	1				
00: 30: 13	MBT_F	0	-53		well it'll be you'll be able to get hold of index for it see if there's anything particularly worth having and if you can't download it online or find it at a university library that's nearby just just order it from the British Library and they'll photoco		

BBC1_MBT_gesture: BBC1_MBT's gestures, such HG and HN

BBC1_MBT_transcript: transcription of BBC1_MBT's utterances.

BBC1_FBS_gesture: BBC1_FBS's gestures, such as HG and HN

BBC1_FBS_transcript: transcription of BBC1_FBS's utterances.

Keys:
FBS_F=female British student's floor-taking, MBT_F=male British teacher's floor-taking,
HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
<$E>pause</$E> = silent pause in conversation, <$G?>= inaudible

As shown above, several steps and procedures needed to be taken to construct this main table. To avoid the redundancy of explanation, only the summary of the data modification has been described in this section. The detail of the modification procedures has been explained in Section 3.2.2.

4.3 Data analysis

4.3.1 Two approaches and their objectives

As described in Section 4.1.3, two approaches were taken in the main study, global pattern analysis and turn structure analysis. Three objectives were set for the global pattern analysis and two objectives were set for the turn structure analysis:

- Global pattern analysis – quantitative analysis
  - Objectives
    1. To summarise the length of speaker status of each participant,
    2. To summarise the numbers of speaker turns of each participant and the average length of speaker turns,
    3. To summarise the numbers of verbal response tokens, such as *erm*, *yeah*, *mm* and *mhm*, and visual response tokens, such as head nods and hand gestures.
- Turn-taking structure analysis – qualitative analysis
  - Objectives
    1. To analyse forms and placements of verbal and visual response tokens in reference to turn-structural episodes and discourse frameworks by integrating the concept of leadtime,
    2. To conduct descriptive analysis on the use of hand gestures with

time-related multimodal transcripts and image captures in reference to multi-functional nature of hand gestures.

To address these objectives, data analysis was conducted quantitatively and qualitatively in the main study. The methods and procedures for the analysis will be discussed in the following sections.

4.3.2 Global pattern analysis

4.3.2.1 Number and length of speaker turns

To measure each participant's speaker status length, only speaker status data needed to be extracted, and leadtime allows us to do this. Speaker status corresponds to leadtimes of greater than zero. Transcript 4.3-1 below is a sample data of BBC1_MBT's speaker status data:

Transcript 4.3-1 A sample extract: BBC1_MBT's speaker status

Timeline_digit	Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_Transcript	BBC1_FBS_Transcript
1	00: 00: 01	MBT_F	0	−2	Go on remind me what you were	Yeah erm.
2	00: 00: 02		1	−1		
7	00: 00: 07	MBT_F	0	−1	W= have you got it there?	
35	00: 00: 35	MBT_F	0	−2	+but you were suggesting the other that you didn't want to do that	+erm.
36	00: 00: 36		1	−1		
66	00: 01: 06	MBT_F	0	−3	Yeah.'Oh right so they're separately interviewed?	+qualities.
67	00: 01: 07		1	−2		
68	00: 01: 08		2	−1	Right.	Yeah.
81	00: 01: 21	MBT_F	0	−15	well it depen= yeah that's not necessarily a problem erm.	
82	00: 01: 22		1	−14		
83	00: 01: 23		2	−13		
84	00: 01: 24		3	−12		No cos.
85	00: 01: 25		4	−11		
86	00: 01: 26		5	−10	Wh= I What what the crucial thing is the sort of circularity of the method so why why were they interviewing the patients were they doing a	

Keys: BBC1_MBT_leadtime=British-British Conversation 1 male British tutor's leadtime,
BBC1_FBS_leadtime= British-British Conversation 1 female British student's leadtime
BBC1_MBT_Transcript=British-British Conversation 1 male British tutor's transcript,
BBC1_FBS_Transcript= British-British Conversation 1 female British student's transcript
MBT_F= Male British tutor taking the floor, FBS_F= Female British student taking the floor

As shown in BBC1_MBT_leadtime in the fourth column, his leadtime has positive positive number values, which means that only BBC1_MBT's speaker status data excluding his listener status data is extracted from the *whole* conversation data. From the extracted data of BBC1_MBT's speaker status, the total length of time of his speaker status in the conversation can be calculated. The numbers of BBC1_MBT's floor-taking can also be acquired by counting an annotated code, 'MBT_F'. This is shown in the third column in the table above. In the sample extract above, BBC1_MBT takes the floor 5 times and the total length of time of his speaker status is 14 seconds. This figure, 14 seconds, is calculated simply by counting the rows of the extracts. These procedures were repeated with the other seven participants in the four conversations. The numbers of floor-taking was also examined in the pilot study with two 10-minute conversation data as shown in Section 3.3.1.1.

Transcript 4.3-2 A pause in BBC1

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 09: 49	MBT_F	0	-32		Yeah. So the the but you're you're prodding the the audience person to to think of it to be persuaded or not.		
00: 09: 50		1	-31	HG			
00: 09: 51		2	-30				
00: 09: 52		3	-29				
00: 09: 53		4	-28	HG			
00: 09: 54		5	-27	HG		HN	
00: 09: 55		6	-26				
00: 09: 56		7	-25			HN	Yeah
00: 09: 57		8	-24				
00: 09: 58		9	-23		Yeah okay.		erm.
00: 09: 59	Pause	-4	-22		<$E> pause </$E>		
00: 10: 00		-3	-21				
00: 10: 01		-2	-20	HN			
00: 10: 02		-1	-19				
00: 10: 03	MBT_F	0	-18		Right. So there's loads of stuff on conceptual metaphor.		

Keys: BBC1_MBT_leadtime=British-British Conversation 1 male British tutor's leadtime,
BBC1_FBS_leadtime= British-British Conversation 1 female British student's leadtime
BBC1_MBT_Transcript=British-British Conversation 1 male British tutor's transcript,
BBC1_FBS_Transcript= British-British Conversation 1 female British student's transcript
HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript 4.3-3 An extract of pauses in BBC1

Timeline	Floor	BJC2_ MBT_le adtime	BJC2_ MJS_le adtime	BJC2_ MBT_g esture	BJC2_MBT_Transcript	BJC2_ MJS_g esture	BJC2_MJS_Transcript
00: 05: 09	Pause	-2	-5		<$E> pause </$E>		
00: 05: 10		-1	-4				Mm.
00: 05: 33	Pause	-7	-11		<$E> pause. </$E>.		
00: 05: 34		-6	-10				
00: 05: 35		-5	-9				
00: 05: 36		-4	-8				
00: 05: 37		-3	-7				
00: 05: 38		-2	-6				
00: 05: 39		-1	-5				
00: 07: 34	Pause	-15	-1		<$E> pause </$E>		
00: 07: 52	Pause	-9	-3		<$E> pause </$E>		
00: 07: 53		-8	-2				
00: 07: 54		-7	-1				
00: 09: 05	Pause	-1	-41		<$E> pause </$E>		
00: 09: 59	Pause	-4	-22		<$E> pause </$E>		
00: 10: 00		-3	-21				
00: 10: 01		-2	-20	HN			
00: 10: 02		-1	-19				

Keys: BBC1_MBT_leadtime=British-British Conversation 1 male British tutor's leadtime,
BBC1_FBS_leadtime= British-British Conversation 1 female British student's leadtime
BBC1_MBT_Transcript=British-British Conversation 1 male British tutor's transcript,
BBC1_FBS_Transcript= British-British Conversation 1 female British student's transcript
<$E>pause</$E>= silent pause in conversation

In terms of pauses in conversation, the total time of pauses in each conversation was also calculated as investigated in the pilot study (see Section 3.3.1.1). Transcript 4.3-2 shows a sample of pauses in BBC1. A pause is observed at 00:09:59 in the timeline, which continues until BBC1_MBT's floor-taking at 00:10:02. Under the condition where both participants' leadtimes were less than 0, the status of silent pauses was extracted as shown in Transcript 4.3-3 above, which shows that both participants' leadtimes in the fourth and fifth columns are negative numbers. This means that both participants share silent pauses in the conversation. With the extracted data of the pauses in the conversation, the total length of the pauses can be measured. In the extract above, for example, the total length of silent pauses is 18 seconds. Again this figure is also acquired by counting the rows in the table. The procedures were repeated with the other three conversation data.

Research has been conducted on silent pauses in English and Japanese

(Hayashi et al. 2002, Maynard 1990, Maynard 1997a, Maynard 1997b). However, allocations of pauses in conversation in relation to turn-taking structure have not been investigated in the previous research. These figures, such as the total length of time of each participant's speaker status and silent pauses, enable us to examine the allocations of the elements in conversation with a concept of time. A comparison was also made between the British-British conversation and the British-Japanese conversation from this perspective.

The results from each conversation are summarised in Table 4.3.3-1 as shown below. The summary of BBC1 indicates that BBC1_MBT's speaker status is 25 minutes 51 seconds in total, while BBC1_FBS's speaker status is 12 minutes 7 seconds. Although BBC1_MBT stays in speaker status more than BBC1_FBS, the number of floor exchanges of each participant is close. BBC1_MBT takes the speaker turns 106 times and BBC1_FBS takes the speaker turns 93 times. The average length of the floor is 14.63 seconds in BBC1_MBT and 7.82 seconds in BBC1_FBS.

Table 4.3.2-1 Summary of lengths and numbers of speaker turns in BBC1

	Speaker status		Speaker turns	
	Length (sec)	Length (HH:MM:SS)	No. turns	Length/turn (sec)
BBC1_MBT	1551	00: 25: 51	106	14.63
BBC1_FBS	727	00: 12: 07	93	7.82
PAUSE	62	00: 01: 02	–	–
Unclassified	0	00: 00: 00	–	–
TOTAL TIME	2340	00: 39: 00	^	^

Keys: BBC1_MBT=British-British Conversation 1 male British tutor,
BBC1_FBS= British-British Conversation 1 female British student

There is 1 minute 2 seconds of pauses in total in BBC1. The same analysis was done with the other three conversation data for comparison and the details in findings from the results are described in the next chapter.

4.3.2.2 Numbers of verbal response tokens

The other issues to be examined in the global pattern analysis were the use of verbal response tokens, such as *erm*, *yeah*, *mm* and *mhm*. Only the transcripts with targeted response tokens were extracted for the analysis.

In Transcript 4.3-4 below, the focus is placed on the use of *yeah* in BBC1_MBT and only the data which includes *yeah* in BBC1_MBT transcript is extracted. In the transcript, *yeah* at 16 seconds in the timeline is uttered -19 seconds in BBC1_MBT's leadtime, which means that 19 seconds before BBC1_MBT's floor-taking, BBC1_MBT utters *yeah*. Using BBC1_MBT's leadtime in the extracted data, the frequency and timing of the response token *yeah* is revealed. Following the procedures developed in the pilot study (see Section 3.3.1.4), the frequency of the use of *yeah* was summarised in five-second time intervals and shown in Table 4.3.2-2 below:

Transcript 4.3-4 Sample extract: *yeah* in BBC1_MBT

Timeline digit	Timeline	Floor	BBC1_MBT leadtime	BBC1_FBS leadtime	BBC1_MBT_Transcript	BBC1_FBS_Transcript
16	00: 00: 16		−19	8	Yeah.	+like health care professional and lay person client patient+
32	00: 00: 32		−3	24	Yeah.	
34	00: 00: 34		−1	26	Yeah+	
66	00: 01: 06	MBT_F	0	−3	Yeah.'Oh right so they're separately interviewed?	+qualities.
81	00: 01: 21	MBT_F	0	−15	well it depen= yeah that's not necessarily a problem erm.	
177	00: 02: 57		−9	6	Yeah.	
184	00: 03: 04		−2	13	Yeah yeah.	+an interviewer+
204	00: 03: 24		−5	9	Right yeah yeah.	
218	00: 03: 38		−24	8	Yeah yeah.	+process it and understand+
219	00: 03: 39		−23	9	Yeah yeah .Yeah yeah.	+metaphor+
223	00: 03: 43		−19	13	Yeah.	
226	00: 03: 46		−16	16	Yeah.	
230	00: 03: 50		−12	20	Yeah.	

Keys: BBC1_MBT_leadtime=British-British Conversation 1 male British tutor's leadtime,
BBC1_FBS_leadtime= British-British Conversation 1 female British student's leadtime
BBC1_MBT_Transcript=British-British Conversation 1 male British tutor's transcript,
BBC1_FBS_Transcript= British-British Conversation 1 female British student's transcript
MBT_F= Male British tutor taking the floor, FBS_F= Female British student taking the floor
"+"= describe the continuous of the sentence, "="= unfinished sentence, <$G?>= inaudible sounds

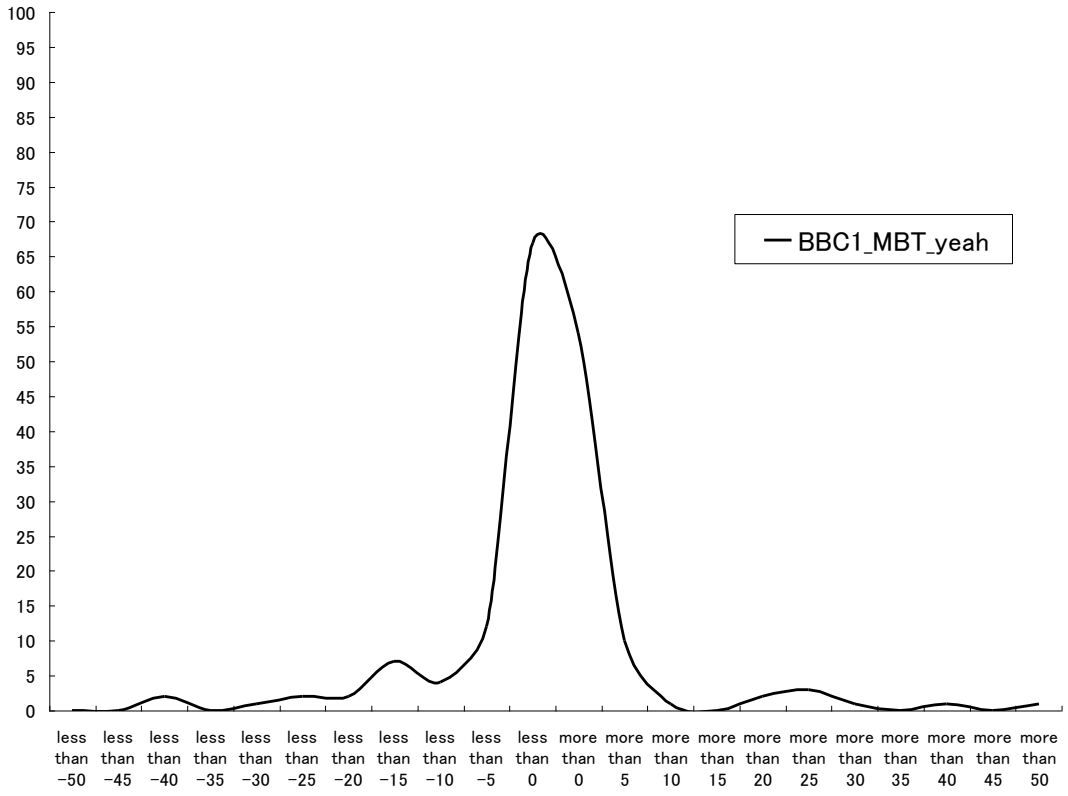
Table 4.3.2-2 Numbers of *yeah* in BBC1_MBT

leadtime	BBC1_MBT_yeah
less than -50	0
less than -45	0
less than -40	2
less than -35	0
less than -30	1
less than -25	2
less than -20	2
less than -15	7
less than -10	4
less than -5	12
less than 0	67
more than 0	54
more than 5	10
more than 10	1
more than 15	0
more than 20	2
more than 25	3
more than 30	1
more than 35	0
more than 40	1
more than 45	0
more than 50	1
	170

Keys: BBC1_MBT=British-British Conversation 1 male British tutor

The table above shows that the numbers of *yeah* at 5 to 1 seconds before floor-taking is 69. BBC1_MBT's use of *yeah* in the next interval, 0 to 4 seconds after floor-taking, is also more than 50. From these figures, it can be interpreted that BBC1_MBT has used *yeah* 5 seconds before and after he takes the floor.

The frequency of the use of *yeah* in the table above was transferred into a graph in order to illustrate the tendency of the use of *yeah* in relation to leadtime visually. Figure 4.3-1 below shows the results of the use of *yeah* in BBC1_MBT. Although the timing where the response token *yeah* is most frequently used in BBC1_MBT is around TTP, he also uses *yeah* when he is in listener status and even in speaker status:



Keys: BBC1_MBT=British-British Conversation 1 male British tutor

Figure 4.3-1 The numbers of *yeah* in BBC1_MBT

All the eight participants' use of the four selected verbal response tokens, *erm*, *yeah*, *mm* and *mhm*, were counted by numbers in relation to the leadtime and compared with each other. Tables and graphs of each participant's use of targeted response tokens were acquired for analysis and the findings are discussed in the next chapter.

4.3.2.3 Numbers of visual response tokens

As examined in the pilot study in Section 3.1.3.3, two targeted visual response tokens, hand gestures and head nods, were counted in relation to leadtime and summarised in tables and figures in the main study. Any hand movements are counted as hand gestures and any vertical head movements are treated as head nods. These two visual response tokens were transcribed and time-stamped by an annotation tool, Transana,

and verbal and visual transcriptions were combined using Microsoft Access. These two visual response tokens in the four 39-minute length conversation data have been annotated and counted in five-second time intervals.

Transcript 4.3-5 Sample transcript from BBC1

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 17: 05	FBS_F	-13	0			HG	It's just too difficult for it.
00: 17: 06		-12	1				
00: 17: 07		-11	2		+literally erm.	HN	
00: 17: 08		-10	3			HG	Yeah the stuff I've seen it's just two word erm+
00: 17: 09		-9	4			HG	
00: 17: 10		-8	5				
00: 17: 11		-7	6	HN	Right.		
00: 17: 12		-6	7		Right yeah yeah.		+metaphors really.
00: 17: 13		-5	8				
00: 17: 14		-4	9				So I mean you get like grammatical metaphors with+
00: 17: 15		-3	10			HG	
00: 17: 16		-2	11	HN	Yeah yeah yeah.		+verbs and stuff.
00: 17: 17		-1	12				
00: 17: 18	MBT_F	0	-3		Yeah that's still clever if you can do that.	HG	

Keys: BBC1 = British-British Conversation 1, BBC1_MBT=British-British Conversation 1 male British tutor
BBC1_FBS=British-British Conversation 1 female British student
FBS_F = Female British Student's taking the floor, MBT_F = Male British tutor's taking the floor
HN = Head nods, HG = Hand gestures

In the fifth column of BBC1_MBT_gesture in Transcript 4.3-5 above, head nods are observed in the timeline 00:17:11 and 00:17:16. In the column of BBC1_FBS_gesture, hand gestures are observed while she is in speaker status. As processed in the pilot study, hand gestures and head nods have also been counted in five-second time intervals in the main study. These values are shown in Table 4.3.2-3 and Figure 4.3-2 below. The table below indicates that BBC1_MBT has used 523 hand gestures in total in the 39-minute conversation data and about half of them are used within 10 seconds after he takes the floor:

Table 4.3.2-3 BBC1_MBT_HG

leadtime	BBC1_MBT_HG
less than -50	0
less than -45	0
less than -40	0
less than -35	0
less than -30	0
less than -25	0
less than -20	0
less than -15	0
less than -10	1
less than -5	1
less than 0	11
more than 0	138
more than 5	89
more than 10	76
more than 15	51
more than 20	32
more than 25	26
more than 30	30
more than 35	16
more than 40	10
more than 45	9
more than 50	33
	523

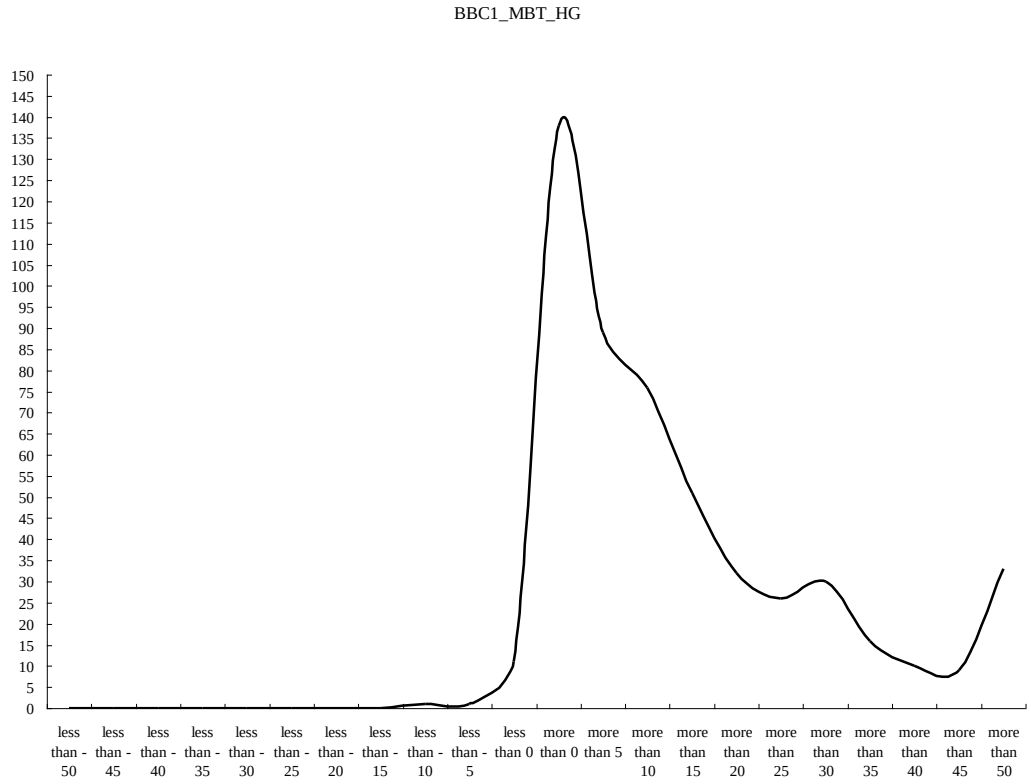


Figure 4.3-2 BBC1_MBT_HG

The graph of BBC1_MBT_HG above also illustrates that the numbers of HGs increases sharply around TTP, which is 0 in leadtime on the X axis. About 10 seconds after he takes the floor, the number of HGs decreases gradually; however, HGs are also used in his speaker status. This tendency supports the findings from the pilot study described in Section 3.3.1.3.

Some existing studies on gestures have also explored functions of gestures in relation to turn-taking structures (Kendon 1972, Knight et al. 2006). However, introducing time-based transcription in research on gestures is a new direction. Placements of HGs and HNs might be related to particular functions of visual response tokens in conversation. Further analysis and discussion on the use of HGs and HNs is given in the next chapter.

4.3.3 Turn structure analysis

Based on O'Keeffe and Adolphs (Knight et al. 2006, O'Keeffe & Adolphs 2008, O'Keeffe et al. 2007), forms and functions of response tokens were analysed in relation to turn-taking structure in the quantitative and descriptive analysis. In terms of forms of response tokens, three types of forms were outlined:

> (1) Minimal response tokens: Short utterances or non-word vocalisations (*yeah*, *mm*)
>
> (2) Non-minimal response tokens: Adverbs and adjectives or short phrases/minimal clauses (*good*, *really*, *is that so?*)
>
> (3) Clustering of response tokens: Both minimal and non-minimal response tokens can occur in pairs or clusters (*yeah mm*, *right fine*)
>
> (O'Keeffe et al. 2007: 143-144)

A clustering of response tokens functions 'to signal a boundary "and" to add satisfaction or agreement or simply to express friendly social support' (ibid: 144, original author's emphasis). These functions have been explored by some researchers (Gardner 1998, Gardner 2002, Maynard 1990), and, currently, four functions of response tokens are focused on:

> (1) Continuers [CN]: Maintaining the flow of discourse.
>
> (2) Convergence tokens [CV]: Markers of agreement/convergence.
>
> (3) Engaged response tokens [EN]: Markers of high engagement where addressee(s) respond on an affective level to the content of the message.
>
> (4) Information receipt tokens [IR]: Markers of points in the discourse where adequate information has been received.
>
> (O'Keeffe & Adolphs 2008 : 84)

The forms and functions of response tokens were analysed in reference to turn-structural episodes introduced by Ohama's (2006) study on a turn shifting mechanism in Japanese conversation (see Table 4.3.3-1 below). This is revised in Section 4.1.6, based on the categorisation developed in the pilot study.

Table 4.3.3-1 Turn-structural episodes

Episode 1	A's turn closing → B's turn-taking
Episode 2	A's turn keeping → B's cut-in
Episode 3	A's turn closing → B's turn refusal → A's turn retaining
Episode 4	A's turn closing → B's turn refusal → A's turn refusal → B's final turn-taking
Episode 5	A's turn giving → B's turn-taking
Episode 6	A's turn giving → B's turn refusal → A's turn retaining
Episode 7	A's turn giving → B's turn refusal → A's turn refusal → B's final turn-taking

(Adapted from Ohama 2006)

The turn-structural analysis was divided into four sections according to the aspects to be focused on:

Aspect 1: verbal response tokens

Aspect 2: verbal response tokens with head nods

Aspect 3: hand gestures

Aspect 4: Turn-structural episodes

The categorisation of floor-taking patterns developed in the pilot study was also applied to the main study. The findings about these four aspects from both quantitative analysis and descriptive analysis are described in the next chapter.

4.3.3.1 Aspect 1: Focusing on verbal response tokens

Functions and forms of verbal response tokens were investigated in relation to the timeline. Three forms of response tokens based on O'Keeffe et al's classifications (2007): minimal response tokens, non-minimal response tokens and clusters, were analysed both numerically and descriptively.

In Transcript 4.3-6 below, for instance, *Right* at 00:02:51 and *Yeah* at 00:02:57 in the timeline are recognised as minimal responses. According to BBC1_MBT's leadtime in the fifth column, those minimal response tokens are uttered more than 5 seconds before the next floor-taking at 00:03:06. *That's right* at 00:03:03 in the timeline is a non-minimal response token followed by *yeah yeah* at 00:03:04 which is a cluster of minimal responses:

Transcript 4.3-6 Sample transcript: Verbal response tokens

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_Transcript	BBC1_FBS_Transcript
00: 02: 51	FBS_F	-15	0	Right.	+at all. Erm and I think it does change the pragmatics perspective of it as well because+
00: 02: 52		-14	1		
00: 02: 53		-13	2		
00: 02: 54		-12	3		
00: 02: 55		-11	4		
00: 02: 56		-10	5		
00: 02: 57		-9	6	Yeah.	
00: 02: 58		-8	7		+erm they're not trying to er come to a shared understanding of something they are explaining it to+
00: 02: 59		-7	8		
00: 03: 00		-6	9		
00: 03: 01		-5	10		
00: 03: 02		-4	11		
00: 03: 03		-3	12	That's right.	
00: 03: 04		-2	13	Yeah yeah.	+an interviewer+
00: 03: 05		-1	14		So.
00: 03: 06	MBT_F	0	-9	Well it's shared but+	

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC1_FBS=British-British Conversation 1 female British student
FBS_F = Female British Student's taking the floor, MBT_F = Male British tutor's taking the floor
BBC1_MBT_CF = BBC1 Male British tutor's conversational functions

The first two minimal response tokens, *right* and *yeah*, can be interpreted as continuers and the last two response tokens, *that's right* and *yeah yeah* can be seen as convergence response tokens to agree with the current speaker. At the same time, the last two response tokens can function as a sign for a boundary of turn exchange. The nature of multi-functionality in response tokens is one of the important issues considered in the qualitative analysis.

By combining research methods from corpus linguistics (Adolphs 2008, Carter & McCarthy 2006), discourse analysis (Carter & McCarthy 1997, McCarthy 2002, McCarthy et al. 2002) and conversation analysis (Heritage 1984a, Heritage 1984b, Sacks 1992), sequences of listenership and speakership were also examined. The relationship between forms and functions of response tokens and listener's transactional goals after taking the floor were taken into consideration. In the

transcript above, for example, a transactional goal of BBC1_MBT's floor-taking at 00:03:06 can be taken as a challenge to BBC1_MBS's previous utterance and offering suggestions on how to deal with her research data. Strategies and patterns of response tokens before their floor-taking with particular transactional goals, such as challenge and expansion, were examined in the current research. The findings are reported in the next chapter.

4.3.3.2 Aspect 2: Focusing on verbal response tokens and head nods

In the pilot study in Section 3.3, the use of visual response tokens, head nods and hand gestures, were analysed. In the main study, collocations of head nods with verbal response tokens were also examined qualitatively in relation to the timeline and turn-taking patterns as described in the previous section.

Transcript 4.3-7 Sample transcript: Head nods with verbal response tokens

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 02: 51	FBS_F	-15	0		Right.		+at all. Erm and I think it does change the pragmatics perspective of it as well because+
00: 02: 52		-14	1				
00: 02: 53		-13	2				
00: 02: 54		-12	3				
00: 02: 55		-11	4				
00: 02: 56		-10	5	HN			
00: 02: 57		-9	6		Yeah.	HG	
00: 02: 58		-8	7				+erm they're not trying to er come to a shared understanding of something they are explaining it to+
00: 02: 59		-7	8				
00: 03: 00		-6	9	HN		HG	
00: 03: 01		-5	10			HG	
00: 03: 02		-4	11	HN			
00: 03: 03		-3	12		That's right.	HG	
00: 03: 04		-2	13		Yeah yeah.		+an interviewer+
00: 03: 05		-1	14				So.
00: 03: 06	MBT_F	0	-9	HG	Well it's shared but+		

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC1_FBS=British-British Conversation 1 female British student
FBS_F = Female British Student's taking the floor, MBT_F = Male British tutor's taking the floor
BBC1_MBT_CF = BBC1 Male British tutor's conversational functions

In Transcript 4.3-7 above, *yeah* at time 00:02:57 follows a head nod at 00:02:56, which can be interpreted as a continuer. Two head nods are observed before the non-minimal response token *that's right* at 00:03:03, which is followed by the cluster *yeah yeah*. These 2 head nods before floor-taking might signal a boundary of turn exchange as with these 2 verbal response tokens. For example, within 6 seconds before BBC1_MBT's floor-taking, 4 response tokens were observed. However, it might be worth noticing that no response token is used 1 second before the speaker takes the floor, and, without overlap, the floor is smoothly moved from BBC1_MBS to BBC1_MBT at 00:03:06. Co-occurrences of head nods with verbal response tokens will be analysed and compared among the participants of the four conversations.

4.3.3.3 Aspect 3: Focusing on hand gestures

In the sample transcription above, hand gestures have also been observed at BBC1_MBT's TTP at 00:03:06. Adapting the method in Carroll (2004), the image was also used for the analysis. Figure 4.3-3 below, for example, describes the movement of hand gestures that occurs at TTP of BBC1_MBT with a discourse marker *well*. At the moment when BBC1_MBT takes the floor, he moves his right hand with his palm open and draws a small circle while saying *well*. This finding supports the observation by Kendon (1972), which reports that body movements can occur *before* speech by introducing the term *speech-preparatory* movement (ibid: 205). At the same time, BBC1_FBS stops her speech and changes her status from speaker to listener. Based on the findings in the pilot study in Section 3.3, the use of hand gestures around a boundary from listener to speaker were examined in the quantitative analysis.

Figure 4.3-3 Sample image: Hand gestures

In the case described above, BBC1_MBT utters 4 response tokens and makes 3 head nods before he takes the floor. At the TTP, he uses a hand gesture accompanied by the discourse marker *well*. In the quantitative analysis, the use of hand gestures in reference to turn structure was examined and a comparison was made among participants. Turn-structural episodes will be described in detail in the next section.

4.3.3.4 Aspect 4: Turn-structural Episodes

As described in Section 4.1.6, seven turn-structural episodes were established based on the pilot study and Ohama's (2006) study (see Table 4.3.3-2 below).

Table 4.3.3-2 Turn-structural episodes

Episode 1	A's turn closing → B's turn-taking
Episode 2	A's turn keeping → B's cut-in
Episode 3	A's turn closing → B's turn refusal → A's turn retaining
Episode 4	A's turn closing → B's turn refusal → A's turn refusal → B's final turn-taking
Episode 5	A's turn giving → B's turn-taking
Episode 6	A's turn giving → B's turn refusal → A's turn retaining
Episode 7	A's turn giving → B's turn refusal → A's turn refusal → B's final turn-taking

(Adapted from Ohama 2006)

As with the pilot study in Section 3.3.2.5, the numbers of these seven turn-taking patterns in each participant's speech were counted and a comparison in the use of these patterns was made in the main study. Means, standard deviations, and variances of listener status in each pattern were calculated. The relationship between the use of verbal and visual response tokens and the turn-taking patterns was investigated.

As reviewed in Section 2.1.3.3, patterns of *discourse sequences* (Baker et al. 2001) and *framework of discourse* (Saft 2007) were concerned in the turn-structural analysis. Saft (2007) pointed out that there are two discourse frameworks observed in a Japanese faculty meeting at university, namely a *reporting framework* and a *discussion framework,* and the chair's use of response tokens leads a transition from a reporting framework to a discussion framework.

In academic tutorials, there also seem to be two frameworks, which I shall refer to as a *commentary framework* and an *explanatory framework*. Commentary frameworks are often used by the tutors to give comments and suggestions to the students, while explanatory frameworks are observed in the students where they are elaborating and explaining their ideas. These two frameworks were applied to the analysis in the main study.

4.4 Summary

The purpose of this chapter was to make a rational link between the pilot study and the main study and to establish the research methods applied to the main study. The key concepts implemented in the current study, such as leadtime and TTP, and discourse frameworks, have been reviewed. As a bridge to the main study, the details of the research data and methods of the analysis in the main study have also been described. The findings from the global pattern analysis and turn structure analysis in

the main study will be reported in the next chapter.

Chapter 5 Results: Main study

5.0 Introduction

In this chapter, findings from the main study will be described. This chapter opens with the findings from the global pattern analysis and moves on to discuss results from the turn structure analysis.

In the first part, occurrences of vocal and visual response tokens are counted and reported in detail with tables and figures in order to provide an overview of the use of response tokens in conversation in academic tutorials based on the pilot study in Section 3.3. The functions and forms of participants' response tokens will be analysed qualitatively in relation to turn-taking patterns in the second part of this chapter.

For the scalability of the research, findings from the main study need to be compared with the findings from the pilot study. In the main study, four sets of conversation data of 39-minute length each will be analysed. The aim of the main study is not to validate the results from the pilot study, rather to develop a model of conversation analysis by applying the model developed through the pilot study to a set of larger data. Discussions and implications of the research together with underpinning theories will also be discussed here.

5.1 Global pattern analysis

In the global pattern analysis, the four sets of conversation data of 39-minute length each were analysed quantitatively based on the methods developed in the pilot study. There were four salient findings reported from the quantitative analysis in the pilot study with two 10-minute length conversation data in Section 3.3.1:

1. In the British-British conversation, the numbers of speaker turns and backchannel turns in conversation were more equal to each participant than the British-Japanese conversation.
2. In the British-British conversation, HGs (hand gestures) were observed more often at TTP. The male Japanese student in the British-Japanese conversation rarely used HGs and the female British tutor in the British-Japanese conversation used HGs continuously while she took the floor of the conversation.
3. The female British student in the British-British conversation used *erm* at TTP several times. The male British tutor in the British-British conversation used *yeah* as a strategy at TTP.
4. In the British-Japanese conversation, the male Japanese student used *mm* 50 times in the 10-minute conversation data, and he also used *mm* constantly when he was in listener status. Conversely, the female British tutor in the British-Japanese conversation did not use *mm* at all. The male British tutor and the female British student rarely used *mm* in the British-British conversation.

These findings from the pilot study were reviewed in the main study to enable comparison. The features of turn structure and the use of response tokens recognised in the pilot study were also focused on in the main study.

5.1.1.1 Number of words uttered

Numbers of words spoken by each participant in the four conversations were counted and compared with the results from the pilot study.

Table 5.1.1.1-1 Conversation data for the pilot study

	Data Length (HH:MM:SS)	Number of words		
		Tutor	Student	Total
C1 (British-British Conversation)	00:10:00	1086	946	2032
C2 (British-Japanese Conversation)	00:10:00	909	100	1009

Table 5.1.1.1-2 Conversation data for global pattern analysis in the main study

	Data Length (HH:MM:SS)	Number of words		
		Tutor	Student	Total
BBC1 (British-British Conversation 1)	00:39:00	5399	2536	7935
BBC2 (British-British Conversation 2)	00:39:00	3301	2489	5790
BCJ1 (British-Japanese Conversation 1)	00:39:00	3158	508	3666
BJC2 (British-Japanese Conversation 2)	00:39:00	4431	654	5085

In the pilot study, 10-minute length data of a British-British conversation and a British-Japanese conversation were analysed. Word counts of the 10-minute pilot study data were about 2,000 in total in the British-British Conversation and about 1,000 in the British-Japanese conversation. In the global pattern analysis in the main study, four sets of conversation data of 39-minute length each were examined. The numbers of words in the British-British conversations were about 5,800 to 8,000 while the British-Japanese conversations had about 3,700 to 5,000 words in total. From this information generated by the pilot study, it can be said that the numbers of words uttered in the British-British conversation is larger than the British-Japanese conversation. Another way of describing this is that, in terms of the numbers of words uttered, the Japanese students tended to contribute to conversation with shorter verbal utterances than the British students, though because this was an instance from a small

data set, generalisations cannot be made.

For instance, the total words in BJC2 was about 5,000, which was close to BBC2; however, BJC2_MJS's number of words uttered was only 654 words while BJC2_MBT uttered 4,431 words. In contrast, BBC2_MBS uttered about 2,500 words and BBC2_MBT uttered only about 1,000 more than the student. As with the case of BJC2, BJC1_MJS spoke only 500 words in the 39-minute length conversation while BJC1_FBT uttered more than 3,000 words in total. In summary, it can be said that the students tended to utter fewer words than the tutors. Compared with the cases of the British-British conversations, inequality of participation between the tutors and the students was greater in the British-Japanese conversations. Again, this is an observation from the small data set with only eight participants, so generalisations cannot be made. With the method established from the current study, however, these features were able to be described.

5.1.2 Number and length of speaker turns

Numbers of speaker turns and length of speaker status of each participant in the four conversation data were examined in the main study. The male British tutor (BBC1_MBT) and the female British student (BBC1_FBS) had a tutorial about her MA dissertation in the British-British Conversation 1 (BBC1). The length of the data was cut into a 39-minute extract. During the tutorial, BBC1_MBT took speaker turns 106 times and BBC1_FBS took speaker turns 93 times as shown in Table 5.1.2-1 below. The tendency for the tutor to take the floor more than the student is observed in BBC1. BBC1 shows one type of conversation style, where the participants exchange shorter turns frequently.

Table 5.1.2-1 BBC1 Summary: No. and length of speaker turns

	Speaker status		Speaker turns	
	Length (sec)	Length (HH:MM:SS)	No. turns	Length/turn (sec)
BBC1_MBT	1551	00: 25: 51	106	14.63
BBC1_FBS	727	00: 12: 07	93	7.82
PAUSE	62	00: 01: 02	–	–
Unclassified	0	00: 00: 00	–	–
TOTAL TIME	2340	00: 39: 00	^	^

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC1_FBS=British-British Conversation 1 female British student

As shown in the table above, the average of BBC1_MBT's speaker turns is 14.63 seconds while BBC1_FBS's speaker status length is about half of BBC1_MBT. Further, BBC1_MBT's speaker status length in total is about 26 minutes, which is more than double that of BBC1_FBS. This supports the observation from the pilot study that tutors were longer in speaker status than students.

BBC2 had a different conversation style in terms of the length of floor-taking as described in Table B.1.1-1 in the appendix. BBC2_MBT took 36 speaker turns, which is a third of BBC1_MBT, while BBC2_MBS took the speaker turns 34 times, which again is about a third of BBC1_FBS. Compared with BBC1, fewer numbers of floor exchanges and longer speaker turns were outstanding features in BBC2. BBC2_MBT's average length of speaker turn was about 43 seconds, which is about three times longer than BBC1_MBT. BBC2_MBS's average speaker turn length was about 23 seconds long, which also about three times longer than BBC1_FBS. Although there seem to be some differences in conversation style between BBC1 and BBC2, BBC2_MBT takes speaker turns more than BBC2_MBS and he is in speaker status longer than BBC2_MBS. BBC2_MBT's total speaker status length is about 26 minutes, which is almost the same as BBC1_MBT. BBC2_MBS's total speaker status length, alternately, is about 13 minutes, which is also close to BBC1_FBS. These are

the similarities between BBC1 and BBC2.

BJC1_FBT took the speaker turns 119 times during the tutorial while BJC1_MJS took only 24 speaker turns as shown in Table B.1.1-2 in the appendix. BJC1_FBT is in speaker status for about 22 minutes in total while BJC1_MBS is in speaker status only for 3 minutes in total. The total length of pauses in the conversation is about 11 minutes 30 seconds, which distinguishes BJC1 from the British-British conversations.

In terms of conversation style, however, some similarities were observed between the British-British conversations and the British-Japanese conversations. From the analysis of the British-British conversations, two conversational styles were recognised in terms of the average length of participants' speaker turns: (1) a shorter turn conversation; and (2) a longer turn conversation. BBC1 was categorised as a shorter turn conversation, where the tutor had about 13-second long speaker turns and the student had about 7-second long speaker turns. In contrast, BBC2 was categorised as a longer turn conversation, where the tutor had about 40 to 50-second long speaker turns and the student had about 20-second long speaker turns. BBC1 and BJC1 seemed to share the first conversational style, shorter turn exchanges, although there were obvious differences in the numbers of floor exchanges and amount of silent pauses between BBC1 and BJC1. BJC1_FBT's average speaker turn length was about 12 seconds, which is close to BBC1_MBT; whereas BJC1_MJS's average speaker turn was about 7 seconds, which is again almost the same as BBC1_FBS. In addition, the fourth conversation data BJC2 had similarities with BBC2 in relation to the conversation style. Both BBC2 and BJC2 can be categorised as longer turn conversation.

As shown in Table B.1.1-3 below, the average speaker turn length of BJC2_MBT is about 54 seconds, which is close to BBC2_MBT's average speaker turn length; whereas BJC2_MJS's average speaker turn length is about 19 seconds, which is similar to BBC2_MBS. Another point to be noted is that the total length of silent pauses in BJC2 is about 1 minute. This is another similarity between BJC2 and the British-British conversations and a difference from BJC1. From this it can be said that more pauses were observed in the British-Japanese conversations than the British-British conversations; and this was truer in BJC1 than in BJC2. The amount of silent pauses was not necessarily a feature of British-Japanese conversations, although placement of pauses might be worth investigating further in relation to turn-taking structure.

Table B.1.1-3 in the appendix shows that BJC2_MBT takes the floor 32 times and BJC2_MJS takes the floor 24 times. This indicates a similar tendency to the British-British conversations, where participants take the speaker turns more equally although BJC2_MBT's total speaker status length is about 30 minutes, which is four times more than BJC2_MJS.

As a summary of the section, the numbers of speaker turns in each participant in the four conversations are described in Table 5.1.2-2 below:

Table 5.1.2-2 Four conversation data: No. and length of speaker turns

	Length	No. of speaker turns		
	(HH:MM:SS)	Teacher	Student	Total
BBC1	00:39:00	106	93	199
BBC2	00:39:00	37	35	72
BJC1	00:39:00	114	24	138
BJC2	00:39:00	32	24	56

Keys: BBC1 = British-British Conversation 1, BBC2 = British-British Conversation 1,
BJC1 = British-Japanese Conversation 1, BJC2 = British-Japanese Conversation 1

Chiasson and Hayes (1993) conducted an experimental study on three types of dyad conversations at university: (1) two college freshmen; (2) a freshman with a senior; and (3) a freshman with a graduate student. The conversations recorded were relational communication, in which the pairs were asked to talk about television programmes. Their study reported that 'freshmen initiated almost twice as many interactions and spent almost twice as much time talking to students of the same status, as compared to those of different status' (Chiasson & Hayes 1993: 13) .

The current research also reflects that age differences and power relationships between the tutors and the students may affect the numbers of turn exchanges and the length of speaker status in conversation. The participants' information is described in Table E1-2.1 in the appendix. Although any generalisation cannot be made from the analysis on the small data sets, in the conversations where differences in their age and power relationships were smaller, such as BBC2 and BJC2, longer turn exchanges were observed, and the numbers of turn exchange were more equal.

5.1.3 Findings about the use of head nods and hand gestures

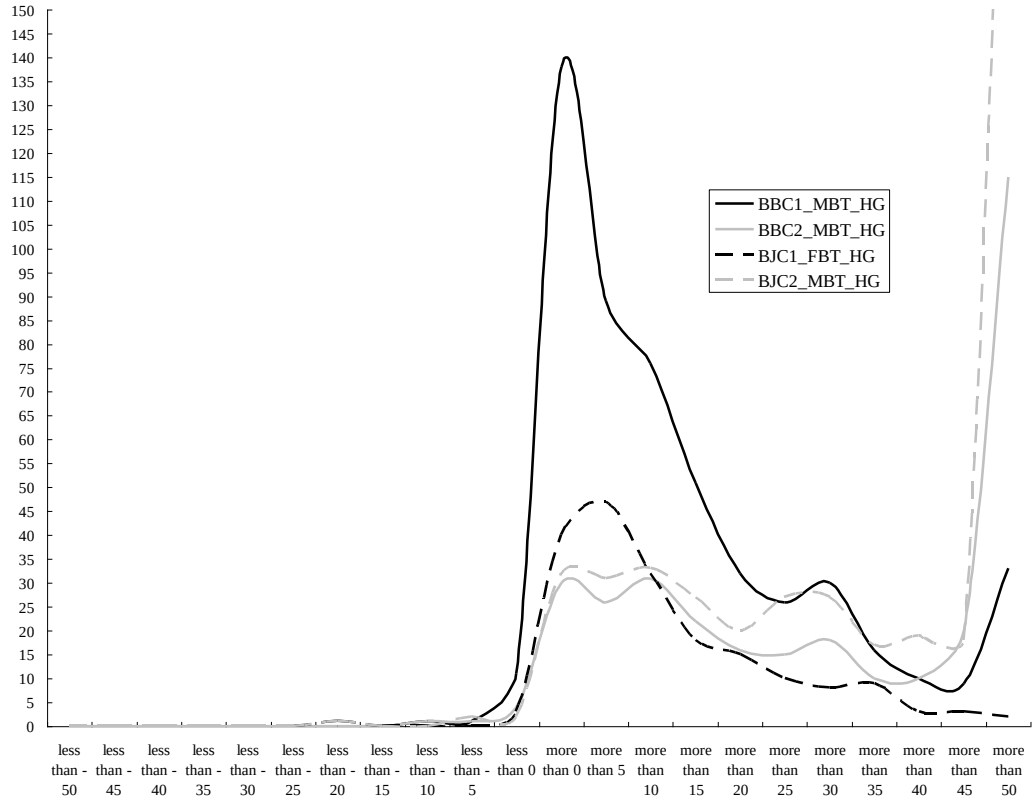
The use of head nods and hand gestures of each participant in the four conversation data were counted and analysed in reference to leadtime. As described in Section 3.1.7, hand gestures were defined as any hand movements observed in the current research, and the time spaces between these movements were counted for the analysis. Several hand gestures that occurred within one second were counted as one hand gesture since a microanalysis of response tokens with smaller time scales, such as in tenth of seconds or milliseconds, was difficult to handle practically at this stage. Time stamps for response tokens were added using the annotation software system manually in the current study, and the same annotation rules were applied for coding

head nods which were classified as any vertical head movements. Head nods that occurred several times within one second were also counted as one head nod in the current study.

5.1.3.1 Placement of hand gestures

As the table and the figure below show, the tutors' use of HGs was observed at TTP and their speaker status:

1. The tutors often used HGs soon after TTP.
2. The tutors used HGs during their speaker status.



Keys: BBC1_MBT_HG = British-British Conversation 1 male British tutor's hand gestures,
BBC2_MBT_HG = British-British Conversation 2 male British tutor's hand gestures
BJC1_FBT_HG = British-Japanese Conversation 1 female British tutor's hand gestures,
BJC2_MBT_HG = British-Japanese Conversation 2 male British tutor's hand gestures

Figure 5.1-1 Tutors' use of hand gestures

The tutors' use of HGs was counted in five-second time intervals as shown in Figure 5.1-1 above (also see Table B.1.2-1 in the appendix). The tutors tended to use hand gestures at TTP and during their speaker status. BBC1_MBT uses HGs more than 500 times in total and BJC2_MBT uses HGs about 490 times, and these latter two results are outstanding numerically. BBC2_MBT also uses HGs more than 300 times in total and BBC2_MBT and BJC1_FBT uses HGs about 200 times. The graph illustrates that BBC2_MBT's use of HGs and BJC2_MBT's use of HGs increase even more than 50 seconds after their floor-taking. These figures do not mean that BBC2_MBT and BJC2_MBT used hand gestures between 'more than 50' and 'more than 55'; rather, that all of the hand gestures used more than 50 seconds after their floor-taking were counted into an interval of 'more than 50'. When BBC2_MBT used hand gestures at 60 seconds, for instance, it was counted into the interval of 'more than 50'. As examined in the previous section, the speaker status length of BBC2_MBT and BJC2_MBT were longer than the other two tutors. This can be one of the reasons why these two tutors' use of HGs was distributed over their longer speaker status.

As the table and the figure of students' use of HGs in the appendix (see Table B.1.2-2 and Figure B.1-1) indicate, there were similarities in the placements of HGs between the tutors and the students. The students also used HGs at TTP more frequently and the numbers of HGs declined until about 25 seconds after their floor-taking. The students' use of HGs in their speaker status was comparatively less than the tutors in numbers. Again, this contrast might be related to a tendency that the students are prone to have the shorter floor than the tutors so that the students' use of HGs in speaker status can be limited.

Although the numbers of the Japanese students' use of HGs were fewer than

the British students, the same tendency in placements of HGs as the British students was observed. BJC1_MJS used hand gestures only 44 times and BJC2_MJS about 160, while the British students used hand gestures more than 200 times in total. This might be related to the numbers and the length of their floor-taking. These features of the students' use of HGs is summarised as follows:

1. Both the British students and the Japanese students often used HGs soon after TTP.
2. Both the British students and the Japanese students used HGs in their speaker status but less than the tutors.

In summary, placement of HGs soon after TTP was also observed in the results from the main study. The same results were reported in the pilot study (see Section 3.3.1.3).

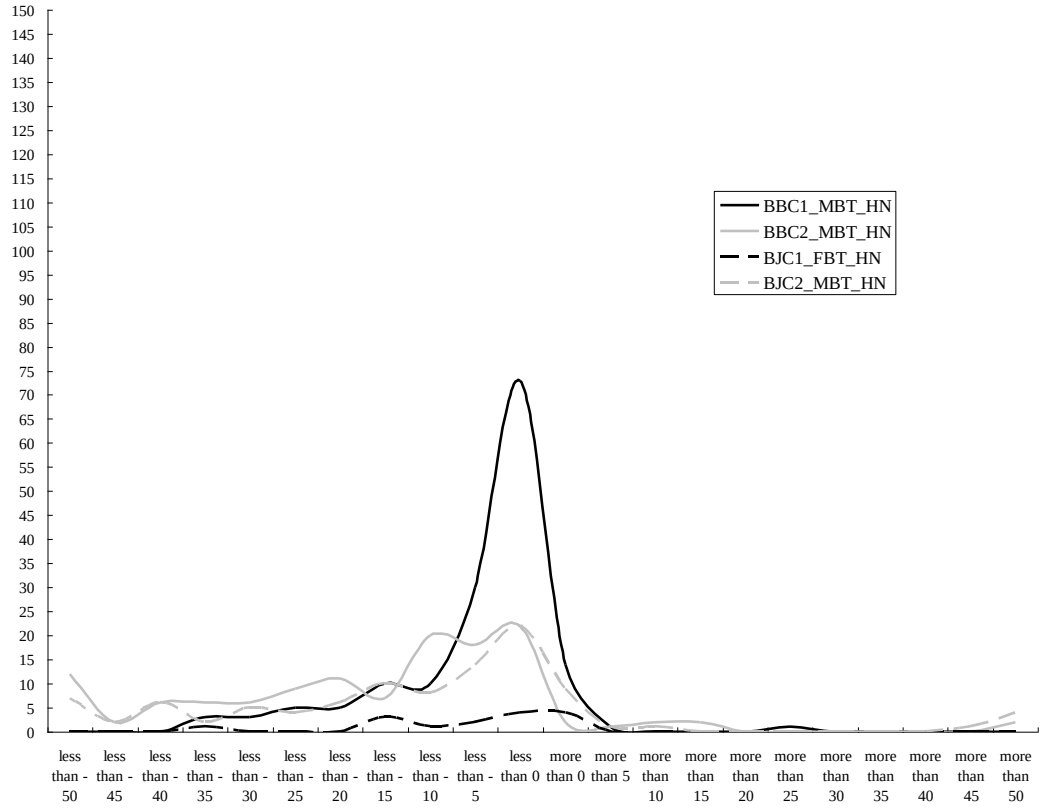
In terms of the relationship between placements and functions of HGs, as studied in Section 2.1.5, these two placements of HGs, namely at TTP and in speaker status, might be related to different functions of conversational gestures: the function to express 'unspoken thoughts' (Goldin-Meadow 1999), and the function related to turn-taking (Kendon 1972, Schegloff 1984). Some HGs might be used to fulfil both of these two purposes. Although functions of HGs cannot be generalised at this early stage, it can be worth noting that HGs can be used as signals of intention to initiate the turn or to secure the floor, and at the same time, HGs can help speakers to convey their ideas and thoughts to listeners.

5.1.3.2 Placement of head nods

The tutors' use of HNs is illustrated in the figure below (also see Table B.1.3-1).

Three of the tutors use HNs more than 100 times in the 39-minute length conversations while BJC1_FBT uses HNs only 15 times in total. In most occurrences, the leadtime of HNs are about 10 seconds before floor-taking. These tendencies were summarised as follows:

1. The tutors often used HNs soon before TTP.
2. The tutors used HNs in their listener status.
3. The tutors used HNs in their speaker status few times.



Keys: BBC1_MBT_HN = British-British Conversation 1 male British tutor's head nods,
BBC2_MBT_HN = British-British Conversation 2 male British tutors' head nods
BJC1_FBT_HN = British-Japanese Conversation 1 female British tutor's head nods,
BJC2_MBT_HN = British-Japanese Conversation 2 male British tutor's head nods

Figure 5.1-2 Tutors' use of head nods

Although the case of BJC1_FBT was different from the others numerically, still the

placements of HNs fit to the descriptions above. BJC1_FBT uses many aspects of visual response tokens. Further analysis on BJC1_FBT's use of HNs is given in the quantitative analysis in the next chapter.

As reviewed in Section 2.1.5, functions of HNs can be more related to listenership than to speakership, and, in addition, the placements of HNs might also suggest a multi-functional nature of response tokens; for HNs can function as continuers, convergent tokens, engaged response tokens and information receipt tokens (Knight et al. 2006). In addition, HNs before TTP might signal a listener's intention to take the next floor. Quantitative and descriptive analyses are conducted on this issue in the next chapter.

Some HNs in speaker status were also observed in the tutors' use of HNs. As shown in Transcript 5.1-1 below, BBC2_MBT uses HNs at timeline 00:27:33 while he is in speaker status. That is followed by BBC2_MBS's utterance *yeah*:

Transcript 5.1-1 Head nods in speaker status

Timeline	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 27: 30	37	-46		Particularly since it's something which is novel+		
00: 27: 31	38	-45				
00: 27: 32	39	-44			HN	
00: 27: 33	40	-43	HN	+that you you can set out your rationale for how the corpora's clean+		Yeah.

Keys: BBC2_MBT=British-British Conversation 2 male British tutor
BBC2_MBS=British-British Conversation 2 male British student
HN = Head nods

From the sample transcript, it can be interpreted that tutors sometimes seem to use HNs in their speaker status as a prompt to confirm agreement or to check their students' understanding. Likewise, the placements of HNs in students are similar to

the tutors. The students use HNs at TTP and in listener status, though a difference between the tutors and the students can be seen in the fact that the numbers of students' use of HNs in listener status is larger than the tutors. This indicates that the students' use of HNs was scattered in listener status and larger numbers of HNs were used in the students' visual response tokens.

These findings were described in the case of the British-Japanese conversation in the pilot study in Section 3.3.1.3, although this tendency was not obvious in the British-British conversation of the pilot study. Through the analysis in the main study, the students' frequent use of HNs was observed both in the British-British conversations and in the British-Japanese conversations as illustrated in the table and the figure in the appendices (see Table B.1.3-2 and Figure B.1-2). As with the tutors' use of HNs, the students also used HNs when they were in speaker status, although the numbers of HNs observed in speaker status were limited:

1. The students often used HNs soon before TTP.
2. The students used HNs in their listener status more than the tutors.
3. The students used HNs in their speaker status very few times.

Both the British students and the Japanese students use HNs more than 250 times in total, and BJC2_MJS in particular uses HNs more than 350 times.

The Japanese students also used HNs at TTP and in listener status. However, the Japanese students' use of HNs in TTP seems to be fewer than the British students. In addition, it can be noted that the numbers of HNs in three of the students, BBC2_MBS(the grey solid line), BJC1_MJS (the black broken line) and BJC2_MJS

(the grey broken line), increase dramatically at an interval of ‘less than -50’. This does not mean that these three students use HNs within 5 seconds between 55 seconds before TTP and 50 seconds before TTP about 100 or 200 times; rather, because all of the HNs used more than 50 seconds before floor-taking were counted into an interval of ‘less than -50’, the three students’ use of HNs produced rather large figures. The indications from the graph and the table are that these three students tended to have longer listener status and use HNs continuously while they were listening to the tutors so that the counts of their HNs in the interval ‘less than -50’ appear as larger figures.

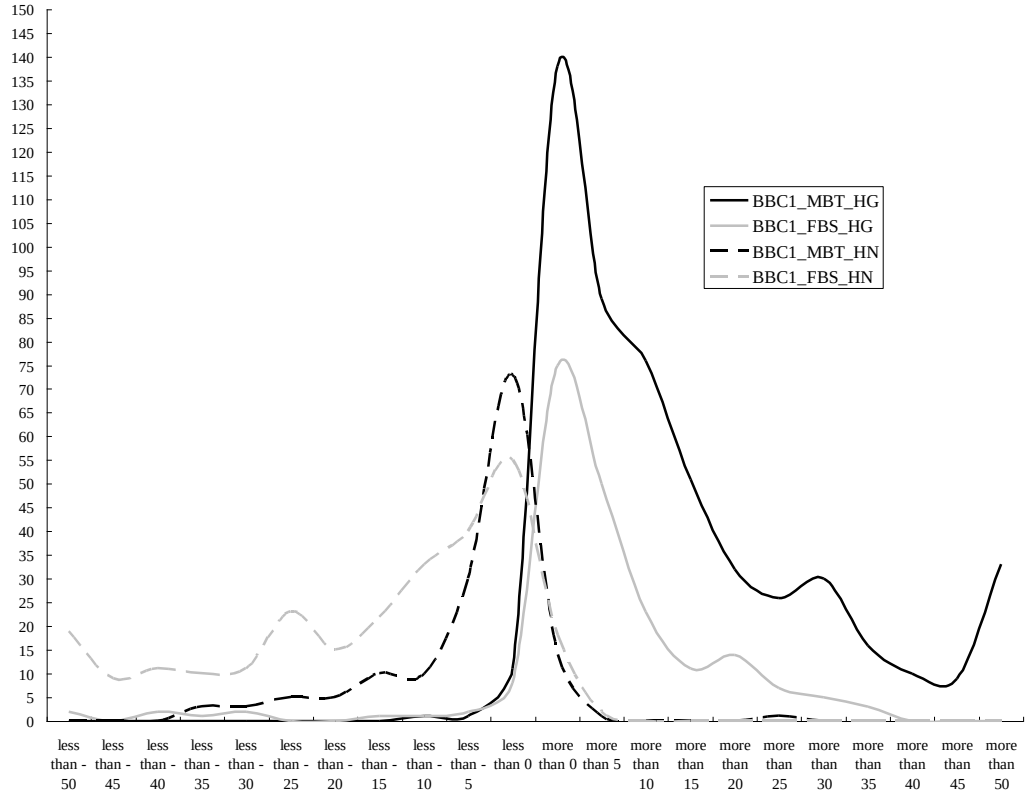
In summary, the use of HNs at TTP and in listener status was found in both the tutors and the students’ use of visual response tokens. The students tended to use HNs in listener status more than the tutors. This tendency was also evident in the Japanese students’ use of HNs.

5.1.3.3 Comparing the use of HGs and HNs in British-British conversation and British-Japanese Conversation

This section offers the results from the same data set of the numbers of HGs and HNs in different presentations. In the previous sections, the use of HGs and HNs was reported with a dichotomy between the tutors and the students. In this section, the use of HGs and HNs in each of the four conversation data will be described in order to highlight differences and similarities between the British-British conversations and the British-Japanese conversations.

In the case of BBC1, symmetry in the placements of HGs and HNs is clearly illustrated in Figure 5.1-3 below (also see Table B.1.4-1). Both BBC1_MBT and BBC1_FBS often use HGs from TTP to 5 to 10 seconds after their floor-taking. This graph resembles the results from the 10-minute data of the British-British

conversation in the pilot reported in Section 3.3.1.3.



Keys: BBC1_MBT_HG = British-British Conversation 1 male British tutor's hand gestures,
BBC1_FBS_HG = British-British Conversation 1 female British student's hand gestures
BBC1_MBT_HN = British-British Conversation 1 male British tutor's head nods,
BBC1_FBS_HN = British-British Conversation 1 female British student's head nods

Figure 5.1-3 Use of HGs and HNs in BBC1

In BBC2, on the other hand, the symmetry in the placements of HGs and HNs is different from BBC1. The peak of the numbers of HGs and HNs was not seen at TTP, but at both ends of the X axis as illustrated in Figure B.1-3 in the appendix (also see Table B.1.4-2). As discussed in the previous section, the average lengths of BBC2_MBT's speaker status and BBC2_MBS's listener status were longer than those of BBC1_MBT and BBC1_FBS. From the Figure B.1-3 in the appendix, these features of a longer turn conversation in BBC2 can be visualized, where BBC2_MBT

frequently used HGs while he was in speaker status, and BBC2_MBS responded to the speaker with continuous HNs.

As observed in the pilot study, placements of HGs and HNs in the British-Japanese conversation appeared different from the British-British conversation. In the main study, the same tendency was recognised, though a symmetrical use of HGs and HNs in the British-British Conversations was not observed in BJC1. A HG was used at TTP and soon after TTP by both BJC1_FBT and BJC1_MJS although the numbers of HGs in both the tutor and the student in BJC1 were fewer than the British-British conversations. The use of HNs at TTP in both BJC1_FBT and BJC1_MJS also seemed to be limited. BJC1_MJS's *continuous* use of HNs is illustrated with a grey broken line in the graph in the appendix (see Table B.1.4-3 and Figure B.1-4).

However, BJC2, the second British-Japanese conversation data in the main study, has similarities with BBC2 in terms of the placements of HGs and HNs. Symmetry in the use of HGs and HNs illustrated in the British-British conversations also appears in BJC2 as shown in the table and the graph in the appendix (see Table B.1.4-4 and Figure B.1-5). Although BJC2_MJS's use of HNs at TTP (the grey split line in the graph) is fewer than the British-British conversations, the use of HGs and HNs in the tutor and the student seems to be balanced as in the British-British Conversations. The notion that the placements of HGs and HNs in the British-Japanese conversation can be assimilated into British-British conversation was not recognisable in the results from the pilot study.

Further analysis on the placements and functions of visual response tokens in reference to differences and similarities between the British-British conversations and the British-Japanese conversations was qualitatively conducted in the turn structure

analysis. Before starting the qualitative analysis, the findings from the quantitative analysis on the use of verbal response tokens will be reported in the next section.

5.1.4 Findings about the use of verbal response tokens

The use of the targeted response tokens was analysed quantitatively, and the numbers of the targeted 4 verbal response tokens, *erm*, *yeah*, *mm* and *mhm*, were counted in five-second time intervals with the methods established in the course of the pilot study. The forms of response tokens, such as clustering, were not taken into account in the quantitative analysis; for occurrences of the targeted response tokens in relation to leadtime were focused on specifically.

A single minimal response token, such as *mm* shown at time 00:34:24 in Transcript 5.1-2 below, was counted as one occurrence of the verbal response token, and added into an interval of 'less than 0' in the table and the graph in the global pattern analysis:

Transcript 5.1-2 Sample transcript: clustering of response tokens

Timeline	Floor	BJC2_ MBT_le adtime	BJC2_ MJS_le adtime	BJC2_MBT_Transcript	BJC2_MJS_Transcript
00: 34: 24		−5	57	Mm.	
00: 34: 25		−4	58		+ of the writer and her readers.
00: 34: 26		−3	59		
00: 34: 27		−2	60		
00: 34: 28		−1	61	Mm. Mm.	
00: 34: 29	MBT_F	0	−91	Yes. Absolutely. Erm I I that's where you I mean that's the way to point how useful you think the idea metafiction is +	

Keys: BJC2_MBT= British-Japanese Conversation 2 male British tutor,
BJC2_MJS = British-Japanese Conversation 2 male Japanese student
MBT_F = British-Japanese Conversation 2 male British tutor's floor-taking

A clustering of response tokens, such as *mm mm* shown at time 00:34:28 in the

timeline in the transcript above, was not distinguished according to the forms. This *mm mm* utterance was counted as an occurrence of *mm* and added to an interval of 'less than 0' in order to measure the distance between the targeted response tokens and TTP by using leadtime.

5.1.4.1 Placement of *erm*

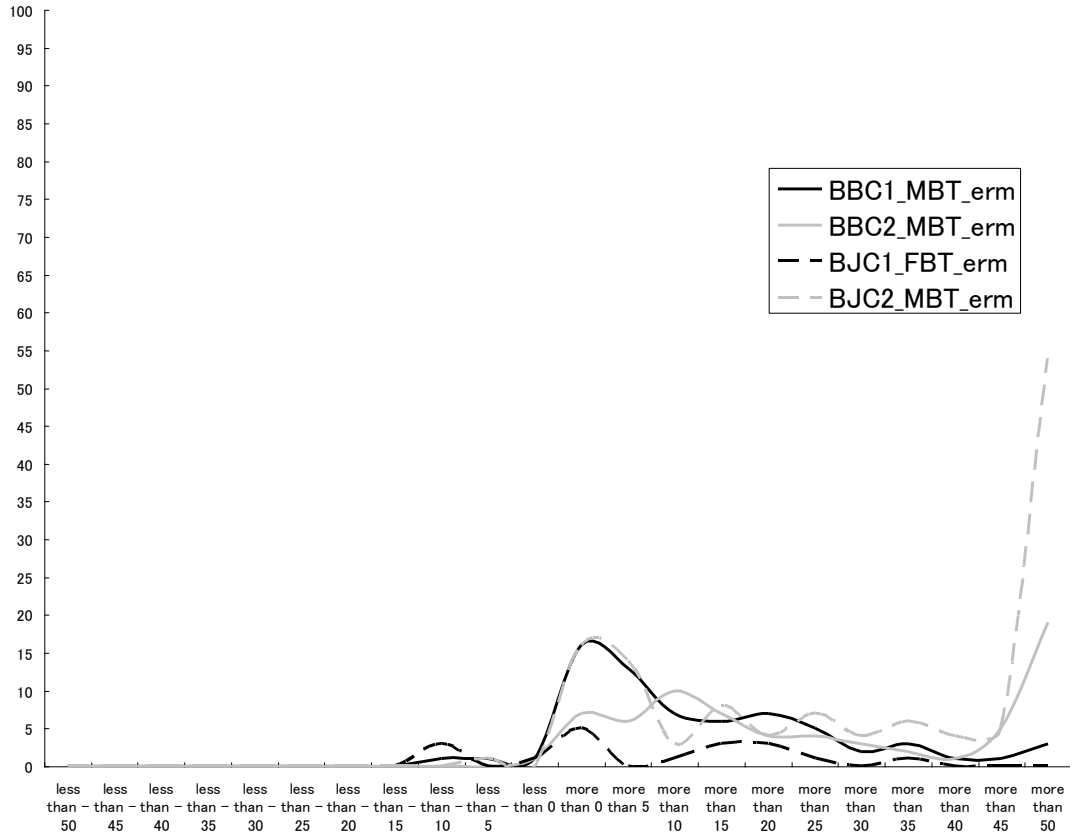
The instances of *erm* in the tutors' utterances are summarised in Figure 5.1-4 below (also see Table B.1.5-1), and some differences in the use of *erm* can be observed between the tutors and the students. The first column of Table B.1.5-1 in the appendix, for example, shows the time intervals from 'less than -50' to 'more than 50', which means 'more than 50 seconds before floor-taking' and 'more than 50 seconds after floor-taking'. The second column shows BBC1_MBT's use of *erm*, the third column shows BBC2_MBT's use of *erm*, and the last two columns describe the numbers of *erm* in BJC1_FBT and BJC2_MBT. The table and the figure show that there were two placements in the tutors' use of *erm*:

1. The tutors used *erm* around TTP.
2. The tutors used *erm* in their speaker status.

The first function of *erm* was reported in the pilot study in Section 3.3.1.4, and the second function was highlighted in the main study through the analysis on longer conversation data.

The numbers in Table B.1.5-1 in the appendix are transferred into Figure 5.1-4 below. The X axis in the figure shows time intervals from 'less than -50' to 'more than 50'. The Y axis alternately illustrates the numbers of the targeted response token,

erm, in this figure.



Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Figure 5.1-4 Tutors' use of *erm*

The black line expresses BBC1_MBT's use of *erm*, the grey line shows BBC2_MBT's use, the broken black line shows BJC1_FBT's use, and BJC2_MBT's use is shown by the grey broken line.

The figure above clearly indicates the first function in the tutors' use of *erm* at TTP. Around more than 0, the tutors use *erm* more frequently to secure the floor. All of the tutors keep using *erm* in their speaker status, as shown in the lines in the figure above, which indicates its second function. In addition, BJC2_MBT uses *erm* in his

speaker status even more than 50 seconds after he takes the floor so that the grey split line increases at 'more than 50'. This means that all of BJC2_MBT's instances of *erm* which uttered more than 50 seconds after he took the floor are included in the interval of 'more than 50'. If BJC2_MBT utters *erm* at 65 seconds after floor-taking, for example, this *erm* is also counted in the interval of 'more than 50'. Items uttered more than 50 seconds before and after floor-taking will be included in to the intervals 'less than -50' and 'more than 50'.

In terms of the *students*' use of *erm*, shown in Figure B.1-6 in the appendix (also see Table B.1.5-2), the students also use it at the floor-taking point. However, not all the students use *erm* during speaker status. BBC2_MBS utters *erm* in his speaker status, which is similar to the tutors' use of *erm* in the second case. BBC1_FBS, in contrast, utters *erm* in her listener status.

Compared with the British students, Japanese students' use of *erm* seems to be limited in numbers. The British students use *erm* more than 50 times in conversations while BJC1_MJS utters *erm* 12 times and BJC2_MJS uses *erm* 14 times. Total numbers of the British students' utterances are about 2500 while that of the Japanese students are about 500 to 600, and therefore the percentage of *erms* in both the British and Japanese students is about 2 %.

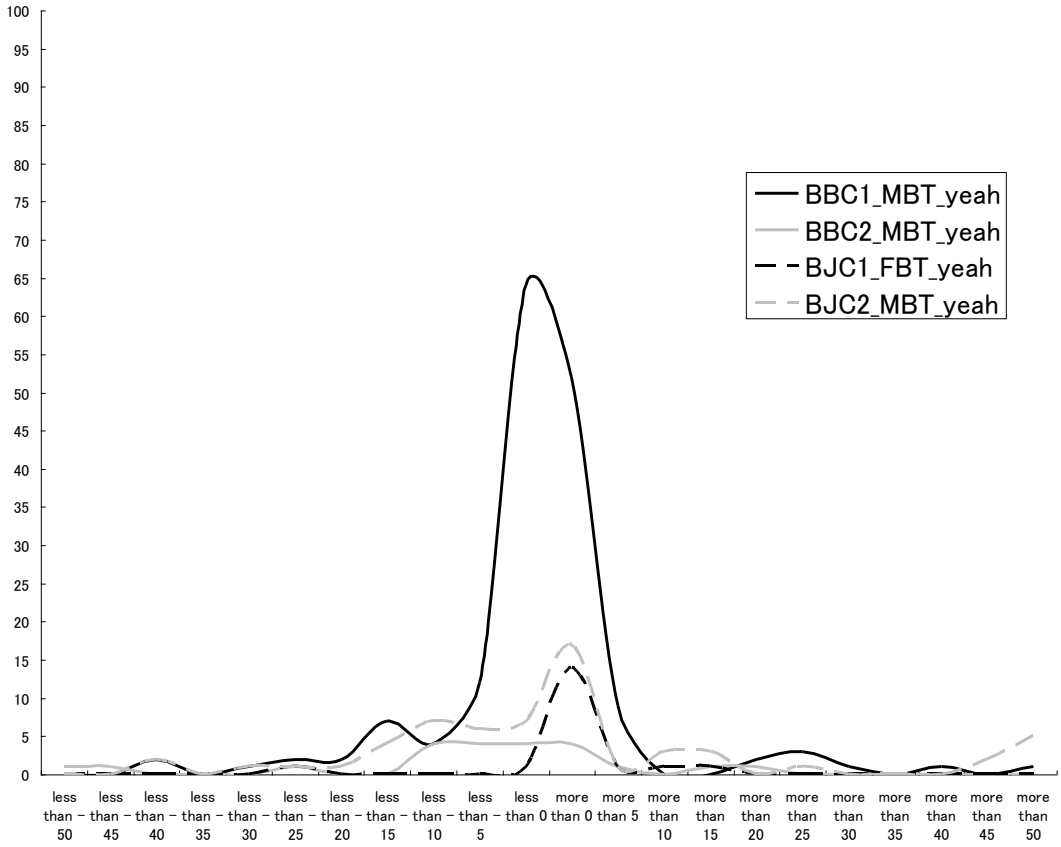
Despite the limitation in the use of *erm* in the Japanese students, however, the figure above shows that the Japanese students also use *erm* at TTP. From this, the following conclusive points can be construed:

1. Both the British and Japanese students used *erm* around TTP.
2. One of the students, BBC1_FBS, often used *erm* in her listener status.

In summary, the tutors tended to use *erm* more than the students. Two placements of *erm* can be recognised in the tutor's use of *erm*: 1) at TTP and 2) at speaker status. The students also used *erm* at TTP, although the use of *erm* at speaker status seems to be limited. BBC1_FBS did use *erm* in her listener status, but this use of *erm* in listener status was not obvious in the other students.

5.1.4.2 Placement of *yeah*

The use of *yeah* in the tutors' speech is illustrated in Figure 5.1-5 below (also see Table B.1.6-1 in the appendix). The table and the figure clearly indicate that the timing when the tutors utter *yeah* is at TTP:



Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Figure 5.1-5 Tutors' use of *yeah*

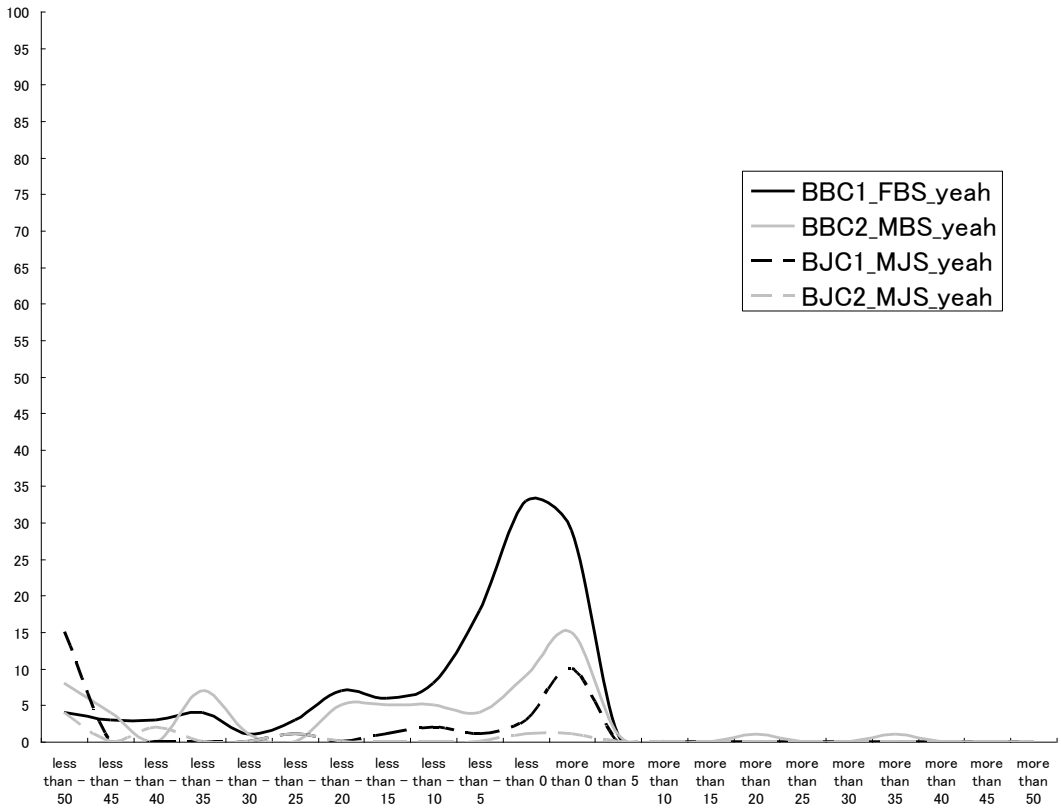
The tutors also use *yeah* in their listener status, at 5 to 10 seconds before their floor-taking if it is more precisely stated. This placement of *yeah* was also reported in the pilot study in Section 3.3.1.4:

1. The tutors used *yeah* around TTP.
2. The tutors used *yeah* in their listener status.

BBC1_MBT in the black solid line in the figure, for example, uses *yeah* about 160 times in total and mostly around TTP. The other three tutors utter *yeah* less frequently, although the tendency to use *yeah* around TTP is apparent. In the case of BBC2_MBT, however, the placements of *yeah* are evenly distributed from 'less than -10' to 'more than 0', which can be seen as different from the other three tutors' usage. Despite this, BBC2_MBT's use of *yeah* can be observed at TTP and soon before TTP.

Comparatively, the students used *yeah* in their listener status more often than the tutors. In Figure 5.1-6 below (also see Table B.1.6-2 in the appendix), BBC1_FBS used *yeah* about 120 times and BBC2_MBS uttered *yeah* 66 times in total. Many of these instances were between 5 seconds before and 5 seconds after TTP. In addition, the use of *yeah* within their listener status was more clearly described in the students' cases. From this, it could be suggested that the students, and in particular the British students, used *yeah* to express their listenership more frequently than the tutors:

1. The students used *yeah* around TTP,
2. The students used *yeah* in their listener status more than the tutors.



Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

Figure 5.1-6 Students' use of *yeah*

In the pilot study, *yeah* was assumed to be characteristic of the male British tutor's choice of response tokens at TTP. Through the analysis in the main study with the larger data set, however, the placement of *yeah* at TTP was found to be in prevalent in both the tutors and the students' response tokens. In addition, the use of *yeah* in listener status distinguished the students' response token *yeah* from the tutors.

In the cases of the Japanese students, fewer instances of *yeah* were observed. BJC1_MJS uttered *yeah* 33 times, and the placement of *yeah* was spread into two areas, at TTP and in listener status respectively, which was similar to the British

students' use of *yeah*. BJC2_MJS used *yeah* only nine times, but, despite this, he also uttered *yeah* at TTP and in listener status. Since the Japanese students sometimes stayed in their listener status longer than 50 seconds, BJC1_MJS's figure at 'less than -50' was 15 and BJC2_MJS's figure at the same interval was 4. Rather than meaning that they uttered *yeah* between 50 seconds after floor-taking to 55 seconds after floor-taking, this means that all of the utterances of *yeah* at more than 50 seconds after floor-taking were included in the interval of 'less than - 50'.

As reviewed in Chapter Two and in Section 3.2.3, *yeah* is defined as an acknowledgement token according to Gardner (2002). From the results, placements of *yeah* seem to be distributed into two areas, namely at TTP and at listener status. *Yeah* in both these two areas can be used as acknowledgement, hence, in the layer of turn-taking, *yeah* in these two placements might have different functions. *Yeah* at TTP might function to convey listener's intention or willingness to become a next speaker while *yeah* at *listener* status might be uttered to show listener's acknowledgement or engagement without attempting to take the next floor of the conversation. Although it is too early to define these functions of *yeah* without conducting further analysis qualitatively, they may be indicative of the multi-functional nature of response tokens.

5.1.4.3 Placement of *mm* and *mhm*

The two response tokens *mm* and *mhm* were also counted and are summarised in the tables and the figures below. These two response tokens were treated as one signal at the stage of the pilot study. However, these two were distinguished as different response tokens in the main study.

Restricted use of *mm* in the tutors was observed as shown in the table and the figure in appendices (see Table B.1.7-1 and Figure B.1-7). Only BJC2_MBT used *mm*

37 times in total and continuously uttered *mm* in his listener status. Aside from these instances, the other tutors rarely used *mm*. Thus, it can be summarised that:

1. The tutors rarely used *mm.*
2. One of the tutors, BBC2_MBT, used *mm* 37 times in total in listener status.

The same tendency was found with *mm* in students' utterances, as illustrated in Figure B.1-8 in the appendix (also see Table B.1.7-2). BBC2_MBS uttered *mm* about 200 times, although the other three students used *mm* less than 10 times in total. BBC2_MBS continuously used *mm* in his listener status even more than 50 seconds before he took the floor, and again the numbers of *mm* uttered more than 50 seconds before floor-taking were counted into the interval of 'less than -50'. As with BJC2_MBT's use of *mm*, this can be one of the cases where personal preference or differences in listenership strategies affect the participants' choice of response tokens in conversation.

According to Handford (2007), the response token *hmm*, which is an equivalent of *mm* but transcribed as *hmm* in CANBEC (Cambridge and Nottingham Business English Corpus), is often used by a senior colleague to a junior colleague in a business setting. BBC2_MBS's frequent use of *mm* can be interpreted as his intention or desire to position himself as equal to his tutor in conversation. BBC2_MBS is aged mid-thirty and working as a part-time lecturer at university while the other three students are aged mid-twenty and have no experience as a lecturer at university. Thus, it can be summarised:

1. The students rarely used *mm.*
2. One of the students, BBC2_MBS, used *mm* about 200 times in total in listener status.

By examining BBC2_MBS's use of *mm* more closely, it can be noticed that there are two placements of *mm*: 1) from 10 seconds before floor-taking to TTP, and 2) from 25 seconds before floor-taking to 30 seconds before floor-taking. From this figure, it might be assumed that *mm* in BBC2_MBS can be used as both a terminator of turn-taking and, conversely, as an acknowledgement with intention to encourage the current speaker to continue talking, as reviewed in Section 2.3.2.

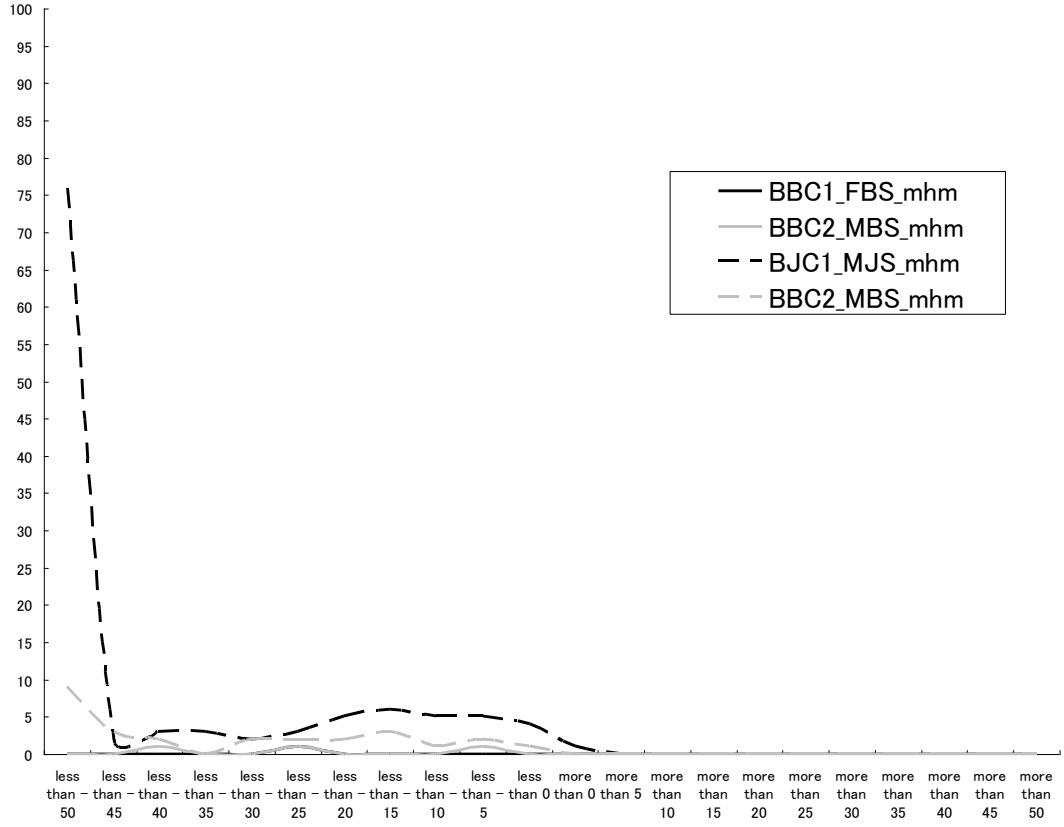
Similarly to their use of *mm*, tutors almost never use a response token *mhm* as shown in the table and the figure in appendices (see Table B.1.7-3 and Figure B.1-9). One of the tutors used *mhm* eight times in total, whereas the other three tutors almost never used *mhm* in the 39-minute length conversations:

1. The tutors almost never used *mhm.*
2. One of the tutors, BJC2_MBT, used *mhm* eight times in total in listener status.

Despite its limited use here, *mhm* also seems to be uttered as continuers from the placement of *mhm* in the table and the figure as reviewed in Chapter Two.

As Figure 5.1-7 below indicates, two British students, BBC1_FBS and BBC2_MBS, seldom used the response token *mhm*, while two Japanese students, BJC1_MJS and BJC2_MJS, kept signalling *mhm* in their listener status (also see

Table B.1.7-4).



Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

Figure 5.1-7 Students' use of mhm

The students' use of *mhm* can be summarised as follows:

1. The students rarely used *mhm.*
2. One of the Japanese students, BJC1_MJS, used *mhm* 115 times in total in listener status.

As one of the most extreme examples, BJC1_MJS used *mhm* more than 100 times in

total. This case has similarities with BBC2_MBS's use of *mm,* for the latter uttered *mm* about 200 times in total as described previously.

In the pilot study, *mm* and *mhm* were transcribed as the same symbol *mm*, and therefore a simple comparison between the findings from the pilot study and the main study in terms of the use of *mm* cannot be made efficiently. Individual preference of *mm,* however, was reported in Section 3.3.1.4 in the case of BJC1_MJS.

There are many ways to interpret BJC1_MJS's frequent use of *mhm* in his listener status and BBC2_MBS's outstanding use of *mm*. BJC1_MJS, for example, might try to adjust his use of response tokens in English as a Second Language by mimicking a native speaker role model, such as BBC2_MBS. Alternatively, BJC1_MJS's use of *mhm* can be explained by L1 transfer as White (1989) and Maynard (1997b) indicate (see Section 2.3.3). It can be said that the case of BJC1_MJS's use of *mhm* supports White (1989) and Maynard (1997b), where they claim that Japanese learners of English deliver more response tokens than native speakers of English. However, the case of BBC2_MBS's use of *mm* might be contrary to their findings since the British student uttered the response token more than the Japanese students. The discussion will be continued in the following chapter.

5.2 Turn structure analysis

There are four sections in the turn structure analysis. Turn-taking patterns will be examined in reference to forms and functions of response tokens qualitatively, with particular focus on the following:

Aspect 1: verbal response tokens

Aspect 2: verbal response tokens with head nods

Aspect 3: hand gestures

Aspect 4: Turn-structural episodes

These four aspects related to turn-taking structure will be analysed in this section.

5.2.1 Focusing on verbal response tokens

Based on O'Keeffe et al. (2007), three forms of verbal response tokens, namely *minimal response tokens*, *non-minimal response tokens* and *clusters*, were counted and summarised in Table 5.2.1-1 below.

Table 5.2.1-1 Forms of verbal response tokens in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
	Number	%	Number	%	Number	%	Number	%
Minimal Response Tokens	90	52%	40	56%	14	54%	57	63%
Non-minimal Response Tokens	15	9%	15	21%	0	0%	3	3%
Clusters	42	24%	6	8%	3	12%	11	12%
Unclassified	27	16%	11	15%	9	35%	19	21%
Total	174	100%	72	100%	26	100%	90	100%

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

BBC1_MBT utters verbal response tokens 174 times in total, which is twice more than BBC2_MBT. Use of verbal response tokens in BJC1_FBT's turns is comparatively limited, and the total number of verbal response tokens in BJC1_FBT is 26, which is less than a third of BJC2_MBT.

Minimal response tokens were most frequently used in the tutors both in the British-British conversations and the British-Japanese conversations. The use of non-minimal response tokens was observed in the tutors in the British-British conversations more than in the British-Japanese conversations, and BBC1_MBT and BJC2_MBT used clusters more than BBC2_MBT and BJC1_FBT.

Both the British students and the Japanese students uttered response tokens more frequently than the tutors as shown in Table 5.2.1-2. BBC2_MBS used verbal response tokens 310 times in total and BB1_FBS uttered verbal response tokens 180 times, which is close to BJC1_MJS, although BJC2_MJS used verbal response tokens only 35 times. It might be worth noting that BJC2_MJS used NH more frequently than the other students as described in Section 5.1.3.2. It can be interpreted from this that BJC2_MJS has chosen HNs rather than verbal response tokens to express his intention to let the tutor continue speaking. This point will be discussed more at a later stage.

Table 5.2.1-2 Forms of verbal response tokens in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS		BJC2_MJS	
	Number	%	Number	%	Number	%	Number	%
Minimal Response Tokens	111	62%	228	74%	152	89%	17	49%
Non-minimal Response Tokens	10	6%	39	13%	1	1%	1	3%
Clusters	14	8%	30	10%	6	4%	3	9%
Unclassified	45	25%	13	4%	11	6%	14	40%
Total	180	100%	310	100%	170	100%	35	100%

Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

Minimal response tokens were more frequently used both by the British and Japanese students than other forms, although there are some differences in the use of non-minimal response tokens and clusters between them. For example, compared with the British students, the use of non-minimal tokens and clusters in the Japanese students is restricted.

5.2.2 Focusing on verbal response tokens with head nods

5.2.2.1 Tutors' use of verbal response tokens with head nods

The use of verbal response tokens with HNs will be examined in this section.

Transcript 5.2-1 below shows BBC1_MBT's use of response tokens before he takes a floor at TTP:

Transcript 5.2-1 BBC1_MBT's response tokens with head nods

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 24: 59	FBS_F	-18	0				Okay. Well I'm gonna find some kind of er principle to f= er select some data from this website.
00: 25: 00		-17	1			HG	
00: 25: 01		-16	2			HG	
00: 25: 02		-15	3	HN			
00: 25: 03		-14	4			HG	
00: 25: 04		-13	5				
00: 25: 05		-12	6	HN	Right.		
00: 25: 06		-11	7				Erm so that I've got something more concrete+
00: 25: 07		-10	8			HG	
00: 25: 08		-9	9	HN	Okay.		
00: 25: 09		-8	10				+in terms of what data I'm going to be using.
00: 25: 10		-7	11		Yeah.		Erm I quite like this idea of doing a comparative study so+
00: 25: 11		-6	12				
00: 25: 12		-5	13				
00: 25: 13		-4	14				
00: 25: 14		-3	15	HN			
00: 25: 15		-2	16	HN	Right.		+I can start off+
00: 25: 16		-1	17		Yeah yeah.		doing that I think and then+
00: 25: 17	MBT_F	0	-59	HG	If you frame that as a sort of research question as as precisely as you can.		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MBT_F = male British tutor's floor-taking

BBC1_MBT utters four different forms of verbal response tokens, such as minimal response tokens *right*, *okay*, *yeah* and a cluster *yeah yeah*. When he utters *right* at 00:25:05, *okay* at 00:25:08 and the second *right* at 00:25:15, he also uses a HN. Furthermore, within 3 seconds before he takes the floor at 00:25:17, he uses 2 HNs and utters *right* and *yeah yeah*. Frequent use of HNs also seems to be a sign of speaker change in this case.

In BBC2_MBT's listener status in Transcript B.2-1 in the appendix, he utters only 2 verbal response tokens, such as *okay right* at 00:33:22 and *okay* at 00:33:27 and also uses 3 HNs, more than 8 seconds before his next turn-taking. Although

BBC2_MBT uses fewer items of response tokens and no collocations between verbal response tokens and HNs was observed, some similarities can be recognised in BBC1_MBT's choice of verbal response tokens with BBC1_ MBT, such as the incremental use of response tokens before TTP.

As shown in Transcript B.2-2 in the appendix, BJC1_FBT used response tokens fewer than the previous two tutors. In a 12-second period of listener status, BJC1_FBT uttered only one non-minimal response token, *alright* at 00:18:22 although she uttered *okay* when she took a turn at 00:18:24, which can be treated as *both* a response token and a discourse marker. Her choice of verbal response tokens is similar to the previous tutors. Further, two HNs were observed; however, no collocations of verbal response tokens with HNs appeared.

BJC2_MBT's response tokens seem to be different from the other three tutors. Frequent use of a minimal response token *mhm* with HNs characterises his use of response tokens. As shown in Transcript B.2-3 in the appendix, the collocation of *mhm* with HNs in BJC2_MBT seems to function as an encouragement of his partner's speaking. In contrast, BJC2_MBT used a minimal response token *mhm* twice, the other minimal response token *mm* once and a cluster *mhm mhm* once in the 20 seconds of listener status. When BJC1_MBT took a speaker turn at 00:19:47, a cluster *mm mm yeah* and a non-minimal token *absolutely* were also uttered. These instances again can be treated as both response tokens and discourse markers.

A variety of uses of verbal response tokens was observed in the tutors' listener status. Response tokens, such as *right, okay, alright* and clusters, seem to function as signals for turn exchange. This observation can support the findings in Carter and McCarthy (2006), which highlights tutors' use of *right* and *okay* at topic or speaker

change in an academic setting. However, collocations between verbal response tokens and HNs seem to be rather arbitrary in the current study; an issue that could be explored further in a larger investigation. In addition, some of the tutors tended to use verbal response tokens with HNs as a sign for turn exchange, whereas others tended to use them as continuers. Non-word response tokens with HNs, in particular, tended to be used as continuers.

5.2.2.2 Students' use of verbal response tokens with head nods

As described in the global pattern analysis in Section 5.1.4, the British students tend to use *yeah* in their listener status more than the tutors and the Japanese students.

Transcript 5.2-2 BBC1_FBS's response tokens with head nods

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 17: 37		3	-18			HN	Yeah I think so.
00: 17: 38		4	-17	SC/mouth	+yeah. Or even is I suppose you could yeah. But yeah you're right that would that would be too narrow for what you want to do.		
00: 17: 39		5	-16			HN	
00: 17: 40		6	-15				
00: 17: 41		7	-14				
00: 17: 42		8	-13				
00: 17: 43		9	-12				
00: 17: 44		10	-11			HN	Yeah it would.
00: 17: 45		11	-10		That's sounding like you're gonna y= erm you're gonna be forced to go down the qualitative route with+		
00: 17: 46		12	-9				
00: 17: 47		13	-8				
00: 17: 48		14	-7	HG			
00: 17: 49		15	-6				
00: 17: 50		16	-5				
00: 17: 51		17	-4				
00: 17: 52		18	-3			HN	
00: 17: 53		19	-2		+less data.		Yeah.
00: 17: 54		20	-1			HN	
00: 17: 55	FBS_F	-3	0			HG	Yeah I think I've already+
00: 17: 56		-2	1	HN	Okay.		+decided that that's what's.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript 5.2-2 of BBC1_FBS's listener status shows several patterns of response tokens with *yeah*, such as clusters *yeah I think so* at time 00:17:37, *yeah it*

would at 00:17:44, and a minimal response token *yeah* at 00:17:53. A HN also occurs with clusters of *yeah.*

Transcript 5.2-3 BBC2_MBS's response tokens with head nods

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 28: 20	MBT_F	0	−19		And I think be be detailed+		+<$G?>
00: 28: 21		1	−18				
00: 28: 22		2	−17	HG			
00: 28: 23		3	−16				
00: 28: 24		4	−15			HN	Mm.
00: 28: 25		5	−14		+about+		
00: 28: 26		6	−13				Mm.
00: 28: 27		7	−12		+the process there. Be trans= you know make it transparent.		
00: 28: 28		8	−11	HG			
00: 28: 29		9	−10				
00: 28: 30		10	−9			HN	Mm. Sure.
00: 28: 31		11	−8		So that y= you know a critical reader would+		
00: 28: 32		12	−7	HG			
00: 28: 33		13	−6				Mm.
00: 28: 34		14	−5		+would want to see+	HN	Mm.
00: 28: 35		15	−4		+well hold on how how's that done.		
00: 28: 36		16	−3			HN	Mm. Yeah+
00: 28: 37		17	−2		Yeah.	HG	
00: 28: 38		18	−1				
00: 28: 39	MBS_F	−5	0			HG	+well an= as long as we've justified erm why things+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

BBC2_MBS used *mm* about 200 times in total as reported in Section 5.1.4.3. In Transcript 5.2-3 above, BBC2_MBS utters *mm* six times within about a 20 second-length listener status. The minimal response token *mm* is uttered with a HN at 00:28:24 and at 00:28:34, and clustered with *sure* at 00:28:30 and with *yeah* at 00:28:36.

A variety of response tokens were observed in the tutors and the British students' listener status. The Japanese students' response tokens, however, tended to be more consistently and monotonously. For example, BJC1_MJS used *mhm* about a hundred times as described in Section 5.1.4.3, and most of them were uttered as a minimal response token not as a cluster as shown in Transcript 5.2-4 below. In the 17-

second length listener status, BJC1_MJS also utters a minimal response token *mhm* four times, two of them occur with HNs. It is worth noting that BJC1_MJS keeps sending this response token, sometimes with HNs, in a constant pace and with a monotonous tone through the conversation. When he takes a speaker turn at 00:22:15, he also uses *ah*, which can be treated as a Japanese response token *aa:* and functions as a change-of-state token. This is followed by a non-minimal response token *alright*, which is a discourse marker often uttered by the tutors. *Alright* was not observed in the two British students' listener status, although further discussion on this point is continued in the next chapter.

Transcript 5.2-4 BJC1_MJS's response tokens with head nods

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 21: 58		14	-17		+ doing exercises. I'm I would think that will make it easier. That if I if everyone is working quietly +		mhm.
00: 21: 59		15	-16	HG		HN	
00: 22: 00		16	-15				
00: 22: 01		17	-14				
00: 22: 02		18	-13	HG			
00: 22: 03		19	-12				
00: 22: 04		20	-11	HG			
00: 22: 05		21	-10	HG		HN	mhm.
00: 22: 06		22	-9		+ and I've got a problem +		
00: 22: 07		23	-8	HG		HN	mhm.
00: 22: 08		24	-7		+ The teacher isn't doing anything. It's easy for me to +		
00: 22: 09		25	-6	HG			
00: 22: 10		26	-5	HG			
00: 22: 11		27	-4				mhm.
00: 22: 12		28	-3		+ say "Excuse me can you come to help me?"	HN	
00: 22: 13		29	-2	HG			
00: 22: 14		30	-1				
00: 22: 15	MJS_F	-9	0				Ah alright. Erm in my opinion Japanese students don't do that <$G?>.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

BJC2_MJS used HNs about 370 times in total, which is the highest figure among the students and the tutors in the four conversation data. BJC2_MJS tended to

use HNs at the same pace through the conversation, as shown in Transcript 5.2-5 below, which is close to BJC1_MJS's use of a minimal response token *mhm*.

Transcript 5.2-5 BBC2_MBS's response tokens with head nods

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 09: 57	MBT_F	0	−21		But you always pick up stuff+	SC/chin	
00: 09: 58		1	−20	HG			
00: 09: 59		2	−19		+because you know erm that's the nature of searching erm that however good our search approach is we will miss some items or new stuff will be produced as well and you you need to keep+	HN	Mm.
00: 10: 00		3	−18				
00: 10: 01		4	−17				
00: 10: 02		5	−16				
00: 10: 03		6	−15	HG			
00: 10: 04		7	−14				
00: 10: 05		8	−13			HN	
00: 10: 06		9	−12				
00: 10: 07		10	−11			HN	
00: 10: 08		11	−10			HN	
00: 10: 09		12	−9	HG			
00: 10: 10		13	−8			HN	
00: 10: 11		14	−7	HG			
00: 10: 12		15	−6		+updating your review.	HN	Mm.
00: 10: 13		16	−5				Mm.
00: 10: 14		17	−4	HG	What's happening with the methodology? Can you just give me a=		
00: 10: 15		18	−3				
00: 10: 16		19	−2				
00: 10: 17		20	−1				
00: 10: 18	MBS_F	−47	0			SC/arm	Yeah. Erm. I I've I'm focusing erm at the moment on er on corpus linguistics as that seemed to be I think quite a major part erm of the analysis. Certainly the an initial part+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MBS_T= Male British student's floor-taking, MBT_F = male British tutor's floor-taking

BJC2_MJS's use of verbal response tokens is fewer than other students. He utters *yes* at 00:38:39 in Transcript B.2-4 in the appendix, which can be treated as a convergence token to answer the previous utterance by BJC2_MBT.

In a conversation with long turn exchanges, such as BBC2 and BJC2, the tutors and the students tended to use *yes* more often than a conversation with short turn exchanges, such as BBC1 and BJC1. BBC2_MBS, for instance, uttered *yes* 28 times in total while BBC1_FBS used *yes* only twice in total. Conversely, BBC1_MBT

and BJC1_MJS did not use *yes* at all (see Appendix C).

By observing the students' listener status, it was noted that there were a few cases where decreases in the use of response tokens were observed. For example, Transcript 5.2-5 above shows that BBC2_MBS uses HNs 6 times and utters a minimal response token *mm* 3 times. However, he stops using visual and verbal response tokens 4 seconds before his next floor-taking. This tendency was also observed in the other three students (also see Transcript B.2-4). The kinds of decreases in the use of verbal and visual response tokens before their next floor-taking will be analysed further in the next section.

In this section, the use of verbal response tokens with HNs has been investigated with time-related transcripts. Compared with the Japanese students, the British tutors and the British students seem to have greater variety in their use of verbal response tokens and more combinations of verbal response tokens with HNs. This inference can support Fung and Carter's (2007) study on Chinese learners' use of response tokens; for, in their study, restricted use of discourse markers was observed in learners of English in Hong Kong when compared with native speakers of British English. The use of verbal response tokens with HNs is examined in reference to turn structure in Section 5.2.4.

5.2.3 Focusing on hand gestures

This section will explore the use of HGs at TTP with transcripts and image captures. As reviewed in Section 2.1.5, four types of conversational gestures have been recognised by Goldin-Meadow (1999), namely, iconic gestures for describing a picture in the speaker's mind; metaphoric gestures for expressing the speaker's idea or thought; beat gestures for adding rhythm, and deictic gestures for pointing to

something. The multi-functional nature of HGs will be taken into consideration in this analysis.

In Transcript 5.2-6 below, BBC1_FBS takes a speaker turn at 00:20:35 with a HG, which is captured in Figure 5.2-1 below. At the same time when BBC1_FBS says *well* at TTP, she raises her right hand, which signals a turn exchange. At 00:20:46, BBC1_MBT has taken a speaker turn with HGs. BBC1_MBT raises his right hand with his palm open and then moves the hand toward left side to express 'the doctor ones' as shown in Figure 5.2-2 and Figure 5.2-3:

Transcript 5.2-6 Hand gestures in BBC1_MBT and BBC1_FBS

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 20: 35	FBS_F	-11	0			HG	Well I was quite interested in that perspective erm but I I don't want to approach er these patient narratives with+
00: 20: 36		-10	1				
00: 20: 37		-9	2				
00: 20: 38		-8	3			HG	
00: 20: 39		-7	4	HN			
00: 20: 40		-6	5			HN	
00: 20: 41		-5	6		Yeah.		
00: 20: 42		-4	7				+the perspective of saying how is it controlling their social perceptions+
00: 20: 43		-3	8	HN			
00: 20: 44		-2	9			HN	
00: 20: 45		-1	10		Yeah.		
00: 20: 46	MBT_F	0	-28	HG	But you could do that with the doctor ones.		+of+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

Figure 5.2-1 BBC1_FBS's HG

Figure 5.2-2 BBC1_MBT's HG (1)

Figure 5.2-3 BBC1_MBT's HG (2)

These hand gestures can function in two layers: one for taking a turn and the other for conveying his idea with his utterances. The latter function can be categorised as metaphoric gesture in Goldin-Meadow (1999).

In Transcript B.2-5 in the appendix, BBC2_MBT's HG at 00:06:17 in the timeline also shows the multi-functional nature of HGs. After BBC2_MBT utters *well I mean* at 00:06:15, he raises his right hand slightly with his palm open toward to BBC2_MBS as shown in the capture of BBC2_MBT's HG in Figure B.2-1 in the appendix. This signals a turn exchange and also expresses his intention to let BBC2_MBS wait and listen to him.

The use of metaphoric gestures was also observed in BBC2_MBS's HG. When BBC2_MBS takes a speaker turn at 00:28:16 in the timeline in Transcript B.2-6 in the appendix, he moves his hand from the right to the centre to describe the utterance: 'that's […] definitely going in methodology'. *That* here deictically refers to the process of cleaning a corpus, which they are discussing in the transcript. To write down the process of cleaning a corpus in the methodology section in his thesis is an abstract idea. However, BBC2_MBS is trying to express this idea with HGs (see Figure B.2-2 and Figure B.2-3 in the appendix). At the same time, he takes a speaker turn in the conversation.

HGs at TTP were also observed in the British-Japanese conversations. In Transcript 5.2-7 below, BJC1_FBT and BJC1_MJS are talking about English classes in Japan. BJC1_FBT uses metaphoric gestures when she takes the floor at 00:17:58 by saying 'So there is no speaking'. BJC1_FBT raises her both hands and opens her palms toward BJC1_MJS as shown in Figure 5.2-4 below. This gesture describes a quiet English class in Japan. At 00:17:59, BJC1_MJS takes a speaker turn with HGs

as described in the transcript above. He slightly raises his right hand with his palm open and draws a small circle several times with his hand as shown in the captures BJC1_MJS's HG (1) and (2) below.

Transcript 5.2-7 Hand gestures in BJC1_FBT and BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 17: 40		2	-15		Okay so okay from this I have a picture of Japanese class working very quietly.		
00: 17: 41		3	-14				
00: 17: 42		4	-13				
00: 17: 43		5	-12				
00: 17: 44		6	-11				
00: 17: 45		7	-10	HT/HG			
00: 17: 46		8	-9				
00: 17: 47		9	-8	HG		HG	mhm.
00: 17: 48		10	-7			HN	
00: 17: 49		11	-6		Is that right?	HN	
00: 17: 50		12	-5		Teacher says "okay open your book and do exercise three."		
00: 17: 51		13	-4	HG			
00: 17: 52		14	-3				
00: 17: 53		15	-2				Y= yeah.
00: 17: 54		16	-1				
00: 17: 55	MJS_F	-3	0		And they are=		Yeah I have experience like that.
00: 17: 56		-2	1	HG			
00: 17: 57		-1	2				
00: 17: 58	FBT_F	0	-1	HG	So there is no speaking.	HG	
00: 17: 59	MJS_F	-25	0			HG	No especially i= if students want to say something er they can do. But normally er just teacher says something.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Figure 5.2-4 BJC1_FBT's HG

Figure 5.2-5 BJC1_MJS's HG (1)

Figure 5.2-6 BJC1_MJS's HG (2)

This HG is subtle, but still signals a turn change. It also describes an action; that is, Japanese students' speech in class, which supports the utterance, 'if students want to say something'.

In BJC2, deictic gestures were observed both in BJC2_MBT and BJC2_MJS. At 00:14:11, BJC2_MJS takes a speaker turn with a HG as shown in Transcript B.2-7 in the appendix. While BJC2_MJS utters 'so this er sorry' at TTP, he points to the paper with his palm down to indicate 'the chapter' as shown in the capture of

BJC2_MJS's HG (see Figure B.2-4 in the appendix). This gesture also functions on two layers: to signal speaker change and to indicate the chapter he needs to clarify. At 00:14:34, BJC2_MBT takes a speaker turn with HNs and HGs. BJC2_MBT also points to the paper with his hand positioned close and downwards, as if holding an invisible pen, while saying 'y= yes you should go' (see Figure B.2-5 in the appendix). This hand gesture also seems to signal a turn exchange and indicate the point on the paper to be revised.

Although the numbers of HGs were limited in the Japanese students, the use of HGs at TTP was observed both in the British-British conversations and in the British-Japanese conversations. Some of the HGs observed at TTP seem to function as two layers, namely turn management and delivering the speaker's idea supportive of utterances. In this section, the use of HGs at TTP has been analysed with the time-related multimodal transcripts and images. The multi-functional nature of HGs has also been observed both in the British-British conversations and the British-Japanese conversations.

5.2.4 Turn-structural episodes

5.2.4.1 Numbers of turn-structural episodes

As described in Section 4.1.6 and Section 4.3.3.4, seven turn-structural episodes were recognised as shown in Table 5.2.4-1 below. This categorisation was applied to the turn structure analysis:

Table 5.2.4-1 Turn-structural episodes

Episode 1	A's turn closing → B's turn-taking
Episode 2	A's turn keeping → B's cut-in
Episode 3	A's turn closing → B's turn refusal → A's turn retaining
Episode 4	A's turn closing → B's turn refusal → A's turn refusal → B's final turn-taking
Episode 5	A's turn giving → B's turn-taking
Episode 6	A's turn giving → B's turn refusal → A's turn retaining
Episode 7	A's turn giving → B's turn refusal → A's turn refusal → B's final turn-taking

(Adapted from Ohama 2006)

Episode 1 shows B's turn-taking after A's turn closing, where B has become the next speaker by self-selection at TRP. In Episode 2, B takes the next speaker turn by self-selection and not at TRP while A is holding the speaker turn, which is labelled as cut-in. Alternately, Episode 3 and Episode 4 show patterns of turn-taking after the partners' turn refusal at TRP, which were introduced in Ohama (2006). Episode 5 is a case of turn-taking after the partner's turn giving, which is referred to as other-selection in Sacks (1974). Episode 6 and Episode 7 are cases of turn refusal after the partners' turn giving, which are added in this current research. In Episode 6 in particular, a turn giver retains the turn after the partner's turn refusal. In Episode 7, the partner's final turn-taking follows a turn giver's turn refusal.

The occurrences of these turn-structural episodes in the conversations between the tutors and the students are summarised in Table 5.2.4-2 and Table 5.2.4-3 below. Episode 1 (A's turn closing → B's turn-taking), Episode 2 (A's turn keeping → B's cut-in) and Episode 5 (A's turn giving → B's turn-taking) can be noted as basic patterns in English conversation since all the participants use these three patterns.

Table 5.2.4-2 Turn-structural episodes in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
	Number	%	Number	%	Number	%	Number	%
Episode 1	33	31%	10	28%	15	13%	14	42%
Episode 2	51	48%	20	56%	3	3%	3	9%
Episode 3	14	13%	0	0%	17	14%	5	15%
Episode 4	1	1%	0	0%	2	2%	0	0%
Episode 5	3	3%	3	8%	0	0%	4	12%
Episode 6	1	1%	0	0%	13	11%	3	9%
Episode 7	0	0%	0	0%	0	0%	0	0%
Unclassified	3	3%	3	8%	69	58%	4	12%
Total	106	100%	36	100%	119	100%	33	100%

Table 5.2.4-3 Turn-structural episodes in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS		BJC2_MJS	
	Number	%	Number	%	Number	%	Number	%
Episode 1	17	18%	9	26%	6	25%	5	21%
Episode 2	37	40%	6	18%	3	13%	3	13%
Episode 3	1	1%	1	3%	2	8%	1	4%
Episode 4	3	3%	0	0%	1	4%	0	0%
Episode 5	33	35%	18	53%	9	38%	13	54%
Episode 6	0	0%	0	0%	0	0%	1	4%
Episode 7	1	1%	0	0%	3	13%	0	0%
Unclassified	1	1%	0	0%	0	0%	1	4%
Total	93	100%	34	100%	24	100%	24	100%

In BBC1_MBT, the majority of the cases were categorised into Episode 1 and Episode 2, hence, Episode 3 has also used 14 times, and Episode 4 (A's turn closing → B's turn refusal → A's turn refusal → B's final turn-taking) and Episode 6 (A's turn giving → B's turn refusal → A's turn retaining) were also shown once each in BBC1_MBT. In the British-Japanese conversations, the numbers of Episode 2 in the tutors was less than in the British-British conversations. Episode 3, Episode 4 and Episode 6 are observed in the tutors in the British-Japanese conversations more than the British-British conversations. All of the students have Episode 3, although the numbers are limited.

Both the British students and the Japanese students tended to take turns with Episode 5 more frequently than the tutors, though the use of Episode 1 in the students

was fewer than the tutors. Additionally, the British students used Episode 2 more than the Japanese students. BBC1_MBS and BJC1_MJS had Episode 4 and Episode 7 (A's turn giving → B's turn refusal → A's turn refusal → B's final turn-taking). Episode 6 appeared more frequently in the British-Japanese conversations than the British-British conversations. These features will be examined with transcripts in detail in the next section.

5.2.4.2 Turn-structural episodes and listenership

5.2.4.2.1 Turn-structural episodes and listenership in BBC1_MBT

In this section, the relationship between turn-structural episodes and length of listener status in each participant will be investigated in reference to both the forms and placements of visual and verbal response tokens. Means, standard deviations (SD, hereafter) and variances of listener status length in each episode will be calculated. In BBC1_MBT as shown in Table 5.2.4-4 below, for example, the mean listener status length in Episode 1 is 8.97 seconds, which means that after about nine seconds listener status, BBC1_MBT takes a speaker turn at TRP by self-selection in average:

Table 5.2.4-4 Turn-structural episodes and listenership in BBC1_MBT

	BBC1_MBT		Mean of Leadtime	SD	Variance
	Number	%	(seconds)		
Episode 1	33	31%	8.97	7.68	58.94
Episode 2	51	48%	7.65	9.39	88.15
Episode 3	14	13%	5.14	10.01	100.12
Episode 4	1	1%	10.00	0.00	0.00
Episode 5	3	3%	2.67	0.47	0.22
Episode 6	1	1%	3.00	0.00	0.00
Episode 7	0	0%	--	--	--
Unclassified	3	3%	--	--	--
Total	106	100%			

As shown in Transcript 5.2-8 below, BBC1_MBT gives a HN at 00:01:42 in the timeline and utters a cluster of verbal response tokens at 00:01:44, *oh I see right*,

which is around 5 seconds before floor-taking. Within 3 seconds before the floor-taking, BBC1_MBT uses a HN again at 00:01:46, and utters a minimal response token *right* at 00:01:47, and then BBC1_MBT gets into his next speaker turn with HGs after BBC1_FBS's turn closing. *Erm* at time 00:01:48 in BBC1_FBS seems to signal the closing of her turn:

Transcript 5.2-8 Episode 1 in BBC1_MBT

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 01: 36	FBS_F	-13	0				No they it's these interviews are on the website as a kind of erm I guess like a self-help forum for people to go and other people to go and visit+
00: 01: 37		-12	1				
00: 01: 38		-11	2			HG	
00: 01: 39		-10	3				
00: 01: 40		-9	4				
00: 01: 41		-8	5			HG	
00: 01: 42		-7	6	HN		HG	
00: 01: 43		-6	7				
00: 01: 44		-5	8		Oh I see right.	HG	+and they're separated out by the particular conditions.
00: 01: 45		-4	9				
00: 01: 46		-3	10	HN			
00: 01: 47		-2	11		Right.	HG	
00: 01: 48		-1	12				Erm.
00: 01: 49	MBT_F	0	-8	HG	So th= the interviews are there because of the condition not because of the any+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

In Episode 2 in BBC1_MBT, the mean listener status length is less than Episode 1. In Transcript 5.2-9 below, after about 8 seconds of listener status, BBC1_MBT takes a speaker turn by cut-in during BBC1_FBS's speaking. Through the observation of the data, two strategies used in Episode 2 in BBC1_MBT were recognised as shown in Transcript 5.2-9 and Transcript 5.2-10 below:

Transcript 5.2-9 Episode 2 in BBC1_MBT: Sample 1

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 05: 44	FBS_F	-9	0	HN	Right.	HN	So I suppose as they would in a+
00: 05: 45		-8	1		Yeah.	HG	
00: 05: 46		-7	2	SC/arm			+erm consultation with a patient but+
00: 05: 47		-6	3				
00: 05: 48		-5	4	HN	Yeah.		
00: 05: 49		-4	5			HG	+trying to generalise it+
00: 05: 50		-3	6	HN	Right.		+for everybody members+
00: 05: 51		-2	7		Yeah.		
00: 05: 52		-1	8	HG	but I+		
00: 05: 53	MBT_F	0	-6		+suppose the interesting thing is what what source domains they're using.	HG	+of the public going on the website

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking, SC/arm=self comfort with arm

Transcript 5.2-10 Episode 2 in BBC1_MBT: Sample 2

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 16: 02	MBT_F	0	-10		I see so you're talking about seven thousand words pretty much+		
00: 16: 03		1	-9				
00: 16: 04		2	-8				
00: 16: 05		3	-7				
00: 16: 06		4	-6				
00: 16: 07		5	-5		+you get about five hundred words+		Is it?
00: 16: 08		6	-4		+of type on an A4 side. Yeah yeah six to seven thousand.		Yeah.
00: 16: 09		7	-3				
00: 16: 10		8	-2				
00: 16: 11		9	-1			HN	
00: 16: 12	FBS_F	-3	0				So they're quite there is quite+
00: 16: 13		-2	1		Okay.		
00: 16: 14		-1	2				+a lot of data I wouldn't+
00: 16: 15	MBT_F	0	-1		And how many interviews are there roughly?		
00: 16: 16	FBS_F	-8	0				Er well for this er for the chronic pain one I was looking at there mus= there are about thirty I think.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

In sample 1, several verbal response tokens with HNs are used before BBC1_MBT's cut-in at 00:05:53 in the timeline. BBC1_MBT's HGs are also observed at TTP in sample 1.

In sample 2 above, no HN in listener status and no HG at TTP are used in BBC1_MBT's cut-in at 00:16:15 in the timeline. Only one minimal response token, *okay*, is uttered at 00:16:13 in this case. They are talking about the data BBC1_FBS

has collected for her MA thesis, and BBC1_MBT is trying to clarify the amount of data BBC1_FBS has obtained in sample 2. For this purpose, BBC1_MBT has cut-in to BBC1_FBS's explanation about data after 3 seconds of listener status. This kind of cut-in strategy can be differentiated from the previous example, and, in particular, the latter strategy can be referred to as *cut-in for clarification*, which encourages the current speaker's talking by giving an inquiry or clarification cue.

As described in Section 4.3.3.4, a discourse framework was considered in the current analysis. Based on Saft (2007), two frameworks in academic tutorials were recognised in this study: a commentary framework and an explanatory framework. All of the transcripts of BBC1_MBT examined above belong to the student's explanatory framework, whereas the first two cases, Episode 1 and Episode 2 with BBC1_MBT's engagement, tend to lead a transition from the student's explanatory framework to the tutor's commentary framework. Cut-in for clarification in sample 2 seems to invoke the student's speaking in the explanatory framework.

Episode 3 appeared in BBC1_MBT 14 times in total. As shown in the table above, after about 5 seconds of pause, BBC1_MBT takes a speaker turn in Episode 3. In Transcript B.2-8 of Episode 3 in BBC1_MBT in the appendix, BBC1_MBT closes his turn with the utterance at 00:09:49 in the timeline and BBC1_FBS uses 2 HNs and utters 2 minimal response tokens, *yeah* and *erm*, without taking the next speaker turn, even though this is a possible speaker change point. Then BBC1_MBT utters *yeah okay* at 00:09:58, which is followed by a silent pause. After this 4-second length pause, BBC1_MBT takes the floor. This is an example of negotiation of speaker change observed in the British-British conversation.

5.2.4.2.2 Turn-structural episodes and listenership in BBC1_FBS

BBC1_FBS also had a wide range of episodes (see Table B.2.3-1 in the appendix). SD and variances of listener status length in each episode in BBC1_FBS were larger than BBC1_MBT so that it can be said that BBC1_FBS's listener status length varied more than BBC1_MBT, and that the means of listener status in each episode were less reliable than BBC1_MBT. Despite the lower reliability of the means of listener status length in each episode, the forms and placements of verbal and visual response tokens in Episode 2, 4 and 5 in BBC1_FBS were analysed qualitatively.

As described in Transcript B.2-9 in the appendix, BBC1_FBS used cut-in in tutor's commentary framework. In the transcript, BBC1_FBS utters 'yeah, I've read that+' with a HG at 00:10:57, which can be a possible point for a turn exchange. However, BBC1_MBT continues speaking. At 00:11:04 in the timeline, BBC1_FBS uses cut-in again, and takes the next speaker turn at this time. In the transcript, BBC1_FBS uses HNs only twice, one at 3 seconds before the second floor-taking, and the second HN at TTP at 00:11:04. In this way, using fewer HNs and verbal response tokens can function to express listener's intention to lead to a turn.

BBC1_FBS's first attempt to take a speaker turn in the transcript can be treated as an aggressive contribution to conversation, and although she fails to become a speaker at this instance, she succeeds in taking the floor 6 seconds after.

Ohama (2006) compared the use of response tokens in Japanese conversation and in English conversation in New Zealand, and reported that both the speaker and the listener contribute to construct a contextual foundation together in English conversation. In contrast, roles of the speaker and the listener are separated in Japanese conversation: the speaker has the role of building a context and the listener

is expected to support the speaker role indirectly with response tokens. As seen in Episode 2, BBC1_FBS responds to the current speaker with an aggressive cut-in twice within the 14 second transcript, and the absence and reduced use of response tokens in the context signals that she is going to contribute to the conversation, not just as a listener but as a speaker.

Episode 5 was observed in the students' cases more than the tutors. As shown in Transcript 5.2-10 in the previous section, Episode 5 was observed within the students' explanatory frameworks.

Transcript 5.2-11 Episode 5 in BBC1_FBS

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 24: 37		36	-22		+Okay so you've again it's it's what it's what you're doing next is is is what I'm trying to get is a sequence a sort of plan for the summer essentially. Erm cos we're gonna need to get together again in a couple of weeks time or next week even+		
00: 24: 38		37	-21	HG			
00: 24: 39		38	-20				
00: 24: 40		39	-19				
00: 24: 41		40	-18	HG			
00: 24: 42		41	-17				
00: 24: 43		42	-16	HG			
00: 24: 44		43	-15				
00: 24: 45		44	-14				
00: 24: 46		45	-13				
00: 24: 47		46	-12	SC/nose		HN	
00: 24: 48		47	-11				
00: 24: 49		48	-10	HG			
00: 24: 50		49	-9				
00: 24: 51		50	-8				
00: 24: 52		51	-7		+er in order+		Yeah.
00: 24: 53		52	-6	HG			
00: 24: 54		53	-5				<$G?> <$H> further along </$H>
00: 24: 55		54	-4		Yeah er so we need a sort of early plan.		
00: 24: 56		55	-3				
00: 24: 57		56	-2				
00: 24: 58		57	-1				
00: 24: 59	FBS_F	-18	0				Okay. Well I'm gonna find some kind of er principle to f= er select some data from this website.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

The other case of Episode 5 appeared at a boundary between the commentary framework and the explanatory framework. For example, in Transcript 5.2-11 above, BBC1_MBT starts talking about her writing time at 00:24:37 in the timeline after mentioning the importance of linking her MA dissertation to her future PhD research for about 30 seconds. Then the tutor tries to move from the commentary framework to an explanatory framework in order to let BBC1_FBS talk about her plan. Before BBC1_FBS accepts taking a turn at 00:24:59, only one HN at 00:22:47 and a minimal response token *yeah* at 00:24:52 are observed in the transcript, which are followed by an overlapping utterance. In addition, fewer response tokens before floor-taking was also observed in Episode 5.

Episode 4 appeared in BBC1_FBS although the number is limited. In Transcript B.2-10 in the appendix, BBC1_MBT suggests a journal to read to BBC1_FBS at 00:29:52. BBC1_FBS cuts in the conversation, and says that she has found the journal but it is held by someone else. BBC1_MBT takes back a speaker turn soon after and says 'how annoying' at 00:30:02. However, after the BBC1_MBT's comment, there is a 5-second pause. BBC1_FBS does not take a speaker turn soon after BBC1_MBT's turn closing and BBC1_MBT also does not retain the speaker turn after BBC1_FBS's refusal to take the turn. Finally, at 00:30:11, BBC1_FBS takes a speaker turn, and BBC1_MBT utters a minimal response token *yeah* after she takes the floor, which is almost at the same timing as BBC1_FBS's floor-taking. This is a moment where the two participants negotiate which is going to be the next speaker. At that moment, no response token is observed, but eye contact is used to fill the silent pause. It is worth noting that this kind of negotiation was also observed in the other British-British conversations, and will be compared with the

cases in the British-Japanese conversations later.

There is only one case of Episode 7 observed in BBC1_FBS. At 00:22:07 in Transcript B.2-11 in the appendix, BBC1_MBT asks BBC1_FBS, 'whether it's worth at this point trying to either exclude either one or the other of those domains [doctors data and patient data] or explicitly recognize that you're doing a comparative study', which is followed by 4 seconds of silent pause. BBC1_FBS does not take a speaker turn soon after BBC1_MBT's turn giving, but BBC1_MBT also refuses to take the speaker turn. At last BBC1_FBS takes the floor at 00:22:25. As with the case of Episode 4, negotiation of speaker changes was observed here.

5.2.4.2.3 Turn-structural episodes and listenership in BBC2_MBT

BBC2_MBT used only three turn-structural episodes: Episode 1, Episode 2 and Episode 5. BBC2_MBT spoke at a slower pace and used fewer response tokens, which characterised BBC2_MBT's way of talking. The standard deviations, means and variances of BBC1_MBT's listener status length in each episode is summarised in Table B.2.3-2 in the appendix. Since the standard deviations and variances in Episode 1 and Episode 2 are large, the means of listener status length in these two episodes cannot be guaranteed. However, from the figures, it can be said that BBC2_MBT is in his listener status longer than BBC1_MBT. The mean of Episode 5 can be reliable since its standard deviation is low although the number of the cases in Episode 5 is limited.

In an example of Episode 1 in BBC2_MBT in Transcript B.2-12 in the appendix, BBC2_MBS reports and explains his experience at a conference where he has presented recently. During BBC2_MBS's speaker turn, BBC2_MBT utters four different verbal response tokens: *mm* at 00:07:02 in the timeline, *sure* at 00:07:11,

yeah at 00:07:19 and *excellent* at 00:07:3, and uses HNs five times. The interval of the last 2 HNs is narrower than the other three. After BBC2_MBT takes a speaker turn, he introduces a new topic by asking about another conference BBC2_MBS is going to attend with 'and is it in December? November?' at 00:07:31. In this case, the use of episode 1 in BBC2_MBT leads not to a boundary of frameworks but to a topic change, and thus the participants remain within the student's explanatory framework. The use of *excellent* in this example supports Carter and McCarthy (2006) as quoted below:

> Adjectives such as *excellent, fine, great, good, lovely, right, perfect* offer positive feedback to the speaker and often mark the boundaries of topics.
> (Carter and McCarthy 2006: 189)

That's good was also observed in BBC2_MBT's utterances before a topic change as shown in Transcript B.2-13 in the appendix, which is a sample transcript for Episode 2 in BBC2_MBT. After giving some comments on BBC2_MBS's narrative about his chairing a home conference, BBC2_MBT moves to the next topic. This case will be explained further.

Episode 2 in BBC2_MBT also has two variations. Transcript B.2-13 in the appendix is a case of episode 2 in the tutor's commentary framework. BBC2_MBT utters a minimal response token *yeah* twice and uses HNs three times within the last 5 seconds before his next speaker turn. Then he takes the floor at 00:02:25 in the timeline, and after BBC2_MBT has given some comments on what BBC2_MBS has

previously said, he moves to the next topic at 00:02:32 and asks whether BBC2_MBS was able to attend John Sinclair's session at the conference. In this case, the topic has changed, and simultaneously BBC2_MBT leads the conversation from a commentary framework to an explanatory framework after he takes the floor.

The second case of Episode 2 in BBC2_MBT appears in the student's explanatory framework. In Transcript B.2-14 in the appendix, BBC2_MBT takes the speaker turn at 00:36:48 after 16 seconds of listener status with 3 HNs, and responds to BBC2_MBS with 'yeah. can you give me some examples?', which encourages BBC2_MBS to speak. This can be recognised as a cut-in for clarification, which is described in BBC1_MBT's Episode 2 in Section 5.4.2.4.2.1.

Some similarities were observed in Episode 2 between BBC1_MBT and BBC2_MBT. There are two functions of Episode 2: the first function is to lead a topic change or a boundary of frameworks, and the second is to stimulate students to speak with a cut-in for clarification. Response tokens seem to be altered according to these functions in Episode 2 and fewer response tokens were recognised as cut-ins for clarification.

5.2.4.2.4 Turn-structural episodes and listenership in BBC2_MBS

Most of the cases in BBC2_MBS were also categorised into three basic patterns of turn structure, Episode 1, Episode 2 and Episode 5 as shown in Table B.2.3-3 in the appendix. Episode 3 was observed only once in BBC2_MBS. The means of listenership status length of BBC2_MJS in Episode 1, Episode 2 and Episode 5 are not reliable since their standard deviations and variances are large. Through the analysis of Transcript B.2-15 in the appendix, similarities and differences in the use of response tokens in Episode 2 and Episode 5 are observed between BBC1_FBS and

BBC2_MBS.

The case of Episode 2 in BBC2_MBS in Transcript B.2-15 in the appendix occurs in the tutor's commentary framework. Before BBC2_MBS has cut-in and taken a speaker turn at 00:22:47, he utters a minimal response token *mm* and a cluster *mm mm* and uses 3 HNs. This differs from the use of response tokens in Episode 2 in BBC1_FBS, for she rarely uses response tokens before turn-taking in Episode 2. Through the cut-in, however, BBC2_MBS also shows his active contribution to the conversation, which is reported as one of the features of English conversation by Ohama (2006).

Transcript B.2-16 in the appendix is an example of Episode 5 in BBC2_MBS. BBC2_MBT is talking about the importance of attending conferences and updating the ideas while writing a thesis, and moves to another topic about methodology with a turn giving instance at 00:10:14. BBC2_MBS takes a speaker turn after his partner's turn giving, and while BBC2_MBS is listening to BBC2_MBT's commentary, he uses several verbal and visual response tokens. However, 4 seconds before his turn-taking at 00:10:18, he stops using response tokens. This tendency was also observed in BBC1_FBS's Episode 5.

5.2.4.2.5 Turn-structural episodes and listenership in BJC1_FBT

A range of turn-structural episodes was observed in BJC1_FBT as shown in Table B.2.3-4 in the appendix. Episode 5, which demonstrates turn giving, and Episode 7, which comprises final turn-taking after partner's turn giving, was not observed in BJC1_FBT. The means of listener status in each episode can be validated from the figures of the SD and variances although occurrences are quite few in Episode 2 and Episode 4. On average, after 11 seconds length listener status, BJC1_FBT takes a

speaker turn by herself at TPR, which is categorised as Episode 1. A typical example displays that, after about a 5-seconds pause, BJC1_FBT retains the next turn in Episode 3. In Episode 6, BJC1_FBT gives a turn to BJC1_MJS, but he disclaims the turn. Thus, she retains the next turn after about 6 seconds of pauses on average. The number of cases categorised into unclassified is 69; most being cases where BJC1_FBT was reading and checking BJC1_MJS's essay during the supervision and BJC1_FBT gave comments on the essay after reading.

As examined in the global pattern analysis in Section 5.1, the number of response tokens used in BJC1_FBT is limited. However, Episode 1 in BJC1_FBT has similarities in the use of response tokens with BBC1_MBT and BBC2_MBT, such as a HN at 00:18:17 and at 00:18:23 in the timeline, and a non-minimal response token *alright* at 2 seconds before she takes the floor at 00:18:24, as shown in Transcript B.2-17 in the appendix. Although the number of response tokens is few, she still shows some HNs for continuers and utters *alright* to express her satisfaction with the information given by BJC1_MJS. These response tokens also lead to a topic change after she takes the floor.

BJC1_FBT also used Episode 2, which is a cut-in, in order to encourage BJC1_MJS's speaking as shown in Transcript B.2-18 in the appendix, although the number of Episode 2 in BJC1_FBT is only 3. At 00:17:58 in the timeline, BJC1_FBT utters 'So there is no speaking', which is not an inquiry but still reinforces the student's speech. BJC1_MJS then continues 'No especially i= if students want to say something er they can do. But normally er just teacher says something'. This can be categorised as a cut-in for clarification as described in the cases of BBC1_MBT and BBC2_MBT.

Episode 3 and Episode 6 appeared in BJC1_FBT often since BJC1_MJS tended to be in silence after BJC1_FBT's turn closing and her turn giving. In Transcript 5.2-12 below, BJC1_FBT comments on BJC1_MJS's essay about classroom observation and tries to let him recognise two different types of observations. At 00:04:56, BJC1_MJS utters a minimal response token *ah*, which is a *change of state* token in Japanese according to Mori (2002). At 00:05:08, BJC1_FBT utters 'you see that's a bit difference' and closes her turn. This is followed by BJC1_MJS's *yeah*, which is a convergence token to BJC1_FBT's previous utterance and can be a signal for a possible shift from a commentary framework to an explanatory framework.

Transcript 5.2-12 Episode 3 and Episode 6 in BJC1_FBT

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 04: 56		20	-26		+ not criticise the teacher teaching +	HN	Ah.
00: 04: 57		21	-25				
00: 04: 58		22	-24	HG			
00: 04: 59		23	-23	HG			mhm.
00: 05: 00		24	-22		+ but the kind of recognise something which will enable me to develop my own teaching as the observer.	HN	
00: 05: 01		25	-21	HG			
00: 05: 02		26	-20				
00: 05: 03		27	-19			HN	
00: 05: 04		28	-18	HG			
00: 05: 05		29	-17				
00: 05: 06		30	-16			HN	
00: 05: 07		31	-15				mhm.
00: 05: 08		32	-14		You see that's a bit difference	HN	
00: 05: 09		33	-13	HG			
00: 05: 10		34	-12				Yeah.
00: 05: 11		35	-11		Erm.	HN	
00: 05: 12	Pause	-2	-10		<$E> pause </$E>		
00: 05: 13		-1	-9			HN	
00: 05: 14	FBT_F	0	-8		What else have you read about observation?		
00: 05: 15		1	-7				
00: 05: 16		2	-6				
00: 05: 17	Pause	-4	-5		<$E> pause </$E>		
00: 05: 18		-3	-4				
00: 05: 19		-2	-3				
00: 05: 20		-1	-2				
00: 05: 21	FBT_F	0	-1		Oh you've got <$H> Winebury </$H>.		
00: 05: 22	MJS_F	-2	0				Yeah I've got.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

However, BJC_MJS refuses to take the next speaker turn. BJC1_FBT then utters *erm* at 00:05:11. After a 2-second silent pause, BJC1_FBT takes the floor and asks a question, 'what else have you read about observation?' which indicates a point of topic change. This is one of the cases of Episode 3 in BJC1_FBT. After BJC1_FBT's inquiry at 00:05:14, BJC1_MJS remains silent, thus refusing to claim the turn after other-selection. BJC1_FBT takes the next turn again after a 4-second silent pause with 'oh, you've got <$H>Winebury</$H>', which is a case of Episode 6.

Episode 4 was observed in BJC1_FBT twice. In Transcript B.2-19 in the appendix, BJC1_FBT and BJC1_MJS are talking about the research questions of his MA dissertation. At 00:16:08, BJC1_MJS answers the question given by BJC1_FBT and closes his turn at 00:16:19, where BJC1_FBT utters a minimal response token *yeah.* After this, there is a silent pause for 2 seconds, and although BJC1_MBT might expect more explanation about the research questions by BJC1_MJS, he does not take the speaker turn and remains silent. BJC1_FBT then takes the next turn at 00:16:22. In the case of Episode 4 in BBC1_FBS, the two participants use eye contact to fill the pause and negotiate who is going to be the next speaker. In the case of BJC1_FBT, both the tutor and the student are looking down and checking the essay during the pause and no eye contact is observed.

5.2.4.2.6 Turn-structural episodes and listenership in BJC1_MJS

BJC1_MJS also used a range of turn-structural episodes as described in Table B.2.3-5 in the appendix. Episode 6, which is a case of retaining the turn after the partner refuses to accept the turn, is the only type that was not observed in BJC1_MJS. The means of listener status length in each episode cannot be guaranteed since the figures of their variances are more than 1000 or samples of the cases are very few. Samples of

BJC1_MJS's listener status in Episode 1, Episode 2, Episode 3, Episode 5 and Episode 7 were examined qualitatively.

Transcript 5.2-13 Episode 1 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 19: 54	FBT_F	0	-33		+ is your question really how frequently does the teacher +		
00: 19: 55		1	-32				
00: 19: 56		2	-31				
00: 19: 57		3	-30				
00: 19: 58		4	-29	HG			
00: 19: 59		5	-28	HG			
00: 20: 00		6	-27		+ ask individual students to answer		mhm.
00: 20: 01		7	-26	HG		HN	
00: 20: 02		8	-25				
00: 20: 03		9	-24	HG	+ questions in the classroom.		mhm.
00: 20: 04		10	-23			HN	
00: 20: 05		11	-22	HG	So if there're students sitting here are they more likely to be asked the questions +		mhm.
00: 20: 06		12	-21			HN	
00: 20: 07		13	-20	HG			
00: 20: 08		14	-19				
00: 20: 09		15	-18			HN	
00: 20: 10		16	-17				mhm.
00: 20: 11		17	-16		+ than somebody sitting in the corner +	HN	
00: 20: 12		18	-15	HG			mhm.
00: 20: 13		19	-14			HN	
00: 20: 14		20	-13		+ is this about teacher nominating people to answer +		
00: 20: 15		21	-12	HG			
00: 20: 16		22	-11	HG			
00: 20: 17		23	-10				mhm.
00: 20: 18		24	-9		+ or is there er er more active interaction where students themselves +		
00: 20: 19		25	-8				
00: 20: 20		26	-7	HG			
00: 20: 21		27	-6				
00: 20: 22		28	-5	HG			
00: 20: 23		29	-4		+ will will say "hey I want to say something".		mhm.
00: 20: 24		30	-3	HG			
00: 20: 25		31	-2	HG			
00: 20: 26		32	-1				mhm.
00: 20: 27	MJS_F	-19	0				Oh my hypothesis is erm it's difficult to er have opportunities to speak in English for students in the large class and then especially erm interaction of between teacher and individual student.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript 5.2-13 above is a sample of episode 1 in BJC1_MJS. BJC1_MJS

utters a minimal response token *mhm* eight times and uses HNs six times in about 30 seconds of listener status, which is different from the cases of Episode 1 in the tutors examined in the previous sections. As mentioned, this monotonous use of verbal response tokens distinguishes the Japanese listeners from the British listeners. About 10 seconds before BJC1_MJS takes the floor, he reduces the use of verbal response tokens and HNs, which can be taken as an implication of a speaker change.

Transcript 5.2-14 Episode 2 and Episode 3 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 17: 40		2	-15		Okay so okay from this I have a picture of Japanese class working very quietly.		
00: 17: 41		3	-14				
00: 17: 42		4	-13				
00: 17: 43		5	-12				
00: 17: 44		6	-11				
00: 17: 45		7	-10	HT/HG			
00: 17: 46		8	-9				
00: 17: 47		9	-8	HG		HN	mhm.
00: 17: 48		10	-7			HN	
00: 17: 49		11	-6		Is that right?	HN	
00: 17: 50		12	-5		Teacher says "okay open your book and do exercise three."		
00: 17: 51		13	-4	HG			
00: 17: 52		14	-3				
00: 17: 53		15	-2				Y= yeah.
00: 17: 54		16	-1				
00: 17: 55	MJS_F	-3	0		And they are=		Yeah I have experience like that.
00: 17: 56		-2	1	HG			
00: 17: 57		-1	2				
00: 17: 58	FBT_F	0	-1	HG	So there is no speaking.	HG	
00: 17: 59	MJS_F	-25	0			HG	No especially i= if students want to say something er they can do. But normally er just teacher says something.
00: 18: 00		-24	1			HN	
00: 18: 01		-23	2				
00: 18: 02		-22	3			HG	
00: 18: 03		-21	4				
00: 18: 04		-20	5	HN		HG	
00: 18: 05		-19	6			HG	
00: 18: 06		-18	7	HN			
00: 18: 07		-17	8				
00: 18: 08		-16	9				
00: 18: 09		-15	10			HG	
00: 18: 10	Pause	-14	-2		<$E> pause </$E>		
00: 18: 11		-13	-1				
00: 18: 12	MJS_F	-12	0				So if teacher er point out to a student maybe the student have opportunity to speak English.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

There were also some similarities between Japanese students and British students. Although the cases were few, BJC1_MJS also used Episode 2, a cut-in, to show his contribution to the conversation. For example, at 00:17:52 in Transcript 5.2-14 above, BJC1_MJS utters *yeah* and takes a speaker turn 2 seconds after. This is an example of Episode 2 in BJC1_MJS. At 00:17:58, BJC1_MJS takes a speaker turn by cutting-in again. He uses HGs to secure the floor this time, and the use of HGs at TTP was also observed in the cases of Episode 2 in the British students. In this example, BJC1_MJS successfully leads a shift from a commentary framework to an explanatory framework. At 00:18:10, there is a 2 seconds of silent pause, which can be a possible point of speaker change; however, BJC1_MJS retains the next speaker turn at 00:18:12. This can be categorised into Episode 3.

Episode 5 was observed in BJC1_MJS most frequently. As described in the cases of BBC1_FBS and BBC2_MBS, the use of response tokens in the students' listener status decreased before turn-taking in Episode 5. This tendency was also observed in BJC1_MJS. As shown in Transcript B.2-20 in the appendix, he utters a minimal response token *mhm* only once before he takes a speaker turn at 00:28:50 after BJC1_FBT's inquiry.

An example of Episode 7 in BJC1_MJS is described in Transcript B.2-21 in the appendix. In the transcript, BJC1_FBT and BJC1_MJS talk about his dissertation topic, which concerns interactions between a tutor and students in English classes in Japan. BJC1_FBT has asked, 'But do they [students] not do that [speaking out in class] because they're working quietly', at 00:22:24, and there is a 3-second silent pause after BJC1_FBT's inquiry. BJC1_MJS first disclaims the next speaker turn, but BJC1_FBT also disclaims the turn. Finally, BJC1_MJS finally takes the speaker turn

at 00:22:31. This kind of negotiation on speaker selection was also observed in BBC1_FBS.

5.2.4.2.7 Turn-structural episodes and listenership in BJC2_MBT

Five turn-structural episodes appeared in BJC2_MBT as described in Table B.2.3-6 in the appendix. From the figures of SD and variances, the means of listener status length in Episode 2, Episode 5 and Episode 6 in BJC1_MBT can be validated, although the samples in these cases are few.

In an example of Episode 1 in BJC2_MBT shown in Transcript B.2-22 in the appendix, three variations of verbal response tokens can be observed, such as minimal response tokens *yeah* and *mm*, and a cluster *mm yeah*. A collocation of a verbal response token *mm* with a HN occurs once at 00:31:15, and, including this, HNs are used five times in his listener status. He utters a cluster *mm great yeah* at TTP with HNs. The frequent use of *mm* shows a similarity with BBC2_MBS's use of response tokens, whereas the less frequent use of *okay* shows a difference from the other tutors. After BJC2_MBT takes a speaker turn at 00:31:35, he comments on BJC2_MJS's previous utterances and then moves onto the next topic. This speaker change leads both a topic change and a shift from an explanatory framework to a commentary framework.

The cases of Episode 2 in BJC2_MBT showed similarities with BBC1_MBT and BBC2_MBT. Although only one of the two types in Episode 2 in the tutors, namely initiation of a framework shift and cut-in for clarification, was observed in BJC1_FBT, BJC2_MBT had both of the two strategies in Episode 2. In Transcript 5.2-15 below, for example, the first pattern of Episode 2 can be observed. BJC2_MBT utters a minimal response token *yeah* with a HN four times in 18 seconds of listener

status to show his engagement to the conversation. BJC2_MBT's floor-taking at 00:37:30 leads a transition from an explanatory framework to a commentary framework within the same topic. In the student's explanatory framework, BJC2_MBT also used a cut-in to encourage BJC2_MJS's speaking as shown at 00:02:04 in Transcript 5.2-16 below. This can be classified as the second strategy of Episode 2 and recognised as *cut-in for clarification*.

Transcript 5.2-15 Episode 2 in BJC2_MBT (1)

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 37: 11	MJS_F	−18	0				+ So if it's possible I will bring for the next meeting +
00: 37: 12		−17	1				
00: 37: 13		−16	2			HG	
00: 37: 14		−15	3				
00: 37: 15		−14	4				
00: 37: 16		−13	5			HG	
00: 37: 17		−12	6	HN	Yeah.		
00: 37: 18		−11	7			HG	+ the chapter of <$G?> +
00: 37: 19		−10	8	HN	Yeah.		
00: 37: 20		−9	9				+ and if possible I will rewrite the chapter +
00: 37: 21		−8	10			HG	
00: 37: 22		−7	11				
00: 37: 23		−6	12				
00: 37: 24		−5	13				
00: 37: 25		−4	14	HN	Yeah.		
00: 37: 26		−3	15			HG	+ add some erm sections +
00: 37: 27		−2	16				
00: 37: 28		−1	17	HN	Yeah.	HG	
00: 37: 29	MBT_F	0	−72		Yeah yeah that's that's fine.		+ to clarify the position.
00: 37: 30		1	−71			HN	
00: 37: 31		2	−70	HG	I mean I think I mean I mean I mean. There's not this this can change all the time.+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript 5.2-16 Episode 2 in BJC2_MBT (2)

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 02: 03	MJS_F	−1	0			HG	Chapter one is the theoretical
00: 02: 04	MBT_F	0	−1		You've written theoretical yeah.		
00: 02: 05	MJS_F	−6	0	HG		HG	Chapter two will be <$G?>.
00: 02: 06		−5	1				
00: 02: 07		−4	2				
00: 02: 08		−3	3	HN		HG	

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Episode 3 appeared in BJC2_MBT five times in total. In Transcript B.2-23 in the appendix, BJC2_MBT closes his speaker turn at 00:27:56 and gets into a 4-second silent pause. This pause can be a possible speaker change point, although BJC2_MJS disclaims the speaker turn. BJC2_MBT then takes the next speaker turn at 00:28:01 and continues his explanation in order to make BJC2_MJS understand what he has meant.

Episode 6 was also used three times in BJC2_MBT. In Transcript B.2-24 in the appendix, BJC2_MBT tries to let BJC2_MJS explain about the concept of 'performance staging and participation' with implicit inquiries given at 00:26:33 and at 00:26:39. However, BJC2_MJS disclaims BJC2_MBT's turn giving and a 2 second pause follows. After the pause, BJC2_MBT utters an explicit inquiry, 'What do you think?' which leads to the student's speech. This is an example of Episode 6 in BJC2_MBT.

Although BBC1_MBT's Episode 3 was observed 14 times in total, frequent use of Episode 3 and Episode 6 by the tutors can be one of the characteristics in the British-Japanese conversations, where negotiation of a speaker change was observed more often than in the British-British conversations.

5.2.4.2.8 Turn-structural episodes and listenership in BJC2_MJS

Most of the cases in BJC2_MJS were categorised into three turn-structural episodes, Episode 1, Episode 2 and Episode 5, which is close to BBC2_MBS. As shown in Table B.2.3-7 in the appendix, the means of listener status length in each episode cannot be reliable since their variances are larger than 1000. As examined in Section 5.1.3, BJC2_MJS used HNs more than 350 times in total within the 39-minute length conversation.

Transcript 5.2-17 below shows an example of Episode 1 in BJC2_MJS. There are no verbal response tokens uttered, however, and HNs are observed 6 times in 33 seconds of listener status. BJC2_MJS constantly gives HNs through the conversation, which is not observed in the British-British conversations.

Transcript 5.2-17 Episode 1 in BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 36: 04	MBT_F	0	−33		And then <$H> lessing </$H> yes absolutely.		
00: 36: 05		1	−32				
00: 36: 06		2	−31		Erm up to you when you wanna how you wanna do this.+		
00: 36: 07		3	−30				
00: 36: 08		4	−29	HG			
00: 36: 09		5	−28				
00: 36: 10		6	−27				
00: 36: 11		7	−26		+I mean you either you could I mean you could do overall revisions now you just clarify what you say here +		
00: 36: 12		8	−25				
00: 36: 13		9	−24	HG			
00: 36: 14		10	−23				
00: 36: 15		11	−22			HN	
00: 36: 16		12	−21	HG			
00: 36: 17		13	−20			HN	
00: 36: 18		14	−19	HG	+ or you could go straight to the next draft on the on the chapter wait put these original and think about this longer +		
00: 36: 19		15	−18				
00: 36: 20		16	−17	HG			
00: 36: 21		17	−16				
00: 36: 22		18	−15	HG		HN	
00: 36: 23		19	−14				
00: 36: 24		20	−13	HG	+ but it's up to you. I mean you might just get it done now and start a fresh chapter <$G?> colour+		
00: 36: 25		21	−12			HN	
00: 36: 26		22	−11	HG			
00: 36: 27		23	−10	HG		HN	
00: 36: 28		24	−9				
00: 36: 29		25	−8	HG			
00: 36: 30		26	−7			HN	
00: 36: 31		27	−6				
00: 36: 32		28	−5	HG	+and them move on to the theories into into put behind the text erm.		
00: 36: 33		29	−4				
00: 36: 34		30	−3	HG			
00: 36: 35		31	−2				
00: 36: 36		32	−1				
00: 36: 37	MJS_F	−32	0				Erm the next meeting will be tenth of August +

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Episode 2 in BJC2_MJS also differed from the other students' cases. In the

cases of Episode 2 in BBC1_FBS, BBC2_MBS and BJC1_MJS, the students tried to show their contribution to the conversation by cutting-in during the tutor's commentary framework. On the other hand, BJC2_MJS used a cut-in to go back to the previous topic, which BJC1_MJS wanted to clarify with his tutor.

Transcript 5.2-18 Episode 2 in BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 13: 53		18	−18		+ John Jennet has got some erm Jo= yeah Jo= John Jennet yeah erm Gerald Jennet. He's got some erm he's got some <$G?> looking forward <$G?>.		
00: 13: 54		19	−17				
00: 13: 55		20	−16				
00: 13: 56		21	−15				
00: 13: 57		22	−14				
00: 13: 58		23	−13				
00: 13: 59		24	−12				
00: 14: 00		25	−11				
00: 14: 01		26	−10				
00: 14: 02		27	−9				
00: 14: 03		28	−8				
00: 14: 04		29	−7				
00: 14: 05		30	−6	HG		HN	
00: 14: 06		31	−5	SC/forehead			
00: 14: 07		32	−4				
00: 14: 08		33	−3				
00: 14: 09		34	−2				
00: 14: 10		35	−1				
00: 14: 11	MJS_F	−23	0			HG	So this er sorry.
00: 14: 12		−22	1		Yeah.		
00: 14: 13		−21	2	HG			It's better I clarify the chapter +
00: 14: 14		−20	3				
00: 14: 15		−19	4			HG	
00: 14: 16		−18	5	HN			
00: 14: 17		−17	6		Yeah.	HG	
00: 14: 18		−16	7			HG	
00: 14: 19		−15	8				+ it will be about anticipation +

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

At 00:14:11 in Transcript 5.2-18 above, BJC2_MJS uses a cut-in and asks whether it is better to clarify a concept of 'anticipation' in the chapter, although BJC2_MBT has been suggesting a reference in his previous speaker turn.

Reduction in the use of response tokens before turn-taking in Episode 5 was observed in BBC1_FBS, BBC2_MBS and BJC1_MJS. However, the case of Episode

5 in BJC2_MJS seems to be an extreme case. As shown in Transcript B.2-25 in the appendix, BJC2_MJS does not use any verbal and visual response tokens before he takes the floor at 00:30:26 after BJC2_MBT's turn giving. Further discussion on this case is given in the next chapter.

5.3 Preferences in turn size and placement of response tokens

5.3.1 Preferences in turn size

As reviewed in Chapter Two, Sacks et al (1974) summarised the principles of conversation from their observation on conversation. One of the principles they offered is that 'turn size is not fixed, but varies'. From the results of the current research, however, an implication can be made that there seem to be *preferences* in turn length in a certain context. In this section, the time length of the particular participants' listener status will be described from the data observation using some referential statistics as supplemental data.

Transcript 5.3-1 Episode 1 in BBC1_MBT: Listener status length

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 11: 04	FBS_F	-10	0	SC/eye		HN	Yeah and there is a whole there's quite a lot of essays in the Andrew Ortony book about+
00: 11: 05		-9	1				
00: 11: 06		-8	2				
00: 11: 07		-7	3				
00: 11: 08		-6	4	HN			
00: 11: 09		-5	5				
00: 11: 10		-4	6	HN	Yeah that's right yeah.		+how you can use them for teaching+
00: 11: 11		-3	7				
00: 11: 12		-2	8	HN	Yeah.	HG	+you know+
00: 11: 13		-1	9	HN	Yeah.		+scientific concepts and stuff.
00: 11: 14	MBT_F	0	-23		Yeah yeah. The orthodoxy in cognitive linguistics is that similes and analogies are just expressions of conceptual metaphor mappings. So in other words+		

Keys: SC/eye = Self-comfort with eyes, HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

In the transcript of Episode 1 in BBC1_MBT above, BBC1_MBT is in listener status for 10 seconds before he takes the speaker turn at time 00:11:14. BBC1_MBT gives a cluster *Yeah that's right yeah* at time 00:11:10, which is -4 in his leadtime, and a minimal response token *Yeah* twice at time 00:11:12 and 00:11:13, which are -2 and -1 in his leadtime. The first response token is uttered after he is in listener status for 5 seconds.

In the transcript of Episode 2 in BBC1_MBT shown below, BBC1_MBT is in his listener status for 8 seconds, which is slightly shorter than the case of Episode 1 above. BBC1_MBT utters a cluster *Right yeah* at time 00:29:57, which is -5 in his leadtime, and a minimal response token *Okay* at 00:30:00, which is -2 in his leadtime.

Transcript 5.3-2 Episode 2 in BBC1_MBT: Listener status length

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 29: 54	FBS_F	-8	0			HN	Yeah I found that online I do= I don't +
00: 29: 55		-7	1				
00: 29: 56		-6	2				
00: 29: 57		-5	3		Right yeah.	HG	+ you can actually get hold of it online but
00: 29: 58		-4	4				
00: 29: 59		-3	5				
00: 30: 00		-2	6		Okay.		you have to subscribe to it or something
00: 30: 01		-1	7	SC/hair			
00: 30: 02	MBT_F	0	-9		How annoying. I wonder who owns it. I wonder if Vernon might take it.		

Keys: SC/hair = Self-comfort with hair, HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

From the comparison between these two occurrences and the data observation on other occurrences of Episode 1 and Episode 2 in BBC1_MBT, the tendency was observed that BBC1_MBT stays in listener status slightly longer in Episode 1 than in Episode 2. This means that BBC1_MBT listens to BBC1_FBS's speech longer in the instances of Episode 1 than Episode 2. Furthermore, in terms of the placements of

response tokens, BBC1_MBT utters the first verbal response token 5 seconds after BBC1_FBS takes the speaker turn in Episode 1. While BBC1_MBT utters the first verbal response tokens 2 seconds after BBC1_FBS takes the speaker turn in Episode 2. The relationship between the use of response tokens and turn sizes in these two episodes will be explored further in the next section.

In the transcript of Episode 1 in BBC2_MBT shown below, BBC2_MBT is in listener status for 19 seconds before he takes the speaker turn at 00:33:29.

Transcript 5.3-3 Episode 1 in BBC2_MBT: Listener status length

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 33: 10	MBS_F	−19	0				For fi= well I'm hoping c= certainly finish this erm month and <$H> would </$H> still combine the writing up of that with the erm connecting the the mental health lit review+
00: 33: 11		−18	1				
00: 33: 12		−17	2				
00: 33: 13		−16	3				
00: 33: 14		−15	4				
00: 33: 15		−14	5	HN			
00: 33: 16		−13	6			HG	
00: 33: 17		−12	7				
00: 33: 18		−11	8	HN			
00: 33: 19		−10	9				
00: 33: 20		−9	10			HG	
00: 33: 21		−8	11	HN			
00: 33: 22		−7	12		Okay. Right.		
00: 33: 23		−6	13			HG	+because that'll be as you say just so many hours available a a week+
00: 33: 24		−5	14				
00: 33: 25		−4	15				
00: 33: 26		−3	16			HG	
00: 33: 27		−2	17		Okay.		+and so the two will be interchangeable. Er+
00: 33: 28		−1	18			HG	
00: 33: 29	MBT_F	0	−59	HG	The the mental the mental health review. This is where you are going to do literature review in terms of mental health+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

BBC2_MBT utters a non-minimal response tokens *Okay Right* once at time 00:33:22, which is -7 in his leadtime, and a minimal response token *Okay* at 00:33:29, which is -

2 in his leadtime. As shown in the transcript, the length of listener status of Episode 2 in BBC2_MBT is about double than that in BBC1_MBT.

In order to describe this phenomenon of preference in turn size in particular turn structural episodes, some referential statistics, such as standard deviations and variances, were applied to the analysis. The cases of the participants' turn-structural episodes described above, Episode 1 in BBC1_MBT, Episode 2 in BBC1_MBT and Episode 2 in BBC2_MBT, were examined alongside the statistics. These episodes have more than 20 instances and their standard deviations are less than 13 as listed in Table 5.3.1-1 below.

Table 5.3.1-1 Preferences in listener status lengths

	No of samples	Listenership length		
		Mean	SD	Variance
BBC1_MBT's episode 1	33	8.97	7.68	58.94
BBC1_MBT's episode 2	55	7.65	9.39	88.15
BBC2_MBT's episode 2	20	16.60	12.66	160.24

Keys: SD = Standard Deviations

As far as I have reviewed there is no equivalent study where lengths of turns in a naturally occurring conversation have been measured by a timeline in a scientific way, so these figures shown in the table above cannot be validated with the existing research. However, from this current investigation, these three cases could be assumed as examples of preference in turn length by a particular participant in a particular context.

BBC1_MBT, for instance, tends to be in listener status about 9 seconds in Episode 1 on average, and the SD is 7.68. The instances shown in the previous

sections validate these statistic (see Transcript 5.2-8 and Transcript 5.3-1) The turn length of his listener status in Episode 2 is 7.65 seconds on average, and the SD is 9.39, which means he tends to be in listener status longer in Episode 1 than Episode 2 as observed in Transcript 5.3-2 (also see Transcript 5.2-9 and Transcript 5.2-10). In the case of BBC2_MBT's Episode 2, the average listener status length is 16.60 seconds, and the SD is 12.66. BBC1_MBT thus prefers shorter listener turns in Episode 2 than BBC2_MBT as observed in Transcript 5.3-3 (also see Transcript B.2-13 and Transcript B.2-14).

5.3.2 Preferences in placement

As shown in Transcript 5.3-1, Transcript 5.3-2 and Transcript 5.3-3 above, BBC1_MBT and BBC2_MBT use response tokens *Yeah, Okay* and *Right,* often as clusters or as minimal response tokens. In this section, frequencies in use of verbal response tokens were examined with the cases of BBC1_MBT's Episode 1, BBC1_MBT's Episode 2 and BBC2_MBT's Episode 2. Preferences in the placement of *yeah* in Episode 1 and Episode 2 in BBC1_MBT were also investigated in detail.

In order to describe the frequency of verbal response tokens used in these three cases, frequent word lists were extracted in Table B.3.1-1, Table B.3.1-2, and Table B.3.1-3 in appendices. Both in the case of BBC1_MBT's Episode 1 and Episode 2, *yeah* and *right* were shown as the top two in the lists of verbal response tokens, and a cluster, *yeah yeah,* follows the two minimal response tokens. In addition, response tokens expressing teachers' agreement to the students' explanation, such as *oh right* or *right yeah,* and response tokens showing their engagement, such as *oh god* or *oh hell,* were observed in Episode 1 and Episode 2. The former can be seen as convergence tokens or continuers depending on the context, whereas the latter can be

recognised as *exclaims* in Edmondson (1981), *strong emotional response* in Maynard (1989) or *engagement tokens* in O'Keeffe et al. (2007).

In BBC2_MBT's Episode 2, the minimal responses *sure* and *yeah* were ranked as top two on the verbal response token list. Minimal response tokens and discourse markers, such as *erm* and *well*, were observed in BBC1_MBT's Episode 2 and BBC2_MBT's Episode 2 above more often than BBC1_MBT's Episode 1. These items can be a part of convergent tokens and also function as a cue for challenge and disagreement to the previous student's utterance.

As described above, the response token *yeah* is the top in the frequency word lists in Episode 1 and Episode 2 in BBC1_MBT. Preferences in the placement of *yeah* in these two episodes were analysed from the data observation and the statistical analysis.

In the case of Episode 1 in BBC1_MBT shown in Transcript 5.3-1 in the previous section, BBC1_MBT utters a cluster *Yeah that's right yeah* 4 seconds before he takes the speaker turn. He also utters a minimal response token *yeah* twice 1 and 2 seconds before the TTP. When the first verbal response token is given, BBC1_MBT has been in listener status for 5 seconds. This is one of the occurrences of Episode 1 in BBC1_MBT where he listens to BBC1_FBS's speech for a while, and then gives the response token *yeah* in several forms, such as a minimal response token or a cluster, soon before he takes the speaker turn.

In the instance of Episode 2 in BBC1_MBT described in Transcript 5.3-2 in the previous section, BBC1_MBT utters a cluster *Right yeah* 5 seconds before he takes the speaker turn, and a minimal response token *Okay* 2 seconds before the TTP. BBC1_MBT has been listening to BBC1_FBS speaking only for 2 seconds before he

utters the first response token *Right yeah*. This instance shows that in Episode 2, BBC1_MBT utters the response token *yeah* soon after he is in listener status.

Another description of these preferences in the placement of *yeah* was provided as the statistic figures in reference to turn structural episodes shown in Table 5.3.2-1 below. More than 20 occurrences of the response token *yeah* were observed in BBC1_MBT's Episode 1 and BBC1_MBT's Episode 2. This is the reason why the placements of *yeah* in these two episodes in BBC1_MBT were chosen for the analysis using a statistical approach.

Table 5.3.2-1 *Yeah* in Episode 1 and 2 in BBC1_MBT (original data)

	No of samples	Listenership length			
		Mean	SD	Variance	SE
BBC1_MBT's episode 1_yeah	22	-6.59	10.51	110.54	2,24
BBC1_MBT's episode 2_yeah	25	-8.56	8.05	64.84	1.61

Keys: SD = Standard Deviations, SE = Standard Errors

In the original data before adjustments, the number of samples of *yeah* in BBC1_MBT's Episode 1 was 22 and the average listener status length was -6.59. In BBC1_MBT's Episode 2, *yeah* was uttered 8.56 seconds before floor-taking in average, which was longer than his case of Episode 1. To make the figure more articulate, the instances that varied by more than or less than 10 seconds from the average were eliminated. The figures after the adjustments were described in Table 5.3.2-2 below.

Table 5.3.2-2 *Yeah* in Episode 1 and 2 in BBC1_MBT (adjusted data)

	No of samples	Listenership length			
		Mean	SD	Variance	SE
BBC1_MBT's episode 1_yeah	20	-3.45	2.50	6.26	0.56
BBC1_MBT's episode 2_yeah	21	-5.81	4.99	24.86	1.09

Keys: SD = Standard Deviations, SE = Standard Errors

The standard errors (SE, hereafter) in BBC1_MBT's Episode 1 were reduced to less than one and about one in Episode 2 in the figures after the adjustment. Although it is said that the SE in scientific experimental research should be less than 0.05 or even smaller in an experimental research, this preliminary human scientific study has no equivalent study to verify to what extent these figures are reliable at this stage.

From the SD and the variances, hence, these two average placements of *yeah* can be taken as meaningful scores. In BBC1_MBT's Episode 1, *yeah* is uttered 3.45 seconds before floor-taking in average. The placement of *yeah* in BBC1_MBT's Episode 1 was closer to the TTP than *yeah* in BBC1_MBT's Episode 2 although the listener status length in BBC1_MBT's Episode 2 was shorter than BBC1_MBT's Episode 1. This means that in Episode 2, BBC1_MBT started using the verbal response token *yeah* soon after achieving listener status than Episode 1.

This preliminary research with a small data set is not able to unveil the placements of many response tokens and cannot generalise the patterns in the use of response tokens fully. However, these descriptions indicate that placements of a particular response token seem to be closely related to turn management. The length of a speaker turn seems not to be defined only by the speaker's intention to keep or give the floor, but also by negotiation or co-construction of the turn structure with the listener. The relationship between listenership length and the turn-structural episodes

in a particular context were highlighted as one of the critical areas to be explored in future research.

5.4 Salient findings from the main study

5.4.1 Findings from the global pattern analysis

To summarise, findings from the global pattern analysis and the turn-structural analysis will be reviewed. In particular, five findings from the global pattern analysis in the main study can be described as follows:

1. The tutors take the floor of the conversation more than the students. The tutors uttered more words than the students. The tutors held their speaker status more than the students. These tendencies were observed both in the British-British conversations and the British-Japanese conversations.
2. Contribution of each participant to conversation was more equal in the numbers of floor-taking and words uttered in the British-British conversations than in the British-Japanese conversations.
3. In terms of the use of visual response tokens, HGs were often observed *at* TTP and HNs were observed soon *before* TTP. This tendency was more evident in the British-British conversations than in the British-Japanese conversations. The students tended to use HNs when they were in listener status more than the tutors in both the British-British conversations and the British-Japanese conversations.
4. In terms of the use of verbal response tokens, there were differences between the tutors and the students. There were tendencies that the tutors used *erm* in their speaker status more than the students, and that the tutors

used *yeah* at TTP more than the students while the students used *yeah* at their listener status more than the tutors. These tendencies were observed both in the British-British conversations and the British-Japanese conversations.

5. The tutors rarely used *mm* and *mhm* in the conversations, while particular students, namely BBC2_MBS and BJC1_MJS, used *mm* or *mhm* quite often in their listener status.

Aside from these findings, it is worth noting that two conversational styles can be recognised from the average length of time that the participants held the floor in the British-British conversations: (1) longer turn conversation and (2) shorter turn conversation. In a longer turn conversation, the tutor held the floor for about 40 to 50 seconds at a time, and the student held the floor for about 20 seconds at a time. A shorter turn conversation was characterised by short turn exchanges, where the tutor had about 13 seconds long floor and the student had about 7 seconds long floor. BBC1 and BJC1 share the second conversational style, and BBC2 and BJC2 share the first conversational style. Any findings cannot be justified fully from the current study; however, this might be related to age differences and the power relationships between the participants rather than participants' native languages. These features will be taken into consideration in the discussion in the next chapter along with underpinning theories.

Another issue raised through the main study is that of silent pauses. A notion that there are more pauses in the British-Japanese conversations than in the British-British conversations has been excluded from the findings since even in these two

British-Japanese conversations, the amount of silent pauses varies.

5.4.2 Findings from the turn structure analysis

Salient findings from the turn structure analysis in the main study are summarised as follows:

- Aspect 1: Focusing on verbal response tokens
  1. Regarding forms of verbal response tokens, the tutors in the British-British conversations used non-minimal response tokens more than in the British-Japanese conversations. The British students uttered non-minimal response tokens and clusters more than the Japanese students.
  2. Changing items and forms of verbal response tokens seemed to function as a sign to initiate a next turn exchange. This tendency was evident in the British participants more than the Japanese students.
- Aspect 2: Focusing on verbal response tokens with head nods
  1. Changing frequency of the use of verbal and response tokens and HNs seemed to function as a sign to initiate a next turn exchange. Two-ways of changing frequency were observed:

     Strategy 1: Increasing frequency of response tokens

     Strategy 2: Decreasing frequency of response tokens

     These tendencies were observed both in the British-British conversations and in the British-Japanese conversations. Strategy 1 was observed in listenership followed by a turn-taking or cut-in by self-selection with engagement.

Strategy 2 was observed in the tutors' listenership followed by a cut-in for clarification, and the students' listenership followed by the partner's turn giving.

- Aspect 3: Focusing on hand gestures

1. In terms of the use of hand gestures (HGs), multi-functional nature of HGs was recognised at TTP in both the British and Japanese participants. There seemed to be three types of HGs used in TTP:

Type 1: HGs function as initiating a speaker turn

Type 2: HGs function as initiating a speaker turn and 'deictic'

Type 3: HGs function as initiating a speaker turn and 'metaphoric'

Type 1 has one layer in function, namely turn management to initiate a speaker. Type 2 and Type 3 have two layers; turn management plus other functions, such as pointing to some objects or conveying some meaning accompanying speech.

- Aspect 4: Turn-structural episodes

1. In terms of turn-structural analysis, Episode 1 (A's turn closing→B's turn-taking) and Episode 2 (A's turn keeping → B's cut-in) were observed more often in the tutors in the British-British conversations than the tutors in the British-Japanese conversation and the students.
2. The tutors' cases of Episode 2 (A's turn keeping → B's cut-in) seemed to be related to the length of the previous listener status and the particular use of response tokens. The tutors' use of Episode 2 in an explanatory framework

functioned to invoke students' further explanation. The tutors' use of Episode 2 in a commentary framework tended to lead a topic change or a boundary of frameworks.

3. Frequent use of Episode 5 (A's turn giving → B's turn-taking) was a shared tendency among the students both in the British-British and the British-Japanese conversations.
4. Episode 3 (A's turn closing → B's turn refusal → A's turn retaining) was shown in BBC1_MBT and BJC1_FBT more often than other tutors and students.
5. Episode 6 (A's turn giving → B's turn refusal → A's turn retaining) was observed in BJC1_FBT more often than other tutors and students.
6. BJC1_FBT's use of Episode 3 (A's turn closing → B's turn refusal → A's turn retaining) and Episode 6 (A's turn giving → B's turn refusal → A's turn retaining) also seemed to be related to the length of the previous listener status and the use of response tokens.
7. Frequent use of these two episodes, namely Episode 3 and Episode 6, can be thought of as a phenomenon of negotiation for speaker change and a framework shift. There seemed to be a gap between the tutor and student about expectation for a transition from a commentary framework to an explanatory framework in these cases.

Apart from these main findings of the turn-structural analysis, there are three subsidiary findings which are worth noting. Firstly, some differences in the use of response tokens between the tutors and the students were observed:

1. Response tokens which function as a compliment, such as *excellent* and *good*, were uttered only by the tutors.
2. Clusters with *right*, such as *that's right* and *right okay*, were observed in the tutors more than the students.
3. The convergence response token *yes* was used in the longer turn conversations, namely BBC2 and BJC2 more than the shorter turn conversations, namely BBC1 and BJC1.

Secondly, through the course of the conversation, a phenomenon of sharing response tokens between the participants was observed. When a tutor uttered *sure*, for instance, the student in the conversation started using *sure*. Thirdly, a Japanese discourse marker *aa:* was observed several times in the Japanese students' listener status. A minimal response token *aa:* is an equivalent of *oh* in English and functions as change-of-status token in Japanese (Mori 2002) as reviewed in Section 2.3.4.

5.5 Summary

In this chapter, the forms and functions of response tokens in the British-British conversations and the British-Japanese conversations have been compared. Two approaches have been taken; global pattern analysis and turn structure analysis. In the global pattern analysis, general figures in the conversation data, such as the numbers of words and the length of speaker status in each participant, were compared. In addition, placements of verbal and visual response tokens were investigated with leadtime. In the turn structure analysis, seven turn-structural episodes have been applied, which have been developed based on Sacks (1974) and Ohama (2006). The

turn-taking patterns examined in the pilot study of this project have also been applied. In addition, the relationships between the turn-structural episodes and the listener status length were analysed. The use of response tokens in each turn-structural episode were described with transcripts.

Some differences have been identified in the use of cutting-in and the frequency of negotiation for speaker change between the British-British conversations and the British-Japanese conversations. It was recognised that some turn-structural episodes may be closely related to a shift of discourse frameworks and topic changes. In the next chapter, theoretical interpretations will be made on the participants' choice in the use of turn-structural episodes and response tokens observed in the main study.

Chapter 6 Discussion and implications

6.0 Introduction

The discussion of the results of the main study will be followed up in this chapter. The participants' listenership behaviour will be analysed from an interlanguage pragmatic and intercultural communication aspect. The rationales behind the participants' choice of turn-structural episodes and the use of response tokens will be deliberated over in relation to the cultural values which the participants' discourse communities hold. In addition, the implications over the monotonous use of response tokens and frequent occurrences of negotiation of speaker changes in the British-Japanese conversations, particularly in relation to cultural backgrounds and speaker roles, will be considered in comparison with the British-British conversations. The expectations of discourse frameworks in academic tutorials, the power relationship between a tutor and a student, and the concept of multiple identities will also be taken into consideration.

6.1 Context and identity in conversation

6.1.1 Constructing a context in intercultural communication

As reviewed in Section 2.3.4, White (1989) and Maynard (1997a) studied Japanese learners' response tokens in English in comparison with native speakers of American English, and concluded that Japanese learners tend to use more response tokens than native speakers. Some attempts have been made to investigate the reason why Japanese learners use more response tokens. For example, Ohama (2006) compared turn-taking patterns between Japanese and English in New Zealand in reference to response tokens, and concluded that frequent use of response tokens in Japanese conversation reflects the Japanese speaker's presumption about speaker and listener

roles. This is because, in Japanese conversation, the speaker is expected to have a dominant role in conversation and build the foundation by him/herself, while the listener encourages the speaker to do so implicitly by giving continuer response tokens. In English conversation, on the other hand, the speaker and the listener contribute to creating the foundation of the conversation together, and the listener is expected to show *aggressive listenership* to co-construct the context with the speaker.

Kita (1996) and Kita and Ide (2007) studied Japanese face to face conversation in reference to aizuchi, which refers to response tokens. The original meaning of aizuchi is 'hammering iron in turn when a sword master makes a Japanese sword' (Kita 1996: 62, translated by me). Kita highlighted the importance of creating 'a constant turn-taking rhythm' in Japanese conversation, thus:

> In Japanese conversation, participants are expected to create a constant turn-taking rhythm by response tokens. Even when a listener has nothing to contribute to the conversation, he or she can send 'meaningless' response tokens to keep the rhythm. [...] When the conversation is not elaborated by either the speaker or the listener, the participants can fill the 'ma', which are silent pauses, with response tokens.
>
> (ibid)

From the findings of the main study, BJC1_MJS's monotonous use of *mhm* and BJC2_MJS's continuous use of HNs have been recognised. When BJC1_MJS's total use of *mhm* is compared with BBC2_MBS's use of *mm*, there seems not to be any

significant differences between these two students. The forms and placements of *mhm* in BJC1_MJS, however, differ from *mm* in BBC2_MBS, where more clusters of response tokens are observed and the varieties of forms are plotted in his listener status in reference to TTP. The frequent use of *mhm* and HNs in the Japanese students can be interpreted as activation of pragmatic strategies of Japanese conversation.

At the level of *frame* (Ager 1994a, 1994b), *a structure of expectation* (Kramsch 1998) or *schemata* (Cook 1994), different expectations of university tutorials between Japan and the UK were reported by Turner and Hiraga (1996). This layer of analysis can be related to *speech events* in Hymes (1972) and *social events* in Fairclough (2003). Through a contrastive study between British tutor–British student conversations and British tutor–Japanese student conversations in university tutorials, Turner and Hiraga found that there are differences in *strategies of elaboration* between British students and Japanese students, and raised the notion that Japanese tutorials are *knowledge-based* while British tutorials are *thinking-based*.

> While British academic culture is predominantly thinking-centred, valuing the process of critical appraisal by means of such analytical strategies as comparison, evaluation, and probing further, Japanese academic culture is predominantly knowledge centred, valuing the demonstration of knowledge gained by following the correct procedures in adequate detail and technique.
>
> (Turner and Hiraga 1996: 139)

This current study seems to support this summary, for it shows that Japanese students

tend to listen to their tutors' commentary while British students actively elaborate their thought in tutorials. Frequent refusals for a shift away from the tutor's commentary framework to the students' explanatory framework in the British-Japanese conversations support Turner and Hiraga's notion of different frames between Japanese and British tutorials. At certain points, however, Japanese students also elaborate their own opinions in a more aggressive way in the British-Japanese conversations. Multiple identities of Japanese students' attitudes in the British-Japanese tutorials will be discussed further in the next section.

6.1.2 Constructing social identities through intercultural communication

How the Japanese students present themselves in the academic tutorials, especially in their listenership, is one of the issues raised in this research. Since the two Japanese students in the current research data, namely BJC1_MJS and BJC2_MJS, grew up in Japan and had stayed in the UK for about a year when the data was recorded, they are rooted in Japanese culture fundamentally. From a close observation of their listenership patterns, both similarities and differences in the Japanese students' choice and placement of response tokens with the British students were recognised. As shown in Transcript 6.1-1 and Transcript 6.1-2 below, the monotonous use of response tokens *mhm* in BJC1_MJS and HNs in BJC2_MJS were recognised. To construct a harmonious rhythm with speaker turns and listener turns, which is a shared strategy in Japanese conversation as described in Kita (1996), the Japanese students keep sending monotonous response tokens at a consistent pace. In BJC1_MJS's case, he has chosen to use an English response token *mhm* while still maintaining a listenership strategy from his native language. The Japanese students' monotonous

use of response tokens seems to have unconsciously emerged without listeners' strong control as Coupland (2007) observes (see Section 2.4.2.3). This can be a representation of a fragment of the Japanese students' multiple identities.

Transcript 6.1-1 Sample 1: Episode 1 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 19: 54	FBT_F	0	-33		+ is your question really how frequently does the teacher +		
00: 19: 55		1	-32				
00: 19: 56		2	-31				
00: 19: 57		3	-30				
00: 19: 58		4	-29	HG			
00: 19: 59		5	-28	HG			
00: 20: 00		6	-27		+ ask individual students to answer		mhm.
00: 20: 01		7	-26	HG		HN	
00: 20: 02		8	-25				
00: 20: 03		9	-24	HG	+ questions in the classroom.		mhm.
00: 20: 04		10	-23			HN	
00: 20: 05		11	-22	HG	So if there're students sitting here are they more likely to be asked the questions +		mhm.
00: 20: 06		12	-21			HN	
00: 20: 07		13	-20	HG			
00: 20: 08		14	-19				
00: 20: 09		15	-18			HN	
00: 20: 10		16	-17				mhm.
00: 20: 11		17	-16		+ than somebody sitting in the corner +	HN	
00: 20: 12		18	-15	HG			mhm.
00: 20: 13		19	-14			HN	
00: 20: 14		20	-13		+ is this about teacher nominating people to answer +		
00: 20: 15		21	-12	HG			
00: 20: 16		22	-11	HG			
00: 20: 17		23	-10				mhm.
00: 20: 18		24	-9		+ or is there er er more active interaction where students themselves +		
00: 20: 19		25	-8				
00: 20: 20		26	-7	HG			
00: 20: 21		27	-6				
00: 20: 22		28	-5	HG			
00: 20: 23		29	-4		+ will will say "hey I want to say something".		mhm.
00: 20: 24		30	-3	HG			
00: 20: 25		31	-2	HG			
00: 20: 26		32	-1				mhm.
00: 20: 27	MJS_F	-19	0				Oh my hypothesis is erm it's difficult to er have opportunities to speak in English for students in the large class and then especially erm interaction of between teacher and individual student.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript 6.1-2 Sample 2: Episode 1 in BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 36: 04	MBT_F	0	−33		And then <$H> lessing </$H> yes absolutely.		
00: 36: 05		1	−32				
00: 36: 06		2	−31		Erm up to you when you wanna how you wanna do this.+		
00: 36: 07		3	−30				
00: 36: 08		4	−29	HG			
00: 36: 09		5	−28				
00: 36: 10		6	−27				
00: 36: 11		7	−26		+I mean you either you could I mean you could do overall revisions now you just clarify what you say here +		
00: 36: 12		8	−25				
00: 36: 13		9	−24	HG			
00: 36: 14		10	−23				
00: 36: 15		11	−22			HN	
00: 36: 16		12	−21	HG			
00: 36: 17		13	−20			HN	
00: 36: 18		14	−19	HG	+ or you could go straight to the next draft on the on the chapter wait put these original and think about this longer +		
00: 36: 19		15	−18				
00: 36: 20		16	−17	HG			
00: 36: 21		17	−16				
00: 36: 22		18	−15	HG		HN	
00: 36: 23		19	−14				
00: 36: 24		20	−13	HG	+ but it's up to you. I mean you might just get it done now and start a fresh chapter <$G?> colour+		
00: 36: 25		21	−12			HN	
00: 36: 26		22	−11	HG			
00: 36: 27		23	−10	HG		HN	
00: 36: 28		24	−9				
00: 36: 29		25	−8	HG			
00: 36: 30		26	−7			HN	
00: 36: 31		27	−6				
00: 36: 32		28	−5	HG	+and them move on to the theories into into put behind the text erm.		
00: 36: 33		29	−4				
00: 36: 34		30	−3	HG			
00: 36: 35		31	−2				
00: 36: 36		32	−1				
00: 36: 37	MJS_F	−32	0				Erm the next meeting will be tenth of August +

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Although this needed to be clarified with the listeners' retrospective views ideally, retrospective interviews were not conducted in the current study, which is one of its limitations.

Negotiation for a speaker change is the other occurrence often observed in the British-Japanese conversations.

Transcript 6.1-3 Sample 3: Episode 3 and Episode 6 in BJC1_FBT

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 04: 56		20	-26		+ not criticise the teacher teaching +	HN	Ah.
00: 04: 57		21	-25				
00: 04: 58		22	-24	HG			
00: 04: 59		23	-23	HG			mhm.
00: 05: 00		24	-22		+ but the kind of recognise something which will enable me to develop my own teaching as the observer.	HN	
00: 05: 01		25	-21	HG			
00: 05: 02		26	-20				
00: 05: 03		27	-19			HN	
00: 05: 04		28	-18	HG			
00: 05: 05		29	-17				
00: 05: 06		30	-16			HN	
00: 05: 07		31	-15				mhm.
00: 05: 08		32	-14		You see that's a bit difference	HN	
00: 05: 09		33	-13	HG			
00: 05: 10		34	-12				Yeah.
00: 05: 11		35	-11		Erm.	HN	
00: 05: 12	Pause	-2	-10		<$E> pause </$E>		
00: 05: 13		-1	-9			HN	
00: 05: 14	FBT_F	0	-8		What else have you read about observation?		
00: 05: 15		1	-7				
00: 05: 16		2	-6				
00: 05: 17	Pause	-4	-5		<$E> pause </$E>		
00: 05: 18		-3	-4				
00: 05: 19		-2	-3				
00: 05: 20		-1	-2				
00: 05: 21	FBT_F	0	-1		Oh you've got <$H> Winebury </$H>.		
00: 05: 22	MJS_F	-2	0				Yeah I've got.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Episode 3 and Episode 6 are shown in Transcript 6.1-3 above and represent the turn-structural patterns where negotiation for a speaker change occurs. Although the negotiation for a speaker change is observed in the British-British conversations, the frequency of the occurrences is larger in the British-Japanese conversations. The sums of Episode 3 and Episode 6 in the tutors in the British-Japanese conversations are about 25 %, compared to 14 % in BBC1 and none in BBC2.

BJC1_FBT utters 'you see that's a bit different' at 00:05:08, which is a possible speaker change point. However, BJC1_MJS answers with a convergent response token only, *yeah* at 00:05:10, without taking the floor of the conversation. They exchange eye contact soon before BJC1_FBT looks down on the student's essay,

instead of waiting for BJC1_MJS to start speaking. BJC1_FBT then inquiries, 'what else have you read about observation?' Without the participants' retrospective views, only assumptions from the data observation can be made concerning these silent pauses. In the first pause, BJC1_MJS might just think it is appropriate to answer with a response token *yeah* since his tutor is giving him suggestions about his essay in a tutorial and does not expect him to give his opinion at this point, although a British student might add some comments on the tutor's previous speech in a similar situation. The knowledge-based tutorial style in Japan (Turner & Hiraga 1996) can be brought into the tutorial in the UK and affects BJC1_MJS's choice in turn exchanges.

BJC1_MJS's choice to refuse a speaker turn also seems to affect BJC1_FBT's presentation in the tutorial. BJC1_FBT used to teach English in Taiwan for many years and currently runs an MA in English Language Teaching at a university in the UK with many students from China, Taiwan, Korea and Japan. She can be assumed to have some understanding about cultures in East Asia. However, when there are silent pauses; BJC1_FBT tends to retain her turns with additional explanations or inquiries, or to return to her student's essay in order to find a cue to understand what the student thinks, instead of waiting for the student to speak.

In terms of the second pause after the tutor's inquiry, 'what else have you read about observation?' at 00:05:17, BJC1_MJS refuses the turn-giving and looks carefully at the lines of his essay which the tutor is reading. Again although this is only an assumption from the observation, some possible interpretations of the silent pause can be made. For instance, rather than listing the references he has read about observation, BJC1_MJS might choose to remain silent to try to anticipate what point BJC1_FBT is going to discuss next, or which of the references she expects him to

have. This can be interpreted as an emergence of a part of one identity, such as other-centred-ness, and the identity which is rooted in society rather than in the self (Maynard 1993, Yamada 1997) (see Section 2.4.1).

Transcript 6.1-4 Sample 4: Episode 2 and Episode 3 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 17: 40		2	-15		Okay so okay from this I have a picture of Japanese class working very quietly.		
00: 17: 41		3	-14				
00: 17: 42		4	-13				
00: 17: 43		5	-12				
00: 17: 44		6	-11				
00: 17: 45		7	-10	HT/HG			
00: 17: 46		8	-9				
00: 17: 47		9	-8	HG		HN	mhm.
00: 17: 48		10	-7			HN	
00: 17: 49		11	-6		Is that right?	HN	
00: 17: 50		12	-5		Teacher says "okay open your book and do exercise three."		
00: 17: 51		13	-4	HG			
00: 17: 52		14	-3				
00: 17: 53		15	-2				Y= yeah.
00: 17: 54		16	-1				
00: 17: 55	MJS_F	-3	0		And they are=		Yeah I have experience like that.
00: 17: 56		-2	1	HG			
00: 17: 57		-1	2				
00: 17: 58	FBT_F	0	-1	HG	So there is no speaking.	HG	
00: 17: 59	MJS_F	-25	0			HG	No especially i= if students want to say something er they can do. But normally er just teacher says something.
00: 18: 00		-24	1			HN	
00: 18: 01		-23	2				
00: 18: 02		-22	3			HG	
00: 18: 03		-21	4				
00: 18: 04		-20	5	HN		HG	
00: 18: 05		-19	6			HG	
00: 18: 06		-18	7	HN			
00: 18: 07		-17	8				
00: 18: 08		-16	9				
00: 18: 09		-15	10			HG	
00: 18: 10	Pause	-14	-2		<$E> pause </$E>		
00: 18: 11		-13	-1				
00: 18: 12	MJS_F	-12	0				So if teacher er point out to a student maybe the student have opportunity to speak English.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

In the same tutorials, the Japanese students present themselves with similar turn-taking strategies to the British students. As shown in Sample 4 in Transcript 6.1-4 above, BJC1_MJS uses cut-in in the same way the British students do. At this

moment, his voice is raised and the use of HGs is increased. It can be interpreted that a 'British' self in the Japanese student's identities is activated through the conversation. In this context, the British tutor and the Japanese student are discussing a quiet English class in Japan. This topic might invoke a memory when BJC1_MJS *was in* a quiet English class in Japan as an English teacher or as a student, which makes him see himself in the tutorial and be aware of his own attitude.

This is an example of multiple timescales and layered simultaneity in interaction. In addition, BJC1_MJS might think that he is more knowledgeable in this particular topic than the tutor. From the context, BJC1_MJS might deliberately choose to adjust his conversation style to that of native speakers, exercising strong control.

In the case of BJC2_MJS, there are some points where BJC2_MJS presents himself with an aggressive listenership, which is often observed in the British students and tutors. In sample 5 in Transcript 6.1-5 below, for example, BJC2_MJS uses cut-in after his listenership with limited response tokens. BJC2_MJS utters 'sorry' to start his speaking turn at 00:14:11, which is an apology for interrupting the tutor. BJC2_MJS starts a speaker turn without any comments on the previous tutor's utterance, which seems not likely to happen in British-British conversation. However, still BJC2_MJS has led the speaker turn with a discourse marker 'So' and a HG, which is a strategy of starting a speaker turn is often observed in the British-British conversations. BJC2_MJS's transactional goal in this cut-in is to confirm what the tutor meant about clarification in the essay before the tutor starts talking about the current topic.

Transcript 6.1-5 Sample 5: Episode 2 in BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 13: 53		18	−18		+ John Jennet has got some erm Jo= yeah Jo= John Jennet yeah erm Gerald Jennet. He's got some erm he's got some <$G?> looking forward <$G?>.		
00: 13: 54		19	−17				
00: 13: 55		20	−16				
00: 13: 56		21	−15				
00: 13: 57		22	−14				
00: 13: 58		23	−13				
00: 13: 59		24	−12				
00: 14: 00		25	−11				
00: 14: 01		26	−10				
00: 14: 02		27	−9				
00: 14: 03		28	−8				
00: 14: 04		29	−7				
00: 14: 05		30	−6	HG		HN	
00: 14: 06		31	−5	SC/forehead			
00: 14: 07		32	−4				
00: 14: 08		33	−3				
00: 14: 09		34	−2				
00: 14: 10		35	−1				
00: 14: 11	MJS_F	−23	0			HG	So this er sorry.
00: 14: 12		−22	1		Yeah.		
00: 14: 13		−21	2	HG			It's better I clarify the chapter +
00: 14: 14		−20	3				
00: 14: 15		−19	4			HG	
00: 14: 16		−18	5	HN			
00: 14: 17		−17	6		Yeah.	HG	
00: 14: 18		−16	7			HG	
00: 14: 19		−15	8				+ it will be about anticipation +

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

The previous topic, which concerns clarifying the concept of 'anticipation', reminds BJC2_MBT of a book by Gerald Jennet. The tutor is recommending the reference to the student in the transcript above while the previous topic about clarification of 'anticipation' makes BJC2_MJS think about the organisation and revisions of his essay. This situation, where an utterance invokes different issues in the participants' mind in a conversation, can occur in any conversation. However, how the participants treat this situation might be different according to the context. BJC2_MJS could wait for the end of the tutor's speaker turn or the end of the topic before giving the inquiry, but at this time, BJC2_MJS has cut-in to the tutor's previous speaker turn and starts asking about the revisions of the chapter with the concept of anticipation directly. From BJC2_MJS's attitude, it can be assumed that there might be a presumption of

BJC2_MJS that this kind of direct inquiry can be acceptable in English conversation in the context. This seems to work well although it might not be completely adjusted to a native speaker's method of turn-taking. For example, BJC2_MBT is a little startled when BJC2_MJS cuts-in the conversation, but apparently welcomes BJC2_MJS's inquiry and utters *yeah* at 00:14:12 with an encouraging voice and a HG, which signals 'go ahead'.

As discussed in this section, the Japanese students' multiple identities can be identified through their use of response tokens and turn-taking patterns. The relationship between the use of language and constructing social identities is described by Johnstone (2002) as follows:

> Ways of talking produce and reproduce ways of thinking, and ways of thinking can be manipulated via choices about grammar, style, wording, and every other aspect of language.
> (Johnstone 2002: 45)

A sentence from Ochs (2005) is also worth quoting:

> We recognize that social identities have a sociohistorical reality independent of language behaviour, but, in any given actual situation, at any given actual moment, people in those situations are actively constructing their social identities rather than passively living out some cultural prescription for social identity.
> (Ochs 2005:84)

Through the processes of representing themselves in ways of talking, including listenership behaviours and turn-taking patterns, interlocutors construct and reconstruct their social identities in conversation. The emergence of these identities seems rather arbitrary, but the context seems to be a catalyst or a critical determinant of the process. Which part of their multiple identities is invoked might depend on the particular moment in a particular context. At the same time the emergence of participants' identities and the related expressions create and recreate the context.

6.2 Acquisition of interlanguage pragmatics

6.2.1 Face and pragmatic failures

The issues raised in the current research, such as the frequent occurrence of the Japanese students' disclaiming the tutor's turn-giving, and the Japanese students' monotonous use of response tokens, can be categorised into participation domain in the politeness categories by Spencer-Oatey (2000). For the Japanese students in the current study, there might be two ways to be polite in the academic tutorials. One way is to follow the politeness rules which they have acquired from their own culture. The other way is to adjust themselves to the politeness system in the target culture, which they have acquired through language learning and their experiences in the target culture.

As reviewed in Section 2.4.2.2., Thomas (1983:101) distinguishes two types of pragmatic failures. Pragmalinguistic failure refers to the 'attitude of the speaker towards the information', such as grammar and lexical choices, is not understood by the hearer. Sociopragmatic failure, on the other hand, is a misunderstanding of the 'intended illocutionary force and/or attitude of the speaker to the hearer'. The latter is considered in the current study.

Nakane (2006) studied Japanese learners' rather quiet attitudes in English class conversations, and concluded that the Japanese learners' use of silence is their strategy to express a 'don't do the face-threatening act' message in order to avoid a shameful moment in which they say something wrong in class discussion. In this case, silence is used in order to save the students' own positive face. Her study also reported that this attitude is perceived by the lecturers negatively as Japanese students are not participating actively. Nakane's study can be taken as an 'incident' in intercultural communication since interlanguage users' utterances are interpreted by the other participants in conversation negatively because of differences in the quality of face (Spencer-Oatey 2000).

Although it can be only an assumption from the observation, the cases of the Japanese students' monotonous use of response tokens, the frequent occurrences of negotiation for speaker change and the much smaller numbers of utterances observed in the current study could be a possible sociopragmatic failure point in the interlanguage setting. Raising awareness of the differences in listenership behaviour between English conversation and Japanese conversation can be of some help to both learners of English and to native speakers in an interlanguage conversation. In order to avoid pragmatic failure, language learners can adjust listenership behaviour to target cultures with knowledge of conversational rules. However, pragmatic competence of successful language learners in interlanguage settings seems to be not just a matter of adjusting themselves to target cultures, but also necessitates being balanced between *accommodation* with target cultures and *reservation* of their own cultural values.

6.2.2 L1 transfer and activation of L2 pragmatic competence

Iwasaki (1997) analysed Japanese conversations recorded at UCLA talking about the earthquake that hit California, and found a phenomenon of the *loop sequence of backchannels* in Japanese conversation, which refers to 'successive exchanges of backchannel signals, and may be understood as a locally managed turn-taking pattern' (Iwasaki: 661). A sample of the loop sequence of backchannels is quoted in Transcript 6.2-1 below.

Transcript 6.2-1 The 'loop sequence of backchannels' in Iwasaki (1997:668)

EXCERPT 1 (JEQ#3)

	Yasumi (female)	*Sayuri (female)*
1	nn soo yuu koto de nigeru hitsuyoo	
2	wa nai tte yuu huu ni.	
3		nn
4	... omotta kara.	
5	s...(0.9) soo.	
6		**naruhodo nee.**
7	**n::.**	
8		nanka soo.
9		atashitachi wa nee.
10		(H) so uchi no .. tks< apaato mo kekkoo
11		hurui-n desu ne?
12		=chiku nannen ka wakannai-n desu kedo:*

English translation of Excerpt 1

	Yasumi (female)	*Sayuri (female)*
1	so there was no	
2	need to escape	
3		*nn*
4	... so I thought.	
5	...(0.9) so.	
6		**I see what you mean.**
7	**n::**	
8		like
9		we
10		... our apartment is also rather
11		old, you see.
12		= though I don't know how long since it was built.

In the English translation of Excerpt 1 above, Sayuri utters a slightly long response token, 'I see what you mean' at line 6, which is followed by Yasumi's minimal

response token ‘n::’ at line 7. This is a typical example of loop sequence of backchannels in Japanese conversation. Following the notion of the loop sequence in Iwasaki (1997), Kogure (2007) points out that the loop sequence of backchannels is used to maintain a speech stream in Japanese conversation, and nodding and smiling are also used as a part of this loop sequence. These features of Japanese response tokens have been observed in the Japanese students in the interlanguage conversations in the current research. This can be treated as L1 transfer in the British-Japanese conversations.

When and to what extent interlanguage users are able to or would like to accommodate to the target culture might be the next issue to be raised. Day (1998) examined conversations in two workplaces in Sweden and investigated the phenomenon of a ‘resistance’ to being a member of an ethnic group by referring to *membership categorization devices* (MCD). Five ways through which resistance to ethnic group membership are realised are described thus:

1. One can dismiss the relevance of the category;
2. One can minimize the supposed ‘difference’ between categories;
3. One can reconstitute the category so that one is excluded;
4. One can ethnify the Ethnifier;
5. One can resist Ethnification by actively avoiding it.

(Adapted from Day 1998:162-166)

In an intercultural context, participants are continuously choosing their presentation in interlanguage both consciously and unconsciously. In order to avoid incidents in

intercultural communication and to be better communicators, language learners need to present themselves in an 'appropriate' way in a particular context by maintaining a balance between adjusting themselves to the context and retaining their own values and entitlements.

Faerch and Kasper (1987: 112, emphasised by the authors) drew attention to L1 transfer of language learners in their interlanguage use, and attempted to investigate ' "where" learners transfer "what", "how much", "why", and "how" '. In addition, the concept of *foreigner role* was introduced by Faerch and Kasper (1987) as a protector of language learners. By marking 'non-membership in the L2 speech community', interlanguage users protect themselves from 'being assessed on the basis of native speaker norms and expectations' (ibid: 126).

Following Edmondson (1981), Fearch and Kasper (1984) defined two types of communicative knowledge; *declarative knowledge*, which is static knowledge about language and 'not related to specific communicative goals or to language use in real time', and *procedural knowledge*, which is dynamic choices of declarative knowledge for 'reaching specific communicative goals, observing constraints imposed by language processing in real time' (ibid: 125). The latter is closely related to acquisition of interlanguage pragmatics. Furthermore, declarative knowledge and procedural knowledge can be similar to the concepts of *communicative competence* and *strategic competence* in Dell Hymes (1972).

Alred, Byram and Fleming (2003: 3) defined the term *intercultural speaker* as 'the language learner who also acquires knowledge and skills of cultural mediation or interpretation, and not just a linguistic competence modelled on a native speaker'. Johnstone (2002) also expresses her attention to language learners' presentation of

themselves in social life as follows:

> Language ideology is of interest to students of language and to students of social life alike, because beliefs about what language is and how it works can affect languages as well as social relations among speakers.
> (Johnstone 2002: 56)

Kramsch (2008) raises the notion of a balance between language learners' approximation to the conversation styles in the target language, and how they present their own identities as a member of their own culture. This is described as follows:

> Symbolic competence is the ability not only to approximate or appropriate for oneself someone else's language, but to shape the very context in which the language is learned and used.
> (Kramsch 2008: 664)

Carter (2007) also highlights the multiple voices of language learners and the dichotomy between errors and the creative use of language in interlanguage:

> Both the external (target language) and the internal (first-language) voices are multiple, changing, and potentially both in conflict and in productive interchange with one another. Does creativity spring therefore from a co-constructed interactive tension between the two (or

> more) languages and identities and how far are the contexts involved wholly social in character? […]
>
> And another pertinent question is: who is responsible for accepting something as creative? Where are lines drawn between errors and creative uses of language by learners?
>
> (Carter 2007: 605)

In an interlanguage context, or even a monolinguistic context, language learners and their interlocutors, who are sometimes the native speakers of the target language and sometimes language learners with or without sharing the same native language, are creating a transactional and interactional context together through their use of language. At the same time, the context makes language learners and their interlocutors construct and reconstruct their identities which are assumed to be appropriate to the context. Sometimes they transpire to be a successful communicator in a context and sometimes they fail.

Even through the methods to be successful communicators have not been fully revealed, the patterns in the use of response tokens seem to be an important factor in participant's successful communication. In an interlanguage context, language learners seem to be in a situation where they are expected to make efforts, not only to adjust their use of response tokens to the way the native speakers of the target language use them, but also to express themselves as a member of their *own* culture in order to be a balanced interlanguage communicator.

6.3 Summary

Theoretical discussions and interpretations drawn from the results have been proposed in this chapter in reference to the use of response tokens and turn-taking patterns between the Japanese and the British students. Although the relationship between knowledge of a learner's native language and the target language, and the processes of activation of the knowledge in conversation is not a central issue in the current study, some implications have been drawn from the Japanese students' attitude observed in the research. Furthermore, the current research has highlighted both L1 transfer from Japanese, such as loop sequences of backchannels, the use of silence based on the concepts of other-centred-ness and differences in expectations of speaker roles between Japanese and British English, as well as aggressive adaptation to the pragmatic rules of the target language in the Japanese students in interlanguage conversation.

Chapter 7 Conclusion

7.0 Concluding overview

This thesis has investigated naturally occurring conversation placing focus on response tokens in a pedagogic setting using a time-related corpus approach. Based on the literature review in Chapter Two, conversation analysis using a time-related corpus was conducted as a pilot study as reported in Chapter Three. The key concepts of TTP and leadtime were introduced for the analysis. The link between the pilot study and the main study was then made in Chapter Four and the scope of the main study was defined with two data analysis approaches, namely the global pattern analysis and the turn structure analysis. By implementing the research method with the time-related spoken corpus data established in Chapter Three and Chapter Four, the main study reported the placements and forms of target response tokens in reference to leadtime and turn-structural episodes based on Ohama (2006), both quantitatively and qualitatively, in Chapter Five. That invoked some discussions on contextualisation and multiple identities observed in interlanguage pragmatics in Chapter Six.

No study to date has analysed these two areas, namely (1) establishing a new model for conversation analysis and (2) a preliminary research with the research method developed, which can be claimed as original contributions of the current study. In terms of the research method, analysing not only verbal but also visual response tokens with the concept of leadtime makes the current study unique. From this preliminary research, two discourse frameworks, the tutor's commentary framework and the student's explanatory framework, were recognised in the tutor-student academic tutorial sessions. The forms and placements of verbal and visual response

tokens observed in the participants' listenership seemed to function as controllers of discourse framework shifts.

As described in Chapter One, the main question of the current thesis is stated as follows:

> What are the differences and similarities between British-British conversation and British-Japanese conversation in English in the context of academic supervision sessions?

The answer to the question is summarised as follows:

1. Similarities in the use of framework shifts and multi-functional nature of response tokens were recognised between the British-British conversations and the British-Japanese conversation.
2. L1 (first language) transfer in listenership behaviour was observed in the Japanese students' use of response tokens, such as the monotonous use of head nods and more negotiations for speaker selection.

Some research areas were highlighted from the current research. In terms of developing a methodology for language research for example, the alignments of the data can be one of the issues to be improved in further research. In this study, participants' utterances and body movements are aligned on a spreadsheet vertically with the timeline as a primary key. This can be applicable for analysis of a dyad conversation, although it might not be practical to apply to conversation analysis with

more than two participants, since it could be too complex for researchers to analyse the multiple participants' verbal and visual transcripts aligned in several columns horizontally. In addition, because of the alignment, the sequence of the utterance and body movements within one second has vanished.

The other issue to be pointed out is the accuracy in combining verbal data with visual data. Since the timestamps for verbal and visual data were added manually in the current study, there might be time delays between the point where an utterances or a body movement occurred and the point where a timestamp were added. With more advanced technology to capture utterances and body movements automatically, a time-related corpus can be of use for language research with more accuracy in timestamps.

In terms of functions and forms of response tokens, collocations between verbal and visual response tokens were not fully explored in the current study. This can be one of the areas to be highlighted for future research. Analysing the relation between the use of particular response tokens and turn-structural episodes from the perspectives of intercultural communication is another issue to be highlighted for further research. In addition, occurrences of L1 transfer and multiple identities in interlanguage need to be further explored in order to indicate elements to become a successful intercultural communicator.

7.1 Limitations of the research

Although an attempt was made to establish a methodology for conversation analysis with multi-modal spoken language corpus by integrating a concept of leadtime, the current study can be seen as a preliminary study of a time-related corpus approach.

There are several limitations in the current study, such as the amount of the data.

Two British-British conversations and two British-Japanese conversations of 39-minute length were examined in this investigation, and the amount of the research data is adequate to conduct a preliminary study with a new method for conversation analysis but not enough to establish reliability of the findings.

Secondly, there is no Japanese-Japanese conversation data collected for comparison in the current study, although the existing research on Japanese conversations was reviewed in Chapter Two. It would be ideal if the data of three conversation types, namely British-British conversation, British-Japanese conversation and Japanese-Japanese conversation, in the same context could be recorded and analysed. To narrow the focus of the research, however, the two types of conversations in English were used in the current research.

Thirdly, retrospective views from the participants are also missing in the current study. Since the main focus is placed on comparing the use of response tokens between British-British conversation and British-Japanese conversation with a time-related spoken corpus, the participants' retrospective views through oral interviews or written journals were excluded from the scope of this investigation. However, these issues can be improved in a future replication of this study.

7.2 Potentials and drawbacks of a time-related corpus-based approach

A new method for conversation analysis established through the current study will also be reviewed. Potentials and drawbacks of the new research method with a time-related corpus analysis will be described.

The synthesis of time and visual data with verbal data can be seen as a unique asset of the current study. Although it is a preliminary study with multimodal and

time-related corpus approach, some contributions were made for developing a methodology in conversation analysis. In a traditional transcript in conversation analysis, utterances of participants are described in line with many annotations, such as intonations, prolongations, and overlaps as shown in Transcript 7.2-1 below, which is well developed to analyse sequences and adjacency pairs with a limited data.

Transcript 7.2-1 Sample transcript (1)

Sample transcription

```
1       Maude:  I says well it's funny: Mizssi:z uh:↑Schmidt ih you'd
2               think she'd help<.hhh Well (.) Missiz Schmidt was the
3               one she: (0.2) assumed respo:nsibility for the three
4               specials.
5               (o.6)
6       Bea:    Oh↓*::,°°M-hm,°°=
7       Maude:  =Maybe: ↑told me this.
```

(Have 2001: 90)

In the time-related corpus-based approach introduced in the current study, however, utterances and movements of two participants can be described with the timeline and aligned in separate columns horizontally as shown in Transcript 7.2-2 below:

Transcript 7.2-2 Sample Transcript (2)

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 01: 36	FBS_F	-13	0				No they it's these interviews are on the website as a kind of erm I guess like a self-help forum for people to go and other people to go and visit+
00: 01: 37		-12	1				
00: 01: 38		-11	2			HG	
00: 01: 39		-10	3				
00: 01: 40		-9	4				
00: 01: 41		-8	5			HG	
00: 01: 42		-7	6	HN		HG	
00: 01: 43		-6	7				
00: 01: 44		-5	8		Oh I see right.	HG	+and they're separated out by the particular conditions.
00: 01: 45		-4	9				
00: 01: 46		-3	10	HN			
00: 01: 47		-2	11		Right.	HG	
00: 01: 48		-1	12				Erm.
00: 01: 49	MBT_F	0	-8	HG	So th= the interviews are there because of the condition not because of the any+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

The time-related transcripts enable researchers to analyse the length of speaker/listener status, and the instances of response tokens in reference to TTP in seconds as described in the previous chapter. In addition, by extracting participant's listener status as illustrated above, patterns in the use of verbal and visual response tokens in listener status can be synthesised and analysed with a timeline. These are issues which many researchers are aware of but which are difficult to analyse with traditional transcripts.

Some drawbacks of the time-related corpus approach need to be reported. Since the timeline in the time-related transcripts was rounded to the nearest second, utterances of two participants in the same seconds will be shown in the same time line. At 00:01:44 in the timeline in the sample transcript (2) above, for example, BBC1_MBT uttered 'Oh I see right' before the BBC1_FBS's utterance in the same second. However, this sequence vanished because of the alignment of the time-related transcripts. A researcher needs to listen to the audio data to verify the order of the utterances in the same second. The other drawback of the time-related transcripts is the reduction in the numbers of HNs and HGs. When several HNs or HNs occur within one second, these HNs and HGs were counted as one HN or one HG since the timeline in the time-related transcripts is rounded up to the nearest second. If a researcher would like to count a precise number of HNs and HGs in conversation with time-related transcripts, they would be required to develop tenth of a seconds or milliseconds level time-related transcripts. That is excluded from the scope of the current research.

7.3 Future research

A brief discussion on the findings of the current study from an ecolinguistic perspective will be described in this section for future research. As reviewed in Section 2.4, the idea of *frames* (Ager 1994a, 1994b, Brown & Yule 1983), *a structure of expectation* (Kramsch 1998) or *Schemata* (Cook 1994) has been introduced to describe people's assumptions of sociocultural routines and conversation styles in particular situations. The processes in intercultural communication were illustrated as a coding framework where a speaker and a listener encode and decode a range of components in a language system in a particular culture in order to get their intentions across. It is assumed that failures in intercultural communication can be caused by a speaker's mis-encoding or a listener's mis-decoding of the language components used as a medium (Jandt 1995, McKey et al. 1995, Singer 1998).

These interpretations of language systems developed through the early twentieth century are intervened by Structuralism, which states that people in a particular culture are assumed to share the components of the language systems and their activations are expected to be stable and consistent. From an ecological perspective, however, Haugen (1972) defines the late nineteenth century as that in which 'interactions between any given language and its environment' (ibid: 325) were studied and the 'vulnerable' nature of language is recognised. Kramsch (2009: 2) also expressed this vulnerable nature by contrasting a modernist view, where people exchange information of a stable world of objective truth through rational argument, with a post-modernist view, where the words people speak might mean different things and 'the memories these words evoke are not necessarily shared'.

In addition, in the ecolinguistic view, not the text but the *context* where a

language is used is centralised. Context is treated as 'the heart of the matter' (van Lier 2004: 5):

> EL [ecological linguistics] regards context as not just something that surrounds language, but that in fact defines language, while at the same time being defined 'by' it.
> (ibid: 5)

Ecological linguistics enhances the role of context in interaction. Influenced by Diane Larsen-Freeman's (1997) 'complex systems theory', which is inspired by uncertainty in physics and chaos in astronomy, some language educators and linguists in intercultural communication (Blommaert 2005, Kramsch 2008, Kramsch & Whiteside 2008, van Lier 2004) have taken the idea of ecological linguistics as a new approach to language analysis, referring to the terms, 'relativity of self and other', 'multiple timescales', 'adaptations to emergents', 'open-endedness and unfinalizability' and 'fractals' of activities and events (Kramsch & Whiteside 2008: 659).

In reference to his study on discourse analysis in a post-Apartheid context in Africa, Blommaert (2005) raised awareness of the importance of timescales in discourse and introduces the term *layered simultaneity*.

> […] we have to conceive of discourse as subject to 'layered simultaneity': It occurs in a real-time, synchronic event, but it is simultaneously encapsulated in several layers of historicity, some of which are within the grasp of the participants while others remain

invisible but are nevertheless present.

(Blommaert 2005: 130)

Although this aspect cannot be investigated in the current study, the notion of relativity of self and other, layered simultaneity and fractal emergence of parts of the identities in a particular context can be concerned in further research in listenership behaviour in interpragmatic settings. Ecological linguistics in particular can be one of the theoretical frameworks which can provide more relevant and convincing interpretations on language learners' presentation of their multiple identities.

7.4 Summary

The current study investigated forms and placements of verbal and visual response tokens in relation to turn-taking structure by introducing the time-related multi-modal corpus approach. A comparison was made between British tutor-British student conversations and British tutor-Japanese student conversations in face-to-face dyad academic tutorials. From this preliminary research, the multi-functional nature of response tokens was recognised. Response tokens were used as controllers of turn-taking and discourse framework shifts, and simultaneously were uttered as expressions of participants' thoughts.

In interlanguage settings, L1 transfer was observed in the Japanese students' use of response tokens. However, the Japanese students also used the same strategies in the use of response tokens as the British students. As reviewed in this chapter, one of the explanations for multiple identities of interlanguage users can be made by theories in ecolinguistics. Areas of linguistic research, such as turn-taking structure, interlanguage pragmatics and multiple identities in interlanguage users, were

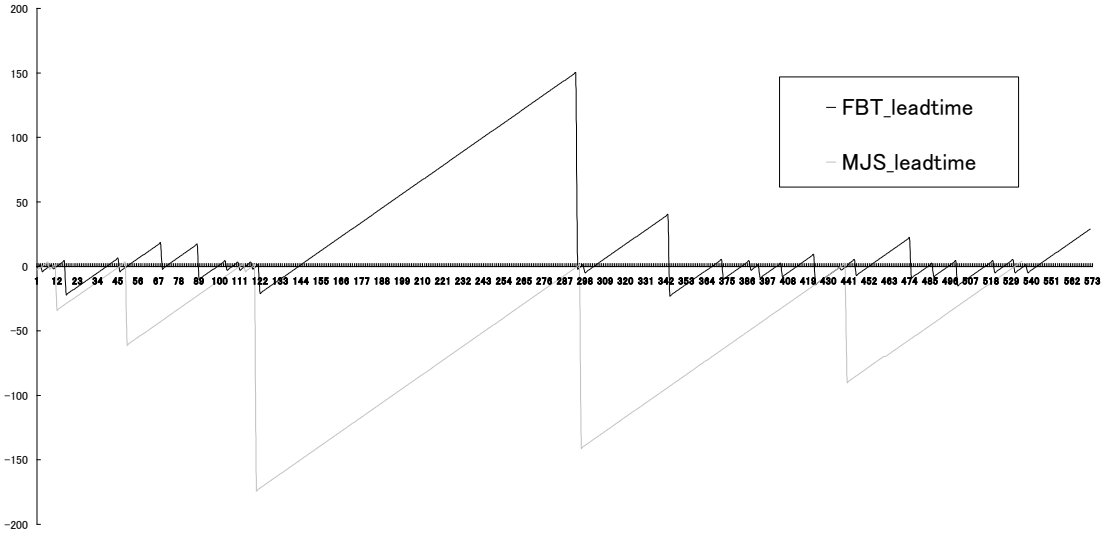
highlighted. An interpretation of multiple identities of language learners from the perspective of ecolinguistics has been suggested in this chapter. These areas can be investigated in future research with this new research method comprising multimodal data and leadtime.

Appendices

Appendix A Tables and figures from the pilot study

A.1 Tables and figures from the global pattern analysis

A.1.1 Equality and inequality in turn-taking



Keys: C2= conversation 2 (British-Japanese conversation), y axis = leadtime (seconds), x axis = timeline (seconds), FBT_leadtime= female British tutor's leadtime, MJS_leadtime= male Japanese student's leadtime,

Figure A.1-1 Numbers of taking the floor and turn in C2

A.1.2 Findings from hand gestures and head nods

Table A.1.2-1 Number of HGs and HNs in C1

leadtime	C1_MBT_HG	C1_FBS_HG	C1_MBT_HN	C1_FBS_HN
less than -50	0	0	0	1
less than -45	0	0	0	1
less than -40	0	0	0	3
less than -35	0	0	1	3
less than -30	0	0	2	3
less than -25	0	0	3	4
less than -20	0	0	3	6
less than -15	0	0	7	6
less than -10	0	2	5	5
less than -5	1	3	12	8
less than 0	2	6	16	8
more than 0	28	22	2	4
more than 5	14	21	0	0
more than 10	11	10	0	0
more than 15	7	9	0	0
more than 20	7	8	0	0
more than 25	4	5	0	0
more than 30	7	1	0	0
more than 35	5	2	0	0
more than 40	4	3	0	0
more than 45	4	1	0	0
more than 50	1	0	0	0
Sum	95	93	51	52

Keys: HG= hand gestures, HN= head nods, C1= conversation 1 (British-British Conversation),
C1_MBT_HG = C1 male British tutor's hand gestures, C1_FBS_HG = C1 female British student's hand gestures,
C1_MBT_HN = C1 male British tutor's head nods, C1_FBS_HN = C1 female British student's head nods

Table A.1.2-2 Number of HGs and HNs in C2

leadtime	C2_FBT_HG	C2_MJS_HG	C2_FBT_HN	C2_MJS_HN
less than -100	0	0	0	10
less than -95	0	0	0	3
less than -90	0	0	0	2
less than -85	0	0	0	1
less than -80	0	0	0	1
less than -75	0	0	0	2
less than -70	0	0	0	3
less than -65	0	0	0	3
less than -60	0	0	0	2
less than -55	0	0	0	3
less than -50	0	0	0	2
less than -45	0	0	0	0
less than -40	0	0	0	3
less than -35	0	0	0	2
less than -30	0	0	0	1
less than -25	0	0	0	3
less than -20	0	0	0	1
less than -15	0	0	0	4
less than -10	0	0	0	3
less than -5	0	0	0	0
less than 0	0	0	0	0
more than 0	1	1	0	1
more than 5	2	0	0	0
more than 10	0	0	0	0
more than 15	0	0	0	0
more than 20	1	0	0	0
more than 25	1	0	0	0
more than 30	0	0	0	0
more than 35	1	0	0	0
more than 40	1	0	0	0
more than 45	2	0	0	0
more than 50	2	0	0	0
more than 55	1	0	0	0
more than 60	3	0	0	0
more than 65	1	0	0	0
more than 70	1	0	0	0
more than 80	1	0	0	0
more than 85	2	0	0	0
more than 90	3	0	0	0
Sum	23	1	0	50

Keys: HG= hand gestures, HN= head nods, C2= conversation 2 (British-Japanese conversation),
C2_FBT_HG = C2 female British tutor's hand gestures, C2_MJS_HG = C2 male Japanese student's hand gestures,
C2_FBT_HN = C2 female British t tutor's head nods, C2_MJS_HN = C2 male Japanese student's head nods

Table A.1.2-3 Means, variances, SD of HGs and HNs in C1

	C1_MBT_HG (n=94)	C1_FBS_HG (n=93)	C1_MBT_HN (n=51)	C1_FBS_HN (n=52)
Mean	15.23	10.46	-11.78	-18.10
Variance	226.80	153.86	96.25	217.50
SD	14.98	12.34	9.71	14.61

Keys: SD= standard deviation, HG= hand gestures, HN= head nods, C1= conversation 1 (British-British Conversation), C1_MBT_HG = C1 male British tutor's hand gestures, C1_FBS_HG = C1 female British student's hand gestures, C1_MBT_HN = C1 male British tutor's head nods, C1_FBS_HN = C1 female British student's head nods

Table A.1.2-4 Means, variances, SD of HGs and HNs in C2

	C2_FBT_HG (n=23)	C2_MJS_HG (n=1)	C2_FBT_HN (n=0)	C2_MJS_HN (n=50)
Mean	58.65	2.00	—	-65.54
Variance	1247.33	—	—	1462.87
Standard deviation	34.54	—	—	37.86

Keys: SD= standard deviation, HG= hand gestures, HN= head nods, C2= conversation 2 (British-Japanese conversation), C2_FBT_HG = C2 female British tutor's hand gestures, C2_MJS_HG = C2 male Japanese student's hand gestures, C2_FBT_HN = C2 female British tutor's head nods, C2_MJS_HN = C2 male Japanese student's head nods,

A.1.3 Findings from *erm*, *yeah* and *mm*

Table A.1.3-1 Numbers of *erm*, *yeah* and *mm* in C1

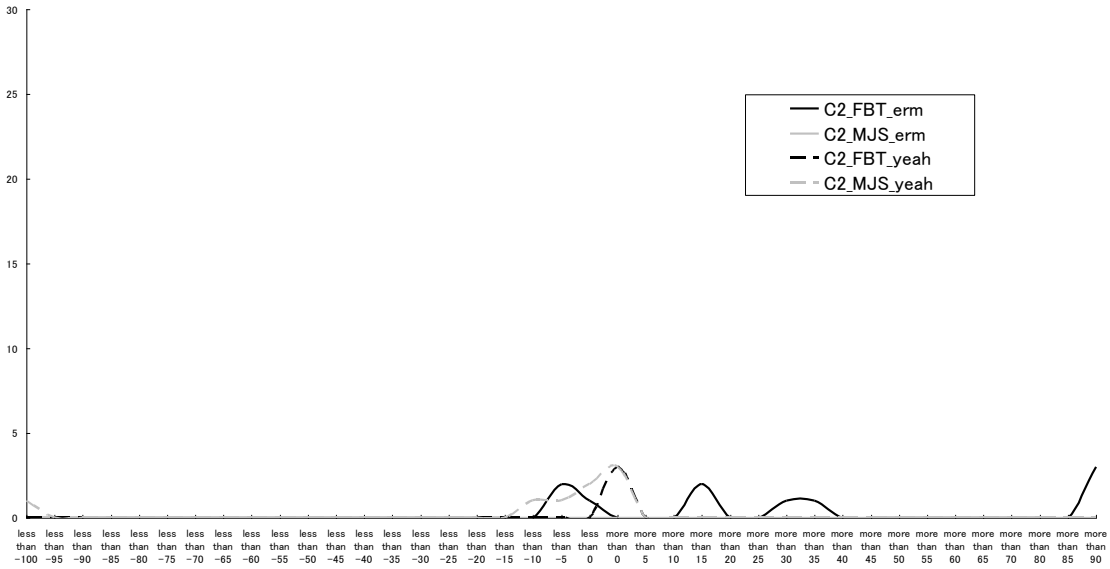
leadtime	C1_MBT_erm	C1_FBS_erm	C1_MBT_yeah	C1_FBS_yeah	C1_MBT_mm	C1_FBS_mm
less than -50	0	0	0	1	0	0
less than -45	0	0	0	1	0	0
less than -40	0	0	0	2	0	0
less than -35	0	0	1	1	0	0
less than -30	0	0	0	0	0	0
less than -25	0	0	1	0	0	0
less than -20	0	0	2	3	0	0
less than -15	0	0	4	2	0	0
less than -10	0	0	1	3	0	0
less than -5	0	0	5	6	1	0
less than 0	0	4	21	8	0	0
more than 0	3	12	5	7	0	0
more than 5	1	1	1	0	0	1
more than 10	2	6	2	0	0	0
more than 15	2	1	0	0	0	0
more than 20	0	0	0	0	0	0
more than 25	2	2	0	0	0	1
more than 30	1	1	0	0	0	0
more than 35	1	1	0	0	0	0
more than 40	1	0	0	0	0	0
more than 45	1	0	0	0	0	0
more than 50	1	0	0	0	0	0
	15	28	43	34	1	2

Keys: C1= conversation 1 (British-British Conversation),
C1_MBT_erm = C1 male British tutor's erm, C1_FBS_erm = C1 female British student's erm,
C1_MBT_yeah = C1 male British tutor's yeah, C1_FBS_yeah = C1 female British student's yeah,
C1_MBT_mm = C1 male British tutor's mm, C1_FBS_mm = C1 female British student's mm,

Table A.1.3-2 Numbers of *erm*, *yeah* and *mm* in C2

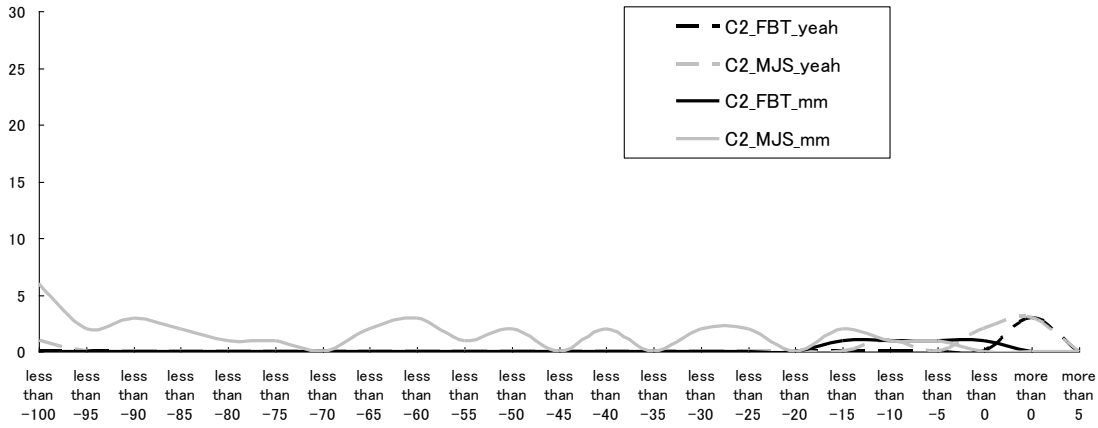
leadtime	C2_FBT_erm	C2_MJS_erm	C2_FBT_yeah	C2_MJS_yeah	C2_FBT_mm	C2_MJS_mm
less than -100	0	0	0	1	0	6
less than -95	0	0	0	0	0	2
less than -90	0	0	0	0	0	3
less than -85	0	0	0	0	0	2
less than -80	0	0	0	0	0	1
less than -75	0	0	0	0	0	1
less than -70	0	0	0	0	0	0
less than -65	0	0	0	0	0	2
less than -60	0	0	0	0	0	3
less than -55	0	0	0	0	0	1
less than -50	0	0	0	0	0	2
less than -45	0	0	0	0	0	0
less than -40	0	0	0	0	0	2
less than -35	0	0	0	0	0	0
less than -30	0	0	0	0	0	2
less than -25	0	0	0	0	0	2
less than -20	0	0	0	0	0	0
less than -15	0	0	0	0	1	2
less than -10	0	0	0	1	1	1
less than -5	2	0	0	0	1	1
less than 0	1	0	0	2	1	0
more than 0	0	0	3	3	0	0
more than 5	0	0	0	0	0	0
more than 10	0	0	0	0	0	0
more than 15	2	0	0	0	0	0
more than 20	0	0	0	0	0	0
more than 25	0	0	0	0	0	0
more than 30	1	0	0	0	0	0
more than 35	1	0	0	0	0	0
more than 40	0	0	0	0	0	0
more than 45	0	0	0	0	0	0
more than 50	0	0	0	0	0	0
more than 55	0	0	0	0	0	0
more than 60	0	0	0	0	0	0
more than 65	0	0	0	0	0	0
more than 70	0	0	0	0	0	0
more than 80	0	0	0	0	0	0
more than 85	0	0	0	0	0	0
more than 90	3	0	0	0	0	0
Sum	10	0	3	6	4	33

Keys: C2= conversation 2 (British-Japanese conversation),
C2_FBT_erm = C2 female British tutor's erm, C2_MJS_erm = C2 male Japanese student's erm,
C2_FBT_yeah = C2 female British tutor's yeah, C2_MJS_yeah = C2 male Japanese student's yeah,
C2_FBT_mm = C2 female British tutor's mm, C2_MJS_mm = C2 male Japanese student's mm



Keys: C2= conversation 2 (British-Japanese conversation),
y axis = the numbers of erm and yeah, x axis = time intervals of leadtime (seconds),
C2_FBT_erm = C2 female British tutor's erm, C2_MJS_erm = C2 male Japanese student's erm,
C2_FBT_yeah = C2 female British tutor's yeah, C2_MJS_yeah = C2 male Japanese student's yeah,

Figure A.1-2 Numbers of *erm* and *yeah* in C2



Keys: C2= conversation 2 (British-Japanese conversation),
y axis = the numbers of yeah and mm, x axis = time intervals of leadtime (seconds),
C2_FBT_yeah = C2 female British tutor's yeah, C2_MJS_yeah = C2 male Japanese student's yeah,
C2_FBT_mm = C2 female British tutor's mm, C2_MJS_mm = C2 male Japanese student's mm

Figure A.1-3 Numbers of *yeah* and *mm* in C2

Table A.1.3-3 Mean, variance, SD of vocal response tokens in C1

	C1_MBT_erm (n=15)	C1_FBS_erm (n=28)	C1_MBT_yeah (n=43)	C1_FBS_yeah (n=34)	C1_MBT_mm (n=1)	C1_FBS_mm (n=2)
Mean	22.27	7.43	-5.35	-12.37	-7.00	17.00
Variance	287.35	123.37	146.42	253.59	–	128.00
SD	16.38	10.91	11.96	15.70	–	8.00

Keys: SD= standard deviation, C1= conversation 1 (British-British Conversation),
C1_MBT_erm = C1 male British tutor's erm, C1_FBS_erm = C1 female British student's erm,
C1_MBT_yeah = C1 male British tutor's yeah, C1_FBS_yeah = C1 female British student's yeah,
C1_MBT_mm = C1 male British tutor's mm, C1_FBS_mm = C1 female British student's mm,

Table A.1.3-4 Mean, variance, SD of vocal response tokens in C2

	C2_FBT_erm (n=10)	C2_MJS_erm (n=0)	C2_FBT_yeah (n=3)	C2_MJS_yeah (n=7)	C2_FBT_mm (n=4)	C2_MJS_mm (n=33)
Mean	41.80	–	2.00	-16.57	-8.50	-67.21
Variance	2689.29	–	4.00	1371.95	87.00	1167.67
Standard deviation	49.20	–	1.63	34.29	8.08	33.65

Keys: SD= standard deviation, C2= conversation 2 (British-Japanese conversation),
C2_FBT_erm = C2 female British tutor's erm, C2_MJS_erm = C2 male Japanese student's erm,
C2_FBT_HG = C2 female British tutor's hand gestures, C2_MJS_HG = C2 male Japanese student's hand gestures,
C2_FBT_HN = C2 female British tutor's head nods, C2_MJS_HN = C2 male Japanese student's head nods,

A.2 Tables and figures from the turn-structural analysis

A.2.1 Collocations of verbal response tokens with visual response tokens

Transcript A.2-1 C1_MBT Pattern A: Response tokens in floor seeker

Timeline	Floor	Pattern	MBT_leadtime	FBS_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
32		A	-3	24	FS	IR		Yeah.
34		A	-1	26	FS	IR		Yeah+
62		A	-4	26	FS	IR		Right. Ah right.
78		A	-3	10	FS	IR		Right.
104		A	-5	8	FS	IR		Oh I see right.
107		A	-2	11	FS	IR		Right.
119		A+	-2	2	FS	IR	HN	Right. Okay.
161		A+	-2	4	FS	IR		Ah right. Okay.
183		A	-3	12	FS	IR		That's right.
184		A	-2	13	FS	IR		Yeah yeah.
306		A	-5	5	FS	IR	SC/neck	Yeah yeah.
308		A	-3	7	FS	IR		Yeah.
350		A	-3	6	FS	iR		Right.
351		A	-2	7	FS	IR		Yeah but I+
468		A	-1	16	FS	IR		Right.
478		A	-3	2	FS	IR		Yeah yeah.
479		A	-2	3	FS	IR	HN	Right yeah yeah yeah.
509		A	-4	8	FS	IR		Yeah yeah.
509		A	-4	8	FS	IR		Yeah.
509		A	-4	8	FS	IR		Yeah yeah.
509		A	-4	8	FS	IR		Yeah.
510		A	-3	9	FS	IR		Yeah.
512		A	-1	11	FS	IR		Yeah.
587		A+	-3	1	FS	IR		Oh right oh okay.

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
MBT_leadtime=male British tutor leadtime, FBS_leadtime= female British student leadtime,
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
MBT_F=MBT floor-taking, =pattern A, A+= a variant of pattern A,
FS= floor seeker, IR= information receipt tokens, HN= head nods, SC/neck= self comfort with neck

Transcript A.2-2 C1_MBT Pattern A: Response tokens in floor-taking

Timeline	Floor	Pattern	MBT_leadtime	FBS_leadtime	C1_MBT_CF	C1_MBT_DF	C1_MBT_gesture	C1_MBT_Transcript
35	MBT_F	A	0	-1	FT		HG	+but you were suggesting the other that you didn't want to do that <$G?>.
66	MBT_F	A	0	-2	FT		HN	Yeah.'Oh right so they're separately interviewed?
81	MBT_F	A	0	-15	FT			well it depen= yeah that's not necessarily a problem erm.
109	MBT_F	A	0	-8	FT		HG	So th= the interviews are there because of the condition not because of the any+
121	MBT_F	A+	0	-36	FT			Ah you might be alright then you see the only worry is if if you're dealing with data that's been set up in an experimental situation by a linguist+
163	MBT_F	A+	0	-8	FT			Right. So it= so it's not going to be a sort of discoursey thing then?
186	MBT_F	A	0	-9	FT		HG	Well it's shared but+
311	MBT_F	A	0	-3	FT			So have you looked at this stuff? I mean is there+
353	MBT_F	A	0	-7	FT			+suppose the interesting thing is what what source domains they're using.
469	MBT_F	A	0	-7	FT			So this is fighting your illness as an+
481	MBT_F	A	0	-20	FT		HG/HS	It's a funny one that isn't it cos this Susan Sontag talks about this that if you if you see your illness as an enemy and and you die of cancer it's your fault+
513	MBT_F	A	0	-6	FT		HG	So are you interested in the conceptual content or the sort of stylistic realisation?
590	MBT_F	A+	0		FT			Yeah. So the the but you're you're prodding the the audience person to to think of it to be persuaded or not.

Keys: C1_MBT=Conversation 1 (NS-NS) male British tutor,
MBT_leadtime=male British tutor leadtime, FBS_leadtime= female British student leadtime,
C1_MBT_CF=C1_MBT conversation function, C1_MBT_DF=C1_MBT discourse function,
MBT_F=MBT floor-taking, =pattern A, A+= a variant of pattern A, FT= floor-taking,
HN= head nods, HG= hand gestures, HG/HS= hand gestures and head shakes

Table A.2.1-1 Pattern B: collocation of verbal and visual response tokens

				C1_MBT With gestture/ All (details)	C1_FBS With gestture/ All (details)	C2_FBT With gestture/ All (details)	C2_MJS With gestture/ All (details)
Pattern B (SS)	FS	HN or HS	Engaged okens	1/1 (HN1)	1/5 (HS1)	--	--
	FT	HG	--	1/1 (HG1)	1/3 (HG1)	--	--

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_FBS=Conversation 1 (British-British conversation) female British student
C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
with gestures/ all = the number of verbal response tokens with gestures/ the total number of verbal response tokens,
details= the details of visual response tokens, SS= Self-selection, LS= listenership, FS= floor seeker, FT= floor-taking,
HN= head nods, HS= hand shakes, HG= hand gestures,

Table A.2.1-2 Pattern C: collocation of verbal and visual response tokens

				C1_MBT With gestture/ All (details)	C1_FBS With gestture/ All (details)	C2_FBT With gestture/ All (details)	C2_MJS With gestture/ All (details)
Pattern C (SS)	LS	HN	Continuers	1/15 (HN1)	11/14 (HN10, HN+HG1)	--	1/2 (HN1)
	FS	HN	Information receipt tokens / engaged tokens	1/5 (HN1)	4/4 (HN2, HN+HG2)	--	0/1
	FT	HG	--	3/3 (HG3)	2/3 (HN1, HG1)	--	1/1 (HG1)

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_FBS=Conversation 1 (British-British conversation) female British student
C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
with gestures/ all = the number of verbal response tokens with gestures/ the total number of verbal response tokens,
details= the details of visual response tokens, SS= Self-selection, LS= listenership, FS= floor seeker, FT= floor-taking,
HN= head nods, HG= hand gestures, HN+HG= head nods and hand gestures

Table A.2.1-3 Pattern D: collocation of verbal and visual response tokens

				C1_MBT With gestture/ All (details)	C1_FBS With gestture/ All (details)	C2_FBT With gestture/ All (details)	C2_MJS With gestture/ All (details)
Pattern D (OS)	LS	HN	Continuers	--	3/5 (HN3)	--	12/21 (12HN)
	FG	--	--	--	--	--	--
	FS	HN or HS	Convergence tokens	--	4/9 (HN2, HN+HG2)	0/0	0/3
	FT	HN or HS with HG	--	--	6/13 HG3, HN1, HN+HG1, HS+HG1)	0/1	0/7

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_FBS=Conversation 1 (British-British conversation) female British student
C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
with gestures/ all = the number of verbal response tokens with gestures/ the total number of verbal response tokens,
details= the details of visual response tokens, SS= Self-selection,
LS= listenership, FG= floor given, FS= floor seeker, FT= floor-taking, HN= head nods, HS= head shakes,
HG= head gestures, HN+HG= head nods and hand gestures, HS+HG= head shakes and hand gestures

Table A.2.1-4 Pattern E: collocation of verbal and visual response tokens

				C1_MBT With gestture/ All (details)	C1_FBS With gestture/ All (details)	C2_FBT With gestture/ All (details)	C2_MJS With gestture/ All (details)
Pattern E (SS)	FS	--	Pause/ discourse markers	--	--	0/7	--
	FT	HT	--	--	--	2/18 (HT2)	--

Keys: C1_MBT=Conversation 1 (British-British conversation) male British tutor,
C1_FBS=Conversation 1 (British-British conversation) female British student
C2_FBT=Conversation 2 (British-Japanese conversation) female British tutor,
C2_MJS=Conversation 2 (British-Japanese conversation) male Japanese student
with gestures/ all = the number of verbal response tokens with gestures/ the total number of verbal response tokens,
details= the details of visual response tokens, SS= Self-selection,
FS= floor seeker, FT= floor-taking, HT= head turns,

Appendix B Tables and figures from the main study

B.1 Tables and figures from the global pattern analysis

B.1.1 Number and length of speaker turns

Table B.1.1-1 BBC2 Summary: No. and length of speaker turns

	Speaker status		Speaker turns	
	Length (sec)	Length (HH:MM:SS)	No. turns	Length/turn (sec)
BBC2_MBT	1553	00: 25: 53	36	43.14
BBC2_MBS	776	00: 12: 56	34	22.82
PAUSE	11	00: 00: 11	–	–
Unclassified	0	00: 00: 00	–	–
TOTAL TIME	2340	00: 39: 00	^	^

Keys: BBC2_MBT=British-British Conversation 2 male British tutor
BBC2_MBS=British-British Conversation 2 male British student

Table B.1.1-2 BJC1 Summary: No. and length of speaker turns

	Speaker status		Speaker turns	
	Length (sec)	Length (HH:MM:SS)	No. turns	Length/turn (sec)
BJC1_FBT	1318	00: 21: 58	119	11.08
BJC1_MJS	189	00: 03: 09	24	7.88
PAUSE	692	00: 11: 32	–	–
Unclassified	141	00: 02: 21	–	–
TOTAL TIME	2340	00: 39: 00	^	^

Keys: BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC1_MJS=British-Japanese Conversation 1 male Japanese student

Table B.1.1-3 BJC2 Summary: No. and length of speaker turns

	Speaker status		Speaker turns	
	Length (sec)	Length (HH:MM:SS)	No. turns	Length/turn (sec)
BJC2_MBT	1772	00: 29: 32	33	53.70
BJC2_MJS	450	00: 07: 30	24	18.75
PAUSE	79	00: 01: 19	–	–
Unclassified	39	00: 00: 39	–	–
TOTAL TIME	2340	00: 39: 00	^	^

Keys: BJC2_MBT=British-Japanese Conversation 2 male British tutor
BJC2_MJS=British-Japanese Conversation 1 male Japanese student

B.1.2 Placement of hand gestures

Table B.1.2-1 Tutors' use of hand gestures

leadtime	BBC1_MBT_HG	BBC2_MBT_HG	BJC1_FBT_HG	BJC2_MBT_HG
less than -50	0	0	0	0
less than -45	0	0	0	0
less than -40	0	0	0	0
less than -35	0	0	0	0
less than -30	0	0	0	0
less than -25	0	0	0	0
less than -20	0	0	1	1
less than -15	0	0	0	0
less than -10	1	1	0	0
less than -5	1	1	0	2
less than 0	11	4	3	2
more than 0	138	30	40	32
more than 5	89	26	47	31
more than 10	76	31	32	33
more than 15	51	22	18	27
more than 20	32	16	15	20
more than 25	26	15	10	27
more than 30	30	18	8	27
more than 35	16	10	9	17
more than 40	10	10	3	19
more than 45	9	20	3	18
more than 50	33	115	2	230
	523	319	191	486

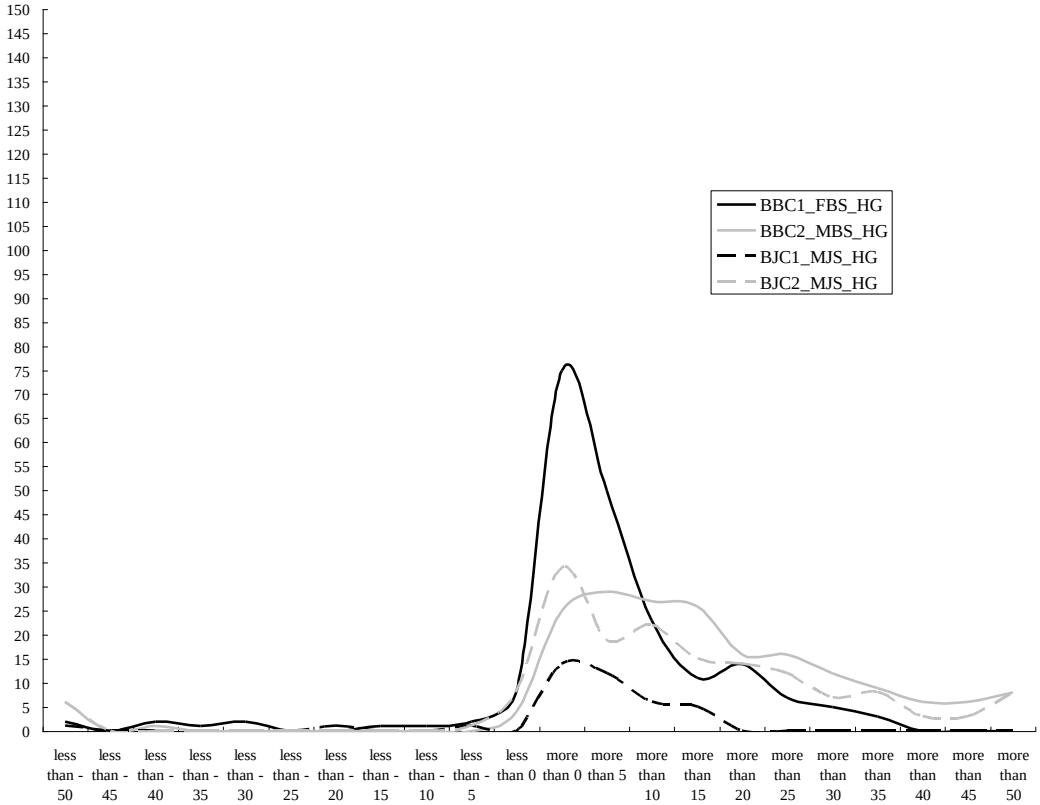
Keys: BBC1_MBT_HG = British-British Conversation 1 male British tutor's hand gestures,
BBC2_MBT_HG = British-British Conversation 2 male British tutor's hand gestures
BJC1_FBT_HG = British-Japanese Conversation 1 female British tutor's hand gestures,
BJC2_MBT_HG = British-Japanese Conversation 2 male British tutors hand gestures

Table B.1.2-2 Students' use of hand gestures

leadtime	BBC1_FBS_HG	BBC2_MBS_HG	BJC1_MJS_HG	BJC2_MJS_HG
less than -50	2	6	1	6
less than -45	0	0	0	0
less than -40	2	1	0	0
less than -35	1	0	0	0
less than -30	2	0	0	0
less than -25	0	0	0	0
less than -20	0	0	1	0
less than -15	1	0	0	0
less than -10	1	0	0	0
less than -5	2	0	1	1
less than 0	8	4	0	8
more than 0	75	25	14	34
more than 5	50	29	12	19
more than 10	23	27	6	22
more than 15	11	26	5	15
more than 20	14	16	0	14
more than 25	7	16	0	12
more than 30	5	12	0	7
more than 35	3	9	0	8
more than 40	0	6	0	3
more than 45	0	6	0	3
more than 50	0	8	0	8
	207	191	40	160

Keys: BBC1_FBS_HG = British-British Conversation 1 female British student's hand gestures,
BBC2_MBS_HG = British-British Conversation 2 male British student's hand gestures
BJC1_MJS_HG = British-Japanese Conversation 1 male Japanese student's hand gestures,
BJC2_MJS_HG = British-Japanese Conversation 2 male Japanese student's hand gestures

Figure B.1-1　Students' use of hand gestures



Keys: BBC1_FBS_HG = British-British Conversation 1 female British student's hand gestures,
BBC2_MBS_HG = British-British Conversation 2 male British student's hand gestures
BJC1_MJS_HG = British-Japanese Conversation 1 male Japanese student's hand gestures,
BJC2_MJS_HG = British-Japanese Conversation 2 male Japanese student's hand gestures

B.1.3 Placement of head nods

Table B.1.3-1 Tutors' use of head nods

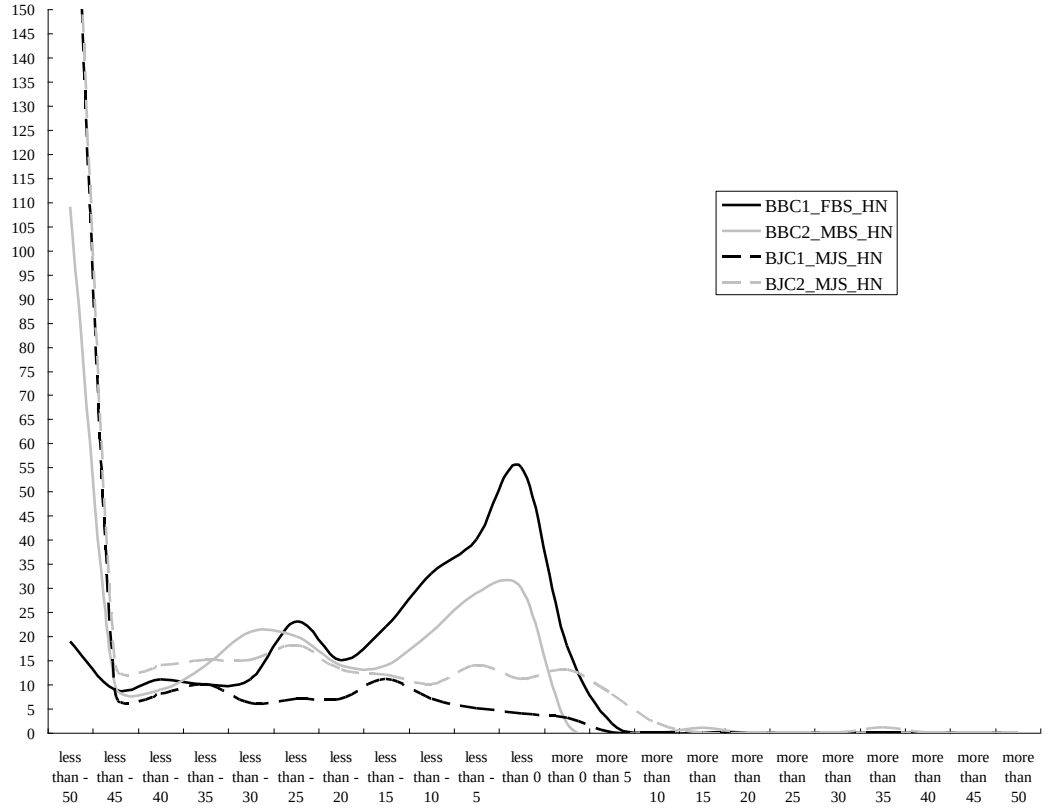
leadtime	BBC1_MBT_HN	BBC2_MBT_HN	BJC1_FBT_HN	BJC2_MBT_HN
less than -50	0	12	0	7
less than -45	0	2	0	2
less than -40	0	6	0	6
less than -35	3	6	1	2
less than -30	3	6	0	5
less than -25	5	9	0	4
less than -20	5	11	0	6
less than -15	10	7	3	10
less than -10	10	20	1	8
less than -5	30	18	2	14
less than 0	73	22	4	22
more than 0	14	2	4	9
more than 5	1	1	0	1
more than 10	0	2	0	1
more than 15	0	2	0	0
more than 20	0	0	0	0
more than 25	1	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	0
more than 40	0	0	0	0
more than 45	0	0	0	1
more than 50	0	2	0	4
	155	128	15	102

Keys: BBC1_MBT_HN = British-British Conversation 1 male British tutor's head nods,
BBC2_MBT_HN = British-British Conversation 2 male British tutor's head nods
BJC1_FBT_HN = British-Japanese Conversation 1 female British tutor's head nods,
BJC2_MBT_HN = British-Japanese Conversation 2 male British tutor's head nods

Table B.1.3-2 Students' use of head nods

leadtime	BBC1_FBS_HN	BBC2_MBS_HN	BJC1_MJS_HN	BJC2_MJS_HN
less than -50	19	109	204	208
less than -45	9	10	8	14
less than -40	11	9	8	14
less than -35	10	14	10	15
less than -30	11	21	6	15
less than -25	23	20	7	18
less than -20	15	14	7	13
less than -15	22	14	11	12
less than -10	33	21	7	10
less than -5	40	29	5	14
less than 0	55	30	4	11
more than 0	18	2	3	13
more than 5	2	0	0	8
more than 10	0	0	0	2
more than 15	0	1	0	0
more than 20	0	0	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	1
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	268	294	280	368

Keys: BBC1_FBS_HN = British-British Conversation 1 female British student's head nods,
BBC2_MBS_HN = British-British Conversation 2 male British student's head nods
BJC1_MJS_HN = British-Japanese Conversation 1 male Japanese student's head nods,
BJC2_MJS_HN = British-Japanese Conversation 2 male Japanese student's head nods



Keys: BBC1_FBS_HN = British-British Conversation 1 female British student's head nods,
BBC2_MBS_HN = British-British Conversation 2 male British student's head nods
BJC1_MJS_HN = British-Japanese Conversation 1 male Japanese student's head nods,
BJC2_MJS_HN = British-Japanese Conversation 2 male Japanese student's head nods

Figure B.1-2　Students' use of head nods

B.1.4 Comparing the use of HGs and HNs in British-British conversation and British-Japanese conversation

Table B.1.4-1 Use of HGs and HNs in BBC1

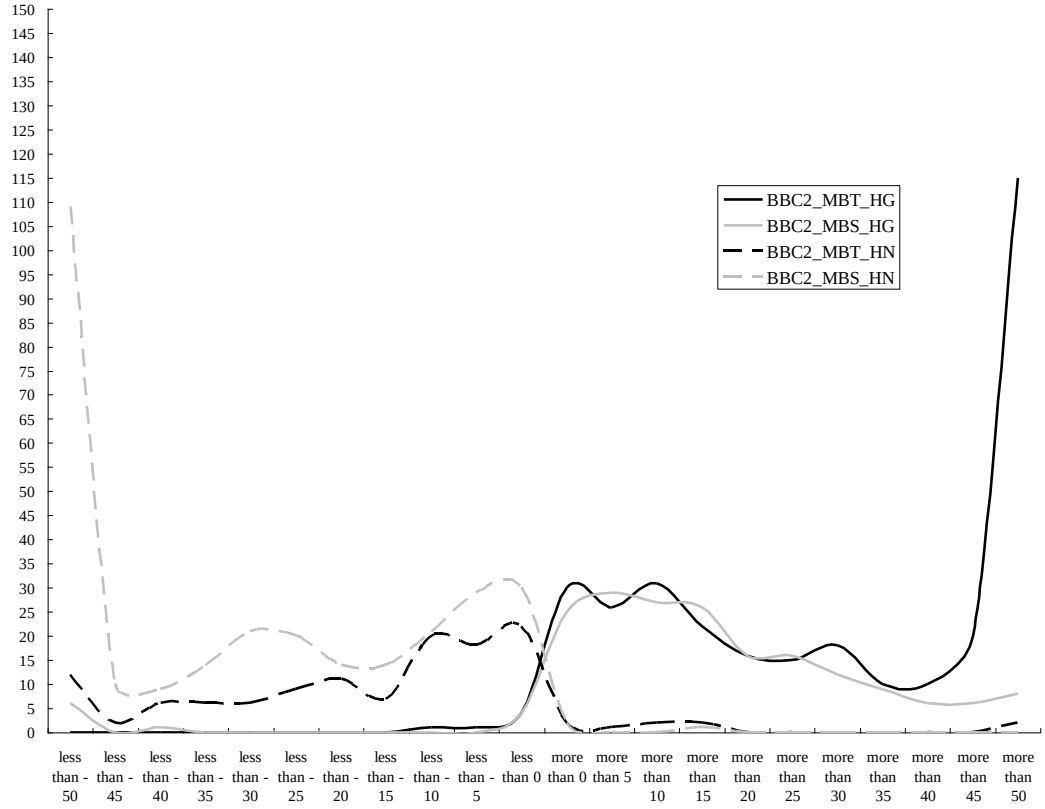
leadtime	BBC1_MBT_HG	BBC1_FBS_HG	BBC1_MBT_HN	BBC1_FBS_HN
less than -50	0	2	0	19
less than -45	0	0	0	9
less than -40	0	2	0	11
less than -35	0	1	3	10
less than -30	0	2	3	11
less than -25	0	0	5	23
less than -20	0	0	5	15
less than -15	0	1	10	22
less than -10	1	1	10	33
less than -5	1	2	30	40
less than 0	11	8	73	55
more than 0	138	75	14	18
more than 5	89	50	1	2
more than 10	76	23	0	0
more than 15	51	11	0	0
more than 20	32	14	0	0
more than 25	26	7	1	0
more than 30	30	5	0	0
more than 35	16	3	0	0
more than 40	10	0	0	0
more than 45	9	0	0	0
more than 50	33	0	0	0
	523	207	155	268

Keys: BBC1_MBT_HG = British-British Conversation 1 male British tutor's hand gestures,
BBC1_FBS_HG = British-British Conversation 1 female British student's hand gestures
BBC1_MBT_HN = British-British Conversation 1 male British tutor's head nods,
BBC1_FBS_HN = British-British Conversation 1 female British student's head nods

Table B.1.4-2 Use of HGs and HNs in BBC2

leadtime	BBC2_MBT_HG	BBC2_MBS_HG	BBC2_MBT_HN	BBC2_MBS_HN
less than -50	0	6	12	109
less than -45	0	0	2	10
less than -40	0	1	6	9
less than -35	0	0	6	14
less than -30	0	0	6	21
less than -25	0	0	9	20
less than -20	0	0	11	14
less than -15	0	0	7	14
less than -10	1	0	20	21
less than -5	1	0	18	29
less than 0	4	4	22	30
more than 0	30	25	2	2
more than 5	26	29	1	0
more than 10	31	27	2	0
more than 15	22	26	2	1
more than 20	16	16	0	0
more than 25	15	16	0	0
more than 30	18	12	0	0
more than 35	10	9	0	0
more than 40	10	6	0	0
more than 45	20	6	0	0
more than 50	115	8	2	0
	319	191	128	294

Keys: BBC2_MBT_HG = British-British Conversation 2 male British tutor's hand gestures,
BBC2_MBS_HG = British-British Conversation 2 male British student's hand gestures
BBC2_MBT_HN = British-British Conversation 2 male British tutor's head nods,
BBC2_MBS_HN = British-British Conversation 2 male British student's head nods



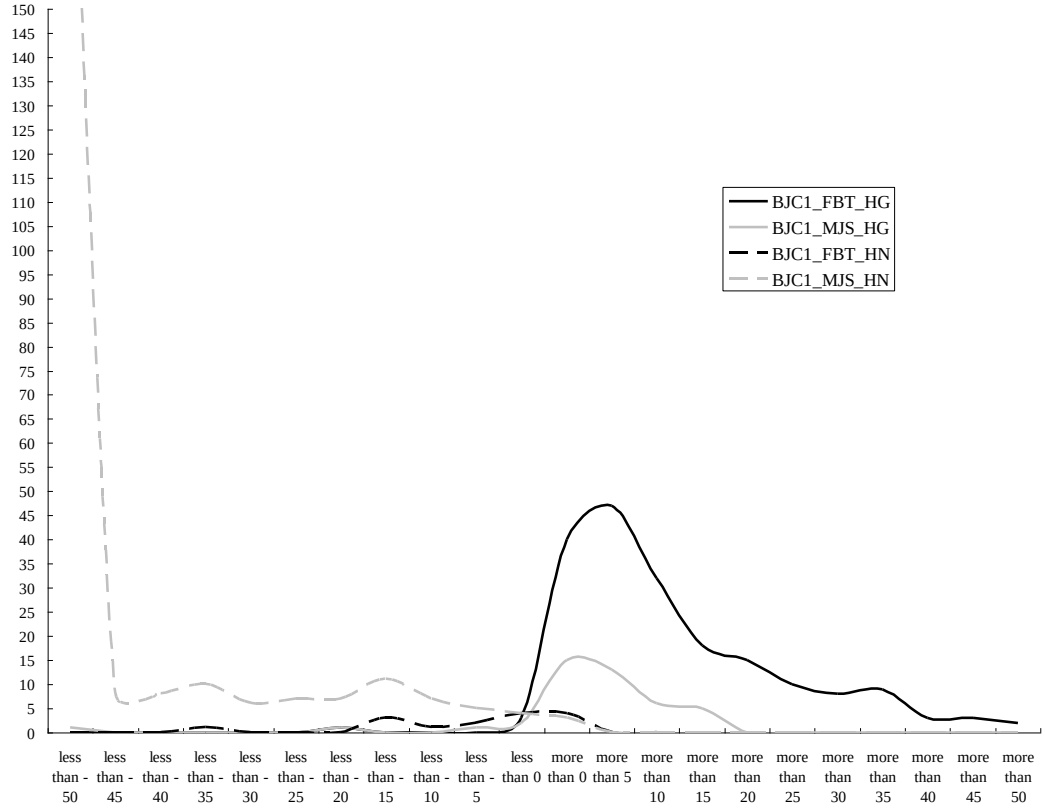
Keys: BBC2_MBT_HG = British-British Conversation 2 male British tutor's hand gestures,
BBC2_MBS_HG = British-British Conversation 2 male British student's hand gestures
BBC2_MBT_HN = British-British Conversation 2 male British tutor's head nods,
BBC2_MBS_HN = British-British Conversation 2 male British student's head nods

Figure B.1-3 Use of HGs and HNs in BBC2

Table B.1.4-3 Use of HGs and HNs in BJC1

leadtime	BJC1_FBT_HG	BJC1_MJS_HG	BJC1_FBT_HN	BJC1_MJS_HN
less than -50	0	1	0	204
less than -45	0	0	0	8
less than -40	0	0	0	8
less than -35	0	0	1	10
less than -30	0	0	0	6
less than -25	0	0	0	7
less than -20	1	1	0	7
less than -15	0	0	3	11
less than -10	0	0	1	7
less than -5	0	1	2	5
less than 0	3	2	4	4
more than 0	40	15	4	3
more than 5	47	13	0	0
more than 10	32	6	0	0
more than 15	18	5	0	0
more than 20	15	0	0	0
more than 25	10	0	0	0
more than 30	8	0	0	0
more than 35	9	0	0	0
more than 40	3	0	0	0
more than 45	3	0	0	0
more than 50	2	0	0	0
	191	44	15	280

Keys: BJC1_FBT_HG = British-Japanese Conversation 1 female British tutor's hand gestures,
BJC1_MJS_HG = British-Japanese Conversation 1 male Japanese student's hand gestures
BJC1_FBT_HN = British-Japanese Conversation 1 female British tutor's head nods,
BJC1_MJS_HN = British-Japanese Conversation 1 male Japanese student's head nods



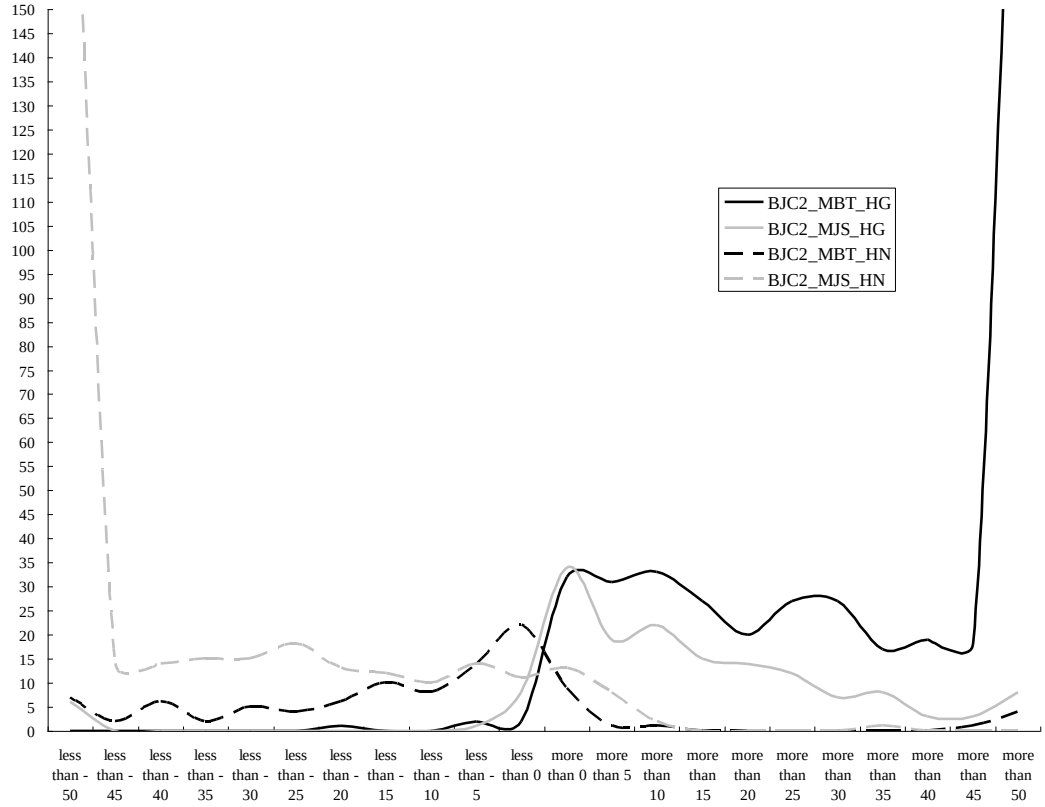
Keys: BJC1_FBT_HG = British-Japanese Conversation 1 female British tutor's hand gestures,
BJC1_MJS_HG = British-Japanese Conversation 1 male Japanese student's hand gestures
BJC1_FBT_HN = British-Japanese Conversation 1 female British tutor's head nods,
BJC1_MJS_HN = British-Japanese Conversation 1 male Japanese student's head nods

Figure B.1-4 Use of HGs and HNs in BJC1

Table B.1.4-4 Use of HGs and HNs in BJC2

leadtime	BJC2_MBT_HG	BJC2_MJS_HG	BJC2_MBT_HN	BJC2_MJS_HN
less than -50	0	6	7	208
less than -45	0	0	2	14
less than -40	0	0	6	14
less than -35	0	0	2	15
less than -30	0	0	5	15
less than -25	0	0	4	18
less than -20	1	0	6	13
less than -15	0	0	10	12
less than -10	0	0	8	10
less than -5	2	1	14	14
less than 0	2	8	22	11
more than 0	32	34	9	13
more than 5	31	19	1	8
more than 10	33	22	1	2
more than 15	27	15	0	0
more than 20	20	14	0	0
more than 25	27	12	0	0
more than 30	27	7	0	0
more than 35	17	8	0	1
more than 40	19	3	0	0
more than 45	18	3	1	0
more than 50	230	8	4	0
	486	160	102	368

Keys: BJC2_MBT_HG = British-Japanese Conversation 2 male British tutor's hand gestures,
BJC2_MJS_HG = British-Japanese Conversation 2 male Japanese student's hand gestures
BJC2_MBT_HN = British-Japanese Conversation 2 male British tutor's head nods,
BJC2_MJS_HN = British-Japanese Conversation 2 male Japanese student's head nods



Keys: BJC2_MBT_HG = British-Japanese Conversation 2 male British tutor's hand gestures,
BJC2_MJS_HG = British-Japanese Conversation 2 male Japanese student's hand gestures
BJC2_MBT_HN = British-Japanese Conversation 2 male British tutor's head nods,
BJC2_MJS_HN = British-Japanese Conversation 2 male Japanese student's head nods

Figure B.1-5 Use of HGs and HNs in BJC2

B.1.5 Placement of *erm*

Table B.1.5-1 Tutors' use of *erm*

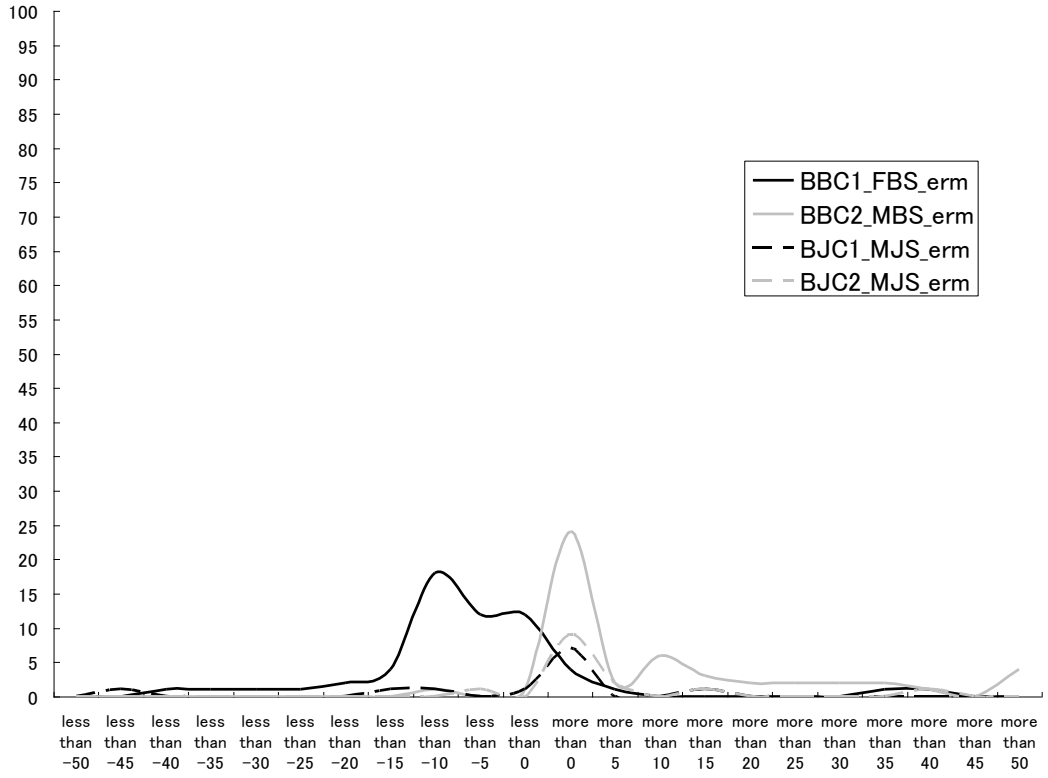
leadtime	BBC1_MBT_erm	BBC2_MBT_erm	BJC1_FBT_erm	BJC2_MBT_erm
less than −50	0	0	0	0
less than −45	0	0	0	0
less than −40	0	0	0	0
less than −35	0	0	0	0
less than −30	0	0	0	0
less than −25	0	0	0	0
less than −20	0	0	0	0
less than −15	0	0	0	0
less than −10	1	0	3	0
less than −5	1	0	0	1
less than 0	1	1	1	0
more than 0	16	7	5	16
more than 5	13	6	0	14
more than 10	7	10	1	3
more than 15	6	7	3	8
more than 20	7	4	3	4
more than 25	5	4	1	7
more than 30	2	3	0	4
more than 35	3	2	1	6
more than 40	1	1	0	4
more than 45	1	5	0	5
more than 50	3	19	0	54
	67	69	18	126

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Table B.1.5-2 Students' use of *erm*

leadtime	BBC1_FBS_erm	BBC2_MBS_erm	BJC1_MJS_erm	BJC2_MJS_erm
less than -50	0	0	0	0
less than -45	0	0	1	0
less than -40	1	0	0	0
less than -35	1	0	0	0
less than -30	1	0	0	0
less than -25	1	0	0	0
less than -20	2	0	0	0
less than -15	4	0	1	0
less than -10	18	1	1	0
less than -5	12	0	0	1
less than 0	12	1	1	0
more than 0	4	24	7	9
more than 5	1	2	0	2
more than 10	0	6	0	0
more than 15	0	3	1	1
more than 20	0	2	0	0
more than 25	0	2	0	0
more than 30	0	2	0	0
more than 35	1	2	0	0
more than 40	1	1	0	1
more than 45	0	0	0	0
more than 50	0	4	0	0
	59	50	12	14

Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student



Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

Figure B.1-6 Students' use of *erm*

B.1.6 Placement of *yeah*

Table B.1.6-1 Tutors' use of *yeah*

leadtime	BBC1_MBT_yeah	BBC2_MBT_yeah	BJC1_FBT_yeah	BJC2_MBT_yeah
less than -50	0	1	0	0
less than -45	0	1	0	0
less than -40	2	0	0	2
less than -35	0	0	0	0
less than -30	1	0	0	1
less than -25	2	1	1	1
less than -20	2	0	0	1
less than -15	7	0	0	4
less than -10	4	4	0	7
less than -5	12	4	0	6
less than 0	64	4	1	7
more than 0	52	4	14	17
more than 5	9	1	1	1
more than 10	0	0	1	3
more than 15	0	1	1	3
more than 20	2	1	0	0
more than 25	3	0	0	1
more than 30	1	0	0	0
more than 35	0	0	0	0
more than 40	1	0	0	0
more than 45	0	0	0	2
more than 50	1	0	0	5
	163	22	19	61

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Table B.1.6-2 Students' use of *yeah*

leadtime	BBC1_FBS_yeah	BBC2_MBS_yeah	BJC1_MJS_yeah	BJC2_MJS_yeah
less than -50	4	8	15	4
less than -45	3	4	0	0
less than -40	3	0	0	2
less than -35	4	7	0	0
less than -30	1	1	0	0
less than -25	3	0	1	1
less than -20	7	5	0	0
less than -15	6	5	1	0
less than -10	8	5	2	0
less than -5	18	4	1	0
less than 0	33	9	3	1
more than 0	29	15	10	1
more than 5	1	1	0	0
more than 10	0	0	0	0
more than 15	0	0	0	0
more than 20	0	1	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	1	0	0
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	120	66	33	9

Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

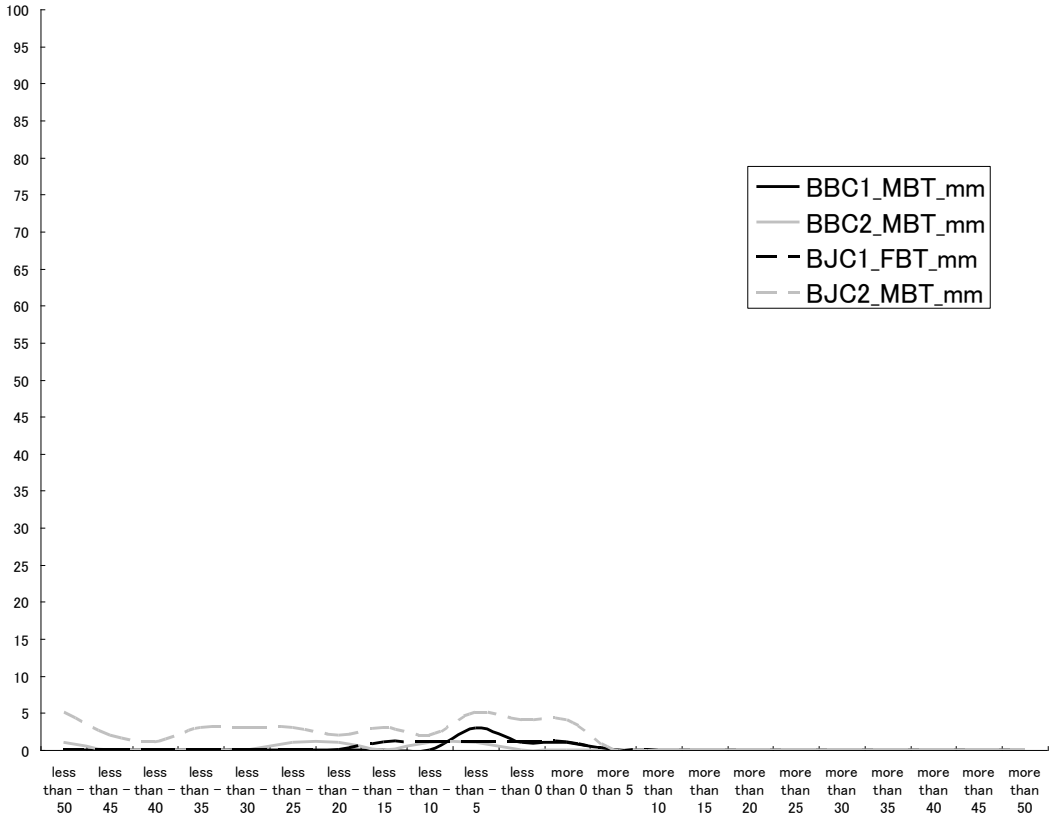
B.1.7 Placement of *mm* and *mhm*

Table B.1.7-1 Tutors' use of *mm*

leadtime	BBC1_MBT_mm	BBC2_MBT_mm	BJC1_FBT_mm	BJC2_MBT_mm
less than -50	0	1	0	5
less than -45	0	0	0	2
less than -40	0	0	0	1
less than -35	0	0	0	3
less than -30	0	0	0	3
less than -25	0	1	0	3
less than -20	0	1	0	2
less than -15	0	0	1	3
less than -10	0	1	1	2
less than -5	3	1	1	5
less than 0	1	0	1	4
more than 0	1	0	1	4
more than 5	0	0	0	0
more than 10	0	0	0	0
more than 15	0	0	0	0
more than 20	0	0	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	0
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	5	5	5	37

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Figure B.1-7　Tutors' use of *mm*

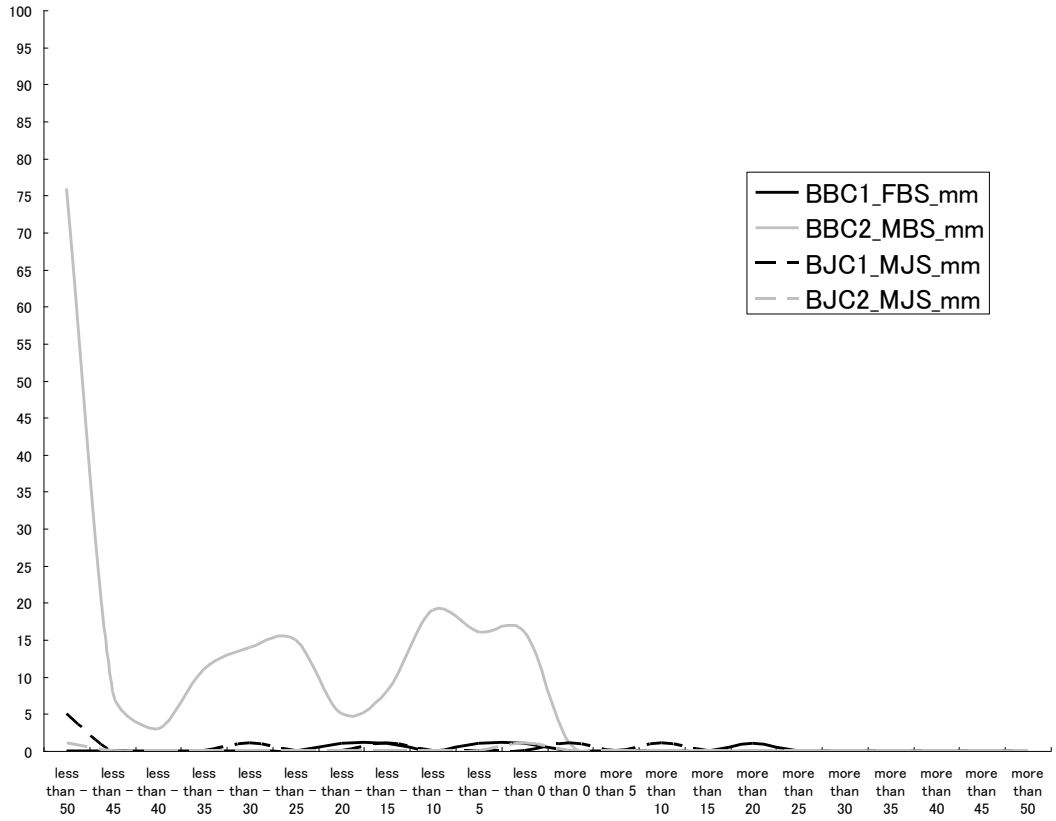


Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Table B.1.7-2 Students' use of *mm*

leadtime	BBC1_FBS_mm	BBC2_MBS_mm	BJC1_MJS_mm	BJC2_MJS_mm
less than -50	0	76	5	1
less than -45	0	8	0	0
less than -40	0	3	0	0
less than -35	0	11	0	0
less than -30	0	14	1	0
less than -25	0	15	0	0
less than -20	1	5	0	0
less than -15	1	8	1	0
less than -10	0	19	0	0
less than -5	1	16	0	0
less than 0	1	16	0	1
more than 0	0	1	1	0
more than 5	0	0	0	0
more than 10	0	0	1	0
more than 15	0	0	0	0
more than 20	1	0	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	0
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	5	192	9	2

Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student



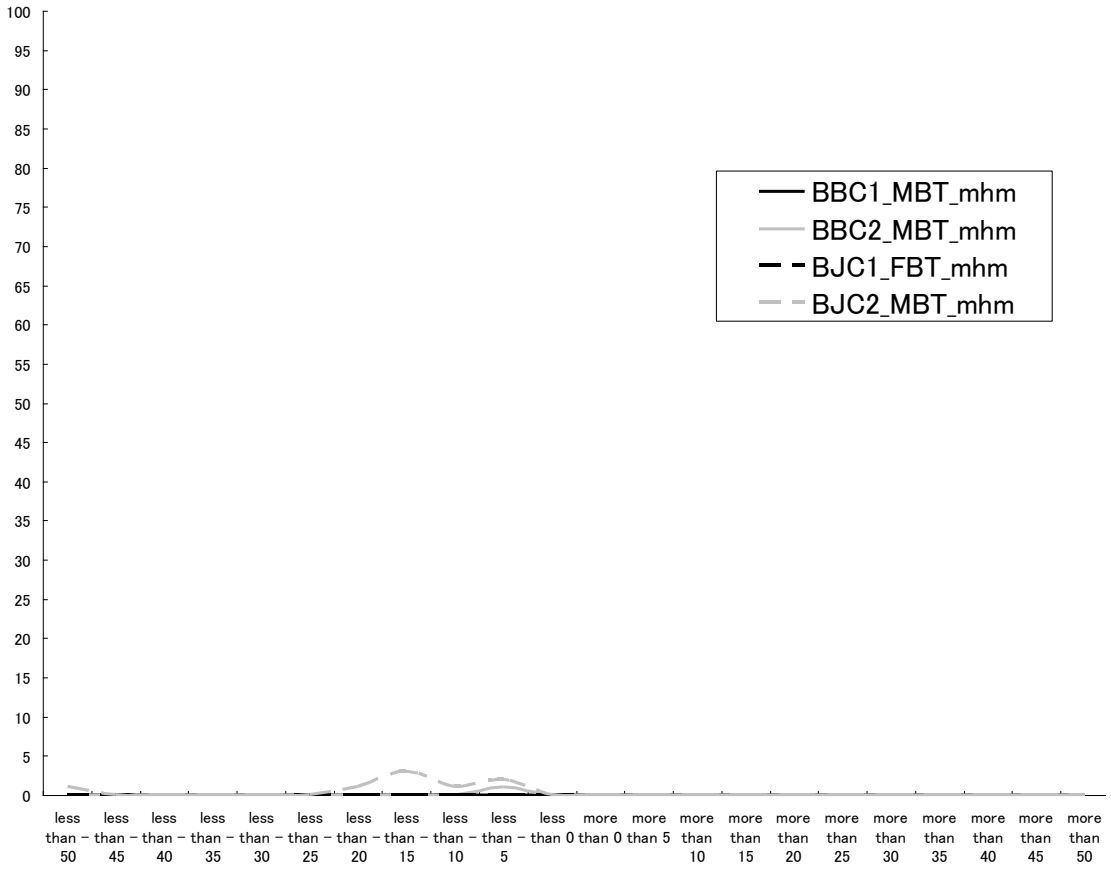
Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

Figure B.1-8 Students' use of *mm*

Table B.1.7-3 Tutors' use of *mhm*

leadtime	BBC1_MBT_mhm	BBC2_MBT_mhm	BJC1_FBT_mhm	BJC2_MBT_mhm
less than -50	0	0	0	1
less than -45	0	0	0	0
less than -40	0	0	0	0
less than -35	0	0	0	0
less than -30	0	0	0	0
less than -25	0	0	0	0
less than -20	0	0	0	1
less than -15	0	0	0	3
less than -10	0	0	0	1
less than -5	0	1	0	2
less than 0	0	0	0	0
more than 0	0	0	0	0
more than 5	0	0	0	0
more than 10	0	0	0	0
more than 15	0	0	0	0
more than 20	0	0	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	0
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	0	1	0	8

Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor



Keys: BBC1_MBT=British-British Conversation 1 male British tutor
BBC2_MBT=British-British Conversation 2 male British tutor
BJC1_FBT=British-Japanese Conversation 1 female British tutor
BJC2_MBT=British-Japanese Conversation 2 male British tutor

Figure B.1-9　Tutors' use of *mhm*

Table B.1.7-4 Students' use of mhm

leadtime	BBC1_FBS_mhm	BBC2_MBS_mhm	BJC1_MJS_mhm	BBC2_MBS_mhm
less than -50	0	0	76	9
less than -45	0	0	2	3
less than -40	0	1	3	2
less than -35	0	0	3	0
less than -30	0	0	2	2
less than -25	1	1	3	2
less than -20	0	0	5	2
less than -15	0	0	6	3
less than -10	0	0	5	1
less than -5	0	1	5	2
less than 0	0	0	4	1
more than 0	0	0	1	0
more than 5	0	0	0	0
more than 10	0	0	0	0
more than 15	0	0	0	0
more than 20	0	0	0	0
more than 25	0	0	0	0
more than 30	0	0	0	0
more than 35	0	0	0	0
more than 40	0	0	0	0
more than 45	0	0	0	0
more than 50	0	0	0	0
	1	3	115	27

Keys: BBC1_FBS=British-British Conversation 1 female British student
BBC2_MBS=British-British Conversation 2 male British student
BJC1_MJS=British-Japanese Conversation 1 male Japanese student
BJC2_MJST=British-Japanese Conversation 2 male Japanese student

B.2 Tables and figures from the turn structure analysis

B.2.1 Focusing on verbal response tokens with head nods

B.2.1.1 Tutors' use of verbal response tokens with head nods

Transcript B.2-1 BBC2_MBT's response tokens with head nods

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 33: 10	MBS_F	−19	0				For fi= well I'm hoping c= certainly finish this erm month and <$H> would </$H> still combine the writing up of that with the erm connecting the the mental health lit review+
00: 33: 11		−18	1				
00: 33: 12		−17	2				
00: 33: 13		−16	3				
00: 33: 14		−15	4				
00: 33: 15		−14	5	HN			
00: 33: 16		−13	6			HG	
00: 33: 17		−12	7				
00: 33: 18		−11	8	HN			
00: 33: 19		−10	9				
00: 33: 20		−9	10			HG	
00: 33: 21		−8	11	HN			
00: 33: 22		−7	12		Okay. Right.		
00: 33: 23		−6	13			HG	+because that'll be as you say just so many hours available a a week+
00: 33: 24		−5	14				
00: 33: 25		−4	15				
00: 33: 26		−3	16			HG	
00: 33: 27		−2	17		Okay.		+and so the two will be interchangeable. Er+
00: 33: 28		−1	18			HG	
00: 33: 29	MBT_F	0	−59	HG	The the mental the mental health review. This is where you are going to do literature review in terms of mental health+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MBT_T= male British tutor's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-2 BJC1_FBT's response tokens with head nods

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 18: 12	MJS_F	-12	0				So if teacher er point out to a student maybe the student have opportunity to speak English.
00: 18: 13		-11	1				
00: 18: 14		-10	2				
00: 18: 15		-9	3				
00: 18: 16		-8	4			HG	
00: 18: 17		-7	5	HN			
00: 18: 18		-6	6			SC/nose	
00: 18: 19		-5	7				
00: 18: 20		-4	8				
00: 18: 21		-3	9				
00: 18: 22		-2	10		Alright.		
00: 18: 23		-1	11	HN			
00: 18: 24	FBT_F	0	-18		okay. Are you going to do this research in Japan?		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript B.2-3 BJC2_MBT's response tokens with head nods

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 19: 27		−20	19	HN	mhm		+ or relationship subjects and objects +
00: 19: 28		−19	20			HG	
00: 19: 29		−18	21				
00: 19: 30		−17	22	HN	mhm		
00: 19: 31		−16	23				
00: 19: 32		−15	24				
00: 19: 33		−14	25				+ so I want to do formulation to authobiographical +
00: 19: 34		−13	26				
00: 19: 35		−12	27			HG	
00: 19: 36		−11	28				
00: 19: 37		−10	29	HN		HG	
00: 19: 38		−9	30		mhm mhm		
00: 19: 39		−8	31			HG	+ and I want to connect biographical aspects to +
00: 19: 40		−7	32			HG	
00: 19: 41		−6	33				
00: 19: 42		−5	34	HN	Mm.		
00: 19: 43		−4	35			HG	+ ethics form of language.
00: 19: 44		−3	36				
00: 19: 45		−2	37	HN			
00: 19: 46		−1	38			HG	
00: 19: 47	MBT_F	0	−170	HN	Mm. Mm. yeah absolutely it's good. it's really good. Erm or you you might want to there is I recommend you here just to+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.1.2 Students' use of verbal response tokens with head nods

Transcript B.2-4 Decrease of response tokens from BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 38: 20		51	−21		+ Erm erm all I said here is suggestions and expansions nothing certainly it's not <$G?> anyway. So that's fine +		
00: 38: 21		52	−20	HG			
00: 38: 22		53	−19			HN	
00: 38: 23		54	−18	HG			
00: 38: 24		55	−17			HN	
00: 38: 25		56	−16	HG			
00: 38: 26		57	−15			HN	
00: 38: 27		58	−14				
00: 38: 28		59	−13	HS		HN	
00: 38: 29		60	−12				
00: 38: 30		61	−11				
00: 38: 31		62	−10	HN			
00: 38: 32		63	−9		+ Erm do you want me to do the same again just email me the the chapter a couple days before.		
00: 38: 33		64	−8				
00: 38: 34		65	−7	HG			
00: 38: 35		66	−6				
00: 38: 36		67	−5	HG			
00: 38: 37		68	−4				
00: 38: 38		69	−3				
00: 38: 39		70	−2			HG	Yes I will +
00: 38: 40		71	−1		Yeah.		
00: 38: 41	MJS_F	−3	0				+ send at least two or three days before.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.2 Focusing on hand gestures

Transcript B.2-5 Hand gestures in BBC2_MBT

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 06: 01	MBS_F	−16	0				Well yeah quite impressed+
00: 06: 02		−15	1	HN	Yeah.		
00: 06: 03		−14	2				+yeah quite impressed with that. And what brought it home is that after our talk finished everyone upped sticks and there was <$E> laugh </$E> two or three people left and this poor presenter+
00: 06: 04		−13	3				
00: 06: 05		−12	4				
00: 06: 06		−11	5				
00: 06: 07		−10	6			HG/SC/nose	
00: 06: 08		−9	7			HG	
00: 06: 09		−8	8				
00: 06: 10		−7	9		Oh yeah.		
00: 06: 11		−6	10				+who was following us. just had this very miniature audience+
00: 06: 12		−5	11			HG	
00: 06: 13		−4	12				
00: 06: 14		−3	13				
00: 06: 15		−2	14		Well I mean+		+that was left.
00: 06: 16		−1	15	HG		HN	
00: 06: 17	MBT_F	0	−41		you're at a you're still at early stages aren't you? Really?		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

Figure B.2-1 BBC2_MBT's HG

Transcript B.2-6 Hand gestures in BBC2_MBS

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 28: 07		74	−9		+in the process. So I wouldn't sort of dismiss that as+		
00: 28: 08		75	−8				
00: 28: 09		76	−7				
00: 28: 10		77	−6	HG			
00: 28: 11		78	−5				No no.
00: 28: 12		79	−4		+as just s= su= a function+	HN	
00: 28: 13		80	−3	HG	+and and then leave it to one side.		Sure yeah.
00: 28: 14		81	−2			HN	
00: 28: 15		82	−1				
00: 28: 16	MBS_F	−4	0				Sure. That's something that I think's definitely going in methodology+
00: 28: 17		−3	1			HG	
00: 28: 18		−2	2				
00: 28: 19		−1	3	HN	Absolutely absolutely.		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

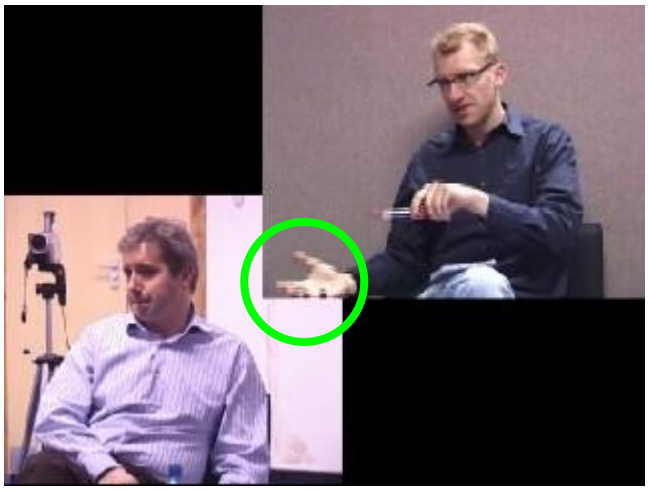


Figure B.2-2 BBC2_MBS's HG (1)

Figure B.2-3 BBC2_MBS's HG (2)

Transcript B.2-7 Hand gestures in BJC2_MBT and BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 13: 53		18	−18		+ John Jennet has got some erm Jo= yeah Jo= John Jennet yeah erm Gerald Jennet. He's got some erm he's got some <$G?> looking forward <$G?>.		
00: 13: 54		19	−17				
00: 13: 55		20	−16				
00: 13: 56		21	−15				
00: 13: 57		22	−14				
00: 13: 58		23	−13				
00: 13: 59		24	−12				
00: 14: 00		25	−11				
00: 14: 01		26	−10				
00: 14: 02		27	−9				
00: 14: 03		28	−8				
00: 14: 04		29	−7				
00: 14: 05		30	−6	HG		HN	
00: 14: 06		31	−5	SC/forehead			
00: 14: 07		32	−4				
00: 14: 08		33	−3				
00: 14: 09		34	−2				
00: 14: 10		35	−1				
00: 14: 11	MJS_F	−23	0			HG	So this er sorry.
00: 14: 12		−22	1		Yeah.		
00: 14: 13		−21	2	HG			It's better I clarify the chapter +
00: 14: 14		−20	3				
00: 14: 15		−19	4			HG	
00: 14: 16		−18	5	HN			
00: 14: 17		−17	6		Yeah.	HG	
00: 14: 18		−16	7			HG	
00: 14: 19		−15	8				+ it will be about anticipation +
00: 14: 20		−14	9				
00: 14: 21		−13	10			HG	
00: 14: 22		−12	11	HN	Yeah.		
00: 14: 23		−11	12				
00: 14: 24		−10	13			HG	+ I will discuss it's like in next chapter +
00: 14: 25		−9	14			HG	
00: 14: 26		−8	15	HN			
00: 14: 27		−7	16			HG	
00: 14: 28		−6	17				
00: 14: 29		−5	18		Yeah.	HG	+ like that I should write in the chapter?
00: 14: 30		−4	19				
00: 14: 31		−3	20	HN		HG	
00: 14: 32		−2	21			HG	
00: 14: 33		−1	22				
00: 14: 34	MBT_F	0	−214	HN/HG	Y= yes you should go. it's up to your topic in the next chapter. Say you could almost say that in the chapter two I'll be exploring this further +		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Figure B.2-4　BJC2_MJS’s HG

Figure B.2-5　BJC2_MBT’s HG

B.2.3 Turn-structural episodes

B.2.3.1 Turn-structural episodes and listenership in BBC1_MBT

Transcript B.2-8 Episode 3 in BBC1_MBT

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 09: 49	MBT_F	0	-32		Yeah. So the the but you're you're prodding the the audience person to to think of it to be persuaded or not.		
00: 09: 50		1	-31	HG			
00: 09: 51		2	-30				
00: 09: 52		3	-29				
00: 09: 53		4	-28	HG			
00: 09: 54		5	-27	HG		HN	
00: 09: 55		6	-26				
00: 09: 56		7	-25			HN	Yeah
00: 09: 57		8	-24				
00: 09: 58		9	-23		Yeah okay.		erm.
00: 09: 59	Pause	-4	-22		<$E> pause </$E>		
00: 10: 00		-3	-21				
00: 10: 01		-2	-20	HN			
00: 10: 02		-1	-19				
00: 10: 03	MBT_F	0	-18		Right. So there's loads of stuff on conceptual metaphor.		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.3.2 Turn-structural episodes and listenership in BBC1_FBS

Table B.2.3-1 Turn-structural episodes and listenership in BBC1_FBS

	BBC1_FBS Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	17	18%	22.71	17.12	293.03
Episode 2	37	40%	16.54	20.99	440.57
Episode 3	1	1%	2.00	0.00	0.00
Episode 4	3	3%	8.00	1.41	2.00
Episode 5	33	35%	10.45	13.72	188.19
Episode 6	0	0%	--	--	--
Episode 7	1	1%	32.00	0.00	0.00
Unclassified	1	1%	--	--	--
Total	93	100%			

Transcript B.2-9 Episode 2 in BBC1_FBS

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 10: 51	MBT_F	0	-6		Yeah yeah. There is a famous old article in 1982 called "Are Scientific Analogies Metaphors?"+		
00: 10: 52		1	-5	HN			
00: 10: 53		2	-4				
00: 10: 54		3	-3				
00: 10: 55		4	-2				
00: 10: 56		5	-1				
00: 10: 57	FBS_F	-1	0			HG	Yeah I've read that+
00: 10: 58	MBT_F	0	-6		+by Dierdre Gentner I think. Yeah. Which is one of the earliest sort of conceptual mapping things that was about the same time as the original Lakoff and Johnson+		
00: 10: 59		1	-5	HG			
00: 11: 00		2	-4	HG			
00: 11: 01		3	-3			HN	
00: 11: 02		4	-2				
00: 11: 03		5	-1	HG			
00: 11: 04	FBS_F	-10	0	SC/eye		HN	Yeah and there is a whole there's quite a lot of essays in the Andrew Ortony book about+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-10 Episode 4 in BBC1_FBS

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 29: 52		12	-2		There's a journal called metaphor and symbol erm+		
00: 29: 53		13	-1				
00: 29: 54	FBS_F	-8	0			HN	Yeah I found that online I do= I don't +
00: 29: 55		-7	1				
00: 29: 56		-6	2				
00: 29: 57		-5	3		Right yeah.	HG	+ you can actually get hold of it online but
00: 29: 58		-4	4				
00: 29: 59		-3	5				
00: 30: 00		-2	6		Okay.		you have to subscribe to it or something
00: 30: 01		-1	7	SC/hair			
00: 30: 02	MBT_F	0	-9		How annoying. I wonder who owns it. I wonder if Vernon might take it.		
00: 30: 03		1	-8				
00: 30: 04		2	-7				
00: 30: 05		3	-6				
00: 30: 06	Pause	-7	-5		<$E> pause </$E>		
00: 30: 07		-6	-4				
00: 30: 08		-5	-3				
00: 30: 09		-4	-2				
00: 30: 10		-3	-1				
00: 30: 11	FBS_F	-2	0		Yeah.		I'll check again cos it'll say where it's <$G?>.
00: 30: 12		-1	1				
00: 30: 13	MBT_F	0	-53		well it'll be you'll be able to get hold of index for it see if there's anything particularly worth having and if you can't download it online or find it at a university library that's nearby just just order it from the British Library and they'll photocopy		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking, SC/hair= Self comfort with hair

Transcript B.2-11 Episode 7 in BBC1_FBS

Timeline	Floor	BBC1_MBT_leadtime	BBC1_FBS_leadtime	BBC1_MBT_gesture	BBC1_MBT_Transcript	BBC1_FBS_gesture	BBC1_FBS_Transcript
00: 22: 07	MBT_F	0	-18	HG	Sorry I'm just wondering whether mm whether it's worth at this point trying to either exclude either one or the other of those domains or explicitly recognise that you're doing a comparative study.		
00: 22: 08		1	-17	SC/mouth			
00: 22: 09		2	-16	HG			
00: 22: 10		3	-15				
00: 22: 11		4	-14	HG			
00: 22: 12		5	-13				
00: 22: 13		6	-12	HG			
00: 22: 14		7	-11				
00: 22: 15		8	-10			HN	
00: 22: 16		9	-9				
00: 22: 17		10	-8	HG			
00: 22: 18		11	-7				
00: 22: 19		12	-6	HG			
00: 22: 20		13	-5			HN	
00: 22: 21	Pause	-12	-4		<$E> pause </$E>		
00: 22: 22		-11	-3				
00: 22: 23		-10	-2				
00: 22: 24		-9	-1		Mm.		
00: 22: 25	FBS_F	-8	0				I mean if it's if I do a comparative study I mean is that something I could do feasibly in the word count or is in the size+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, FBS_T= female British student's floor-taking, MBT_F = male British tutor's floor-taking, SC/hair= Self comfort with hair

B.2.3.3 Turn-structural episodes and listenership in BBC2_MBT

Table B.2.3-2 Turn-structural episodes and listenership in BBC2_MBT

	BBC2_MBT Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	10	28%	27.00	36.78	1352.57
Episode 2	20	56%	16.60	12.66	160.24
Episode 3	0	0%	--	--	--
Episode 4	0	0%	--	--	--
Episode 5	3	8%	7.00	4.90	24.00
Episode 6	0	0%	--	--	--
Episode 7	0	0%	--	--	--
Unclassified	3	8%	--	--	--
Total	36	100%			

Transcript B.2-12 Episode 1 in BBC2_MBT

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 06: 58	MBS_F	-33	0				Yeah and and it's nice just to go to a conference like that+
00: 06: 59		-32	1			SC/hands	
00: 07: 00		-31	2				
00: 07: 01		-30	3				
00: 07: 02		-29	4	HN	Mm.		+and give a paper because it's something that if one goes into academia has to do and erm yeah it's always er a big step to take.
00: 07: 03		-28	5				
00: 07: 04		-27	6				
00: 07: 05		-26	7				
00: 07: 06		-25	8			HG	
00: 07: 07		-24	9	HN			
00: 07: 08		-23	10				
00: 07: 09		-22	11				
00: 07: 10		-21	12				
00: 07: 11		-20	13		Sure.		
00: 07: 12		-19	14				And some people do it easily and some don't and so it's nice to have a few under the belt as it were.
00: 07: 13		-18	15			HG	
00: 07: 14		-17	16	HN			
00: 07: 15		-16	17				
00: 07: 16		-15	18				
00: 07: 17		-14	19				
00: 07: 18		-13	20				
00: 07: 19		-12	21		Yeah.		
00: 07: 20		-11	22				And er that's good. Also the feedback at the end was very good. Lots of questions which I think er was reassuring. It shows that people are interested or at least feigning interest.
00: 07: 21		-10	23				
00: 07: 22		-9	24			SC/ear	
00: 07: 23		-8	25				
00: 07: 24		-7	26	HN			
00: 07: 25		-6	27				
00: 07: 26		-5	28				
00: 07: 27		-4	29				
00: 07: 28		-3	30	HN			
00: 07: 29		-2	31				
00: 07: 30		-1	32		Excellent.		
00: 07: 31	MBT_F	0	-15		And is it December? November?+		Erm.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-13 Episode 2 in BBC2_MBT: Sample 1

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 02: 11		13	−5		+in terms of the chairing which I think is good for confidence, isn't it?		Mm yeah.
00: 02: 12		14	−4	HG		HN	
00: 02: 13		15	−3	HG			
00: 02: 14		16	−2				
00: 02: 15		17	−1				Definitely+
00: 02: 16	MBS_F	−9	0		Yeah		+yeah yeah that's right and it's something that one has to do and er.
00: 02: 17		−8	1				
00: 02: 18		−7	2				
00: 02: 19		−6	3			HG	
00: 02: 20		−5	4	HN			
00: 02: 21		−4	5	HN	Yeah.		Yeah it's just nice to be part of er the the team there <$G?>.
00: 02: 22		−3	6			HG	
00: 02: 23		−2	7	HN			
00: 02: 24		−1	8				
00: 02: 25	MBT_F	0	−15		That's good especially you know for a home conference.		
00: 02: 26		1	−14	HG			
00: 02: 27		2	−13			HN	Mm.
00: 02: 28		3	−12		It did. It was very friendly+		
00: 02: 29		4	−11	HG		HN	Mm.
00: 02: 30		5	−10	HG	+but it was also very interesting, wasn't it?		
00: 02: 31		6	−9			HN	
00: 02: 32		7	−8		What did you make of er John Sinclair'a erm +	HN	Definitely.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-14 Episode 2 in BBC2_MBT: Sample 2

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 36: 32	MBS_F	−16	0			HG/SC/arm	Well I thi= think one thing is to erm identify er key words which communicate erm mental health issues. Erm and it seems to me that there's a fair amount that communicates some form or shape of depression. Erm+
00: 36: 33		−15	1				
00: 36: 34		−14	2				
00: 36: 35		−13	3				
00: 36: 36		−12	4				
00: 36: 37		−11	5	HN			
00: 36: 38		−10	6				
00: 36: 39		−9	7	HN			
00: 36: 40		−8	8				
00: 36: 41		−7	9				
00: 36: 42		−6	10				
00: 36: 43		−5	11				
00: 36: 44		−4	12				
00: 36: 45		−3	13				
00: 36: 46		−2	14	HN			
00: 36: 47		−1	15				
00: 36: 48	MBT_F	0	−1		Yeah. Can you give me some examples?		
00: 36: 49	MBS_F	−36	0				Yeah there's quite a lot lot of metaphoric use. Low and h= high and erm people talking about hating themselves+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.3.4 Turn-structural episodes and listenership in BBC2_MBS

Table B.2.3-3 Turn-structural episodes and listenership in BBC2_MBS

	BBC2_MBS Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	9	26%	65.56	52.74	2781.58
Episode 2	6	18%	58.67	49.99	2499.22
Episode 3	1	3%	1.00	0.00	0.00
Episode 4	0	0%	--	--	--
Episode 5	18	53%	33.28	38.67	1495.09
Episode 6	0	0%	--	--	--
Episode 7	0	0%	--	--	--
Unclassified	0	0%	--	--	--
Total	34	100%			

Transcript B.2-15 Episode 2 in BBC2_MBS

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 22: 37		50	−10		+that erm practitioners are very aware of+		
00: 22: 38		51	−9				
00: 22: 39		52	−8				
00: 22: 40		53	−7			HN	
00: 22: 41		54	−6				
00: 22: 42		55	−5			HN	Mm.
00: 22: 43		56	−4		+of their words+		Mm. Mm.
00: 22: 44		57	−3		+and the impact of their words so er+		
00: 22: 45		58	−2			HN	
00: 22: 46		59	−1				
00: 22: 47	MBS_F	−3	0				They seemed quite interested in corpus tools.
00: 22: 48		−2	1			HG	
00: 22: 49		−1	2				
00: 22: 50	MBT_F	0	−115		Well, I think I think I think Srikant would probably er he would probably admit that as a data management tool+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-16 Episode 5 in BBC2_MBS

Timeline	Floor	BBC2_MBT_leadtime	BBC2_MBS_leadtime	BBC2_MBT_gesture	BBC2_MBT_Transcript	BBC2_MBS_gesture	BBC2_MBS_Transcript
00: 09: 57	MBT_F	0	−21		But you always pick up stuff+	SC/chin	
00: 09: 58		1	−20	HG			
00: 09: 59		2	−19		+because you know erm that's the nature of searching erm that however good our search approach is we will miss some items or new stuff will be produced as well and you you need to keep+	HN	Mm.
00: 10: 00		3	−18				
00: 10: 01		4	−17				
00: 10: 02		5	−16				
00: 10: 03		6	−15	HG			
00: 10: 04		7	−14				
00: 10: 05		8	−13			HN	
00: 10: 06		9	−12				
00: 10: 07		10	−11			HN	
00: 10: 08		11	−10			HN	
00: 10: 09		12	−9	HG			
00: 10: 10		13	−8			HN	
00: 10: 11		14	−7	HG			
00: 10: 12		15	−6		+updating your review.	HN	Mm.
00: 10: 13		16	−5				Mm.
00: 10: 14		17	−4	HG	What's happening with the methodology? Can you just give me a=		
00: 10: 15		18	−3				
00: 10: 16		19	−2				
00: 10: 17		20	−1				
00: 10: 18	MBS_F	−47	0			SC/arm	Yeah. Erm. I I've I'm focusing erm at the moment on er on corpus linguistics as that seemed to be I think quite a major part erm of the analysis. Certainly the an initial part+

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MBS_T= male British student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.3.5 Turn-structural episodes and listenership in BJC1_FBT

Table B.2.3-4 Turn-structural episodes and listenership in BJS1_FBT

	BJC1_FBT Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	15	13%	11.07	13.14	172.60
Episode 2	3	3%	3.33	1.25	1.56
Episode 3	17	14%	4.73	3.59	12.86
Episode 4	2	2%	16.00	2.00	4.00
Episode 5	0	0%	--	--	--
Episode 6	13	11%	5.50	3.55	12.58
Episode 7	0	0%	--	--	--
Unclassified	69	58%	--	--	--
Total	119	100%			

Transcript B.2-17 Episode 1 in BJC1_FBT

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 18: 12	MJS_F	-12	0				So if teacher er point out to a student maybe the student have opportunity to speak English.
00: 18: 13		-11	1				
00: 18: 14		-10	2				
00: 18: 15		-9	3				
00: 18: 16		-8	4			HG	
00: 18: 17		-7	5	HN			
00: 18: 18		-6	6			SC/nose	
00: 18: 19		-5	7				
00: 18: 20		-4	8				
00: 18: 21		-3	9				
00: 18: 22		-2	10		Alright.		
00: 18: 23		-1	11	HN			
00: 18: 24	FBT_F	0	-18		okay. Are you going to do this research in Japan?		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript B.2-18 Episode 2 in BJC1_FBT

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 17: 40		2	-15		Okay so okay from this I have a picture of Japanese class working very quietly.		
00: 17: 41		3	-14				
00: 17: 42		4	-13				
00: 17: 43		5	-12				
00: 17: 44		6	-11				
00: 17: 45		7	-10	HT/HG			
00: 17: 46		8	-9				
00: 17: 47		9	-8	HG		HG	mhm.
00: 17: 48		10	-7			HN	
00: 17: 49		11	-6		Is that right?	HN	
00: 17: 50		12	-5		Teacher says "okay open your book and do exercise three."		
00: 17: 51		13	-4	HG			
00: 17: 52		14	-3				
00: 17: 53		15	-2				Y= yeah.
00: 17: 54		16	-1				
00: 17: 55	MJS_F	-3	0		And they are=		Yeah I have experience like that.
00: 17: 56		-2	1	HG			
00: 17: 57		-1	2				
00: 17: 58	FBT_F	0	-1		So there is no speaking.	HG	
00: 17: 59	MJS_F	-25	0	HS/HG			No especially i= if students want to say something er they can do. But normally er just teacher says something.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript B.2-19 Episode 4 in BJC1_FBT

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 15: 49		25	-19	HG	+ tell me what you are putting here?		
00: 15: 50		26	-18			HN	
00: 15: 51		27	-17				Ah well.
00: 15: 52	Pause	-14	-16		<$E> pause </$E>	SC/chin	
00: 15: 53		-13	-15				
00: 15: 54		-12	-14				
00: 15: 55		-11	-13				
00: 15: 56		-10	-12				
00: 15: 57		-9	-11				
00: 15: 58		-8	-10				
00: 15: 59		-7	-9				
00: 16: 00		-6	-8				
00: 16: 01		-5	-7				
00: 16: 02		-4	-6				
00: 16: 03		-3	-5				
00: 16: 04		-2	-4				
00: 16: 05		-1	-3				
00: 16: 06	FBT_F	0	-2		What is your research question? Have you got		
00: 16: 07		1	-1				
00: 16: 08	MJS_F	-14	0				Ah yeah. How fre= yeah here how frequent the interaction between teacher and teenager students in a large class.
00: 16: 09		-13	1				
00: 16: 10		-12	2				
00: 16: 11		-11	3				
00: 16: 12		-10	4				
00: 16: 13		-9	5			HG	
00: 16: 14		-8	6				
00: 16: 15		-7	7				
00: 16: 16		-6	8				
00: 16: 17		-5	9			SC/chin	
00: 16: 18		-4	10				
00: 16: 19		-3	11		yeah.		
00: 16: 20	Pause	-2	-96		<$E> pause </$E>		
00: 16: 21		-1	-95				
00: 16: 22	FBT_F	0	-94		I think that's needed to be much more clearly stated.		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking, SC/chin = self comfort with chin

B.2.3.6 Turn-structural episodes and listenership in BJC1_MJS

Table B.2.3-5 Turn-structural episodes and listenership in BJC1_MJS

	BJC1_MJS Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	6	25%	58. 67	42.39	1797.22
Episode 2	3	13%	120.00	109.33	11954.00
Episode 3	2	8%	3.00	1.00	1.00
Episode 4	1	4%	18.00	0.00	0.00
Episode 5	9	38%	105.67	101.24	10248.89
Episode 6	0	0%	--	--	--
Episode 7	3	13%	44.50	37.50	1406.25
Unclassified	0	0%	--	--	--
Total	24	100%			

Transcript B.2-20 Episode 5 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 28: 25		6	-25		Do you mean that others may not others may not have opportunity to interact with a teacher at all? So for these students +		
00: 28: 26		7	-24				
00: 28: 27		8	-23				
00: 28: 28		9	-22				
00: 28: 29		10	-21				
00: 28: 30		11	-20				
00: 28: 31		12	-19				
00: 28: 32		13	-18				
00: 28: 33		14	-17				
00: 28: 34		15	-16				
00: 28: 35		16	-15	HT			
00: 28: 36		17	-14				mhm.
00: 28: 37		18	-13	HT	+ studying quietly is good or these students actually prefer to study quietly and therefore there are unwillingness to speak it's reinforced.		
00: 28: 38		19	-12				
00: 28: 39		20	-11				
00: 28: 40		21	-10	HG			
00: 28: 41		22	-9	HG			
00: 28: 42		23	-8				
00: 28: 43		24	-7	HG			
00: 28: 44		25	-6				
00: 28: 45		26	-5				
00: 28: 46		27	-4	HG			
00: 28: 47		28	-3				
00: 28: 48		29	-2	HG			
00: 28: 49		30	-1				
00: 28: 50	MJS_F	-17	0				Mhm. yeah I think not always erm some students like study individually and quietly and some students want to speak English <$G?> yeah.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

Transcript B.2-21 Episode 7 in BJC1_MJS

Timeline	Floor	BJC1_FBT_leadtime	BJC1_MJS_leadtime	BJC1_FBT_gesture	BJC1_FBT_Transcript	BJC1_MJS_gesture	BJC1_MJS_Transcript
00: 22: 24	FBT_F	0	-7		But do they not do that because they're working quietly?		
00: 22: 25		1	-6				
00: 22: 26		2	-5	HG			
00: 22: 27		3	-4				
00: 22: 28	Pause	-18	-3		<$E> pause </$E>		mhm.
00: 22: 29		-17	-2			HN	
00: 22: 30		-16	-1				
00: 22: 31	MJS_F	-15	0				I think they have chance to talk with teacher but mm.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, FBT_F = female British tutor's floor-taking

B.2.3.7 Turn-structural episodes and listenership in BJC2_MBT

Table B.2.3-6 Turn-structural episodes and listenership in BJC2_MBT

	BJC2_MBT Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	14	42%	27.93	23.38	546.78
Episode 2	3	9%	11.33	7.41	54.89
Episode 3	5	15%	28.30	48.91	2391.76
Episode 4	0	0%	--	--	--
Episode 5	4	12%	8.00	8.77	77.00
Episode 6	3	9%	4.33	1.25	1.56
Episode 7	0	0%	--	--	--
Unclassified	4	12%	--	--	--
Total	33	100%			

Transcript B.2-22 Episode 1 in BJC2_MBT

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 30: 57	MJS_F	−38	0			HG	The the the nature of language is arbitrary he says +
00: 30: 58		−37	1	HN			
00: 30: 59		−36	2			HG	
00: 31: 00		−35	3	HN			
00: 31: 01		−34	4		Yeah.		
00: 31: 02		−33	5			HG	
00: 31: 03		−32	6				+ but er the relationship between text and history and er text and it's form+
00: 31: 03		−32	6				+ but er the relationship between text and history and er text and it's form+
00: 31: 04		−31	7				
00: 31: 05		−30	8				
00: 31: 06		−29	9			HG	
00: 31: 07		−28	10			HG	
00: 31: 08		−27	11				
00: 31: 09		−26	12			HG	
00: 31: 10		−25	13				
00: 31: 11		−24	14				
00: 31: 12		−23	15				
00: 31: 13		−22	16			HG	
00: 31: 14		−21	17			HG	
00: 31: 15		−20	18	HN	Mm.		
00: 31: 16		−19	19			HG	+require the reader to read in the certain frame +
00: 31: 17		−18	20				
00: 31: 18		−17	21				
00: 31: 19		−16	22			HG	
00: 31: 20		−15	23				
00: 31: 21		−14	24	HN		HG	
00: 31: 22		−13	25		Mm. Yeah.		
00: 31: 23		−12	26				So this kind of requirement from text +
00: 31: 24		−11	27				
00: 31: 25		−10	28				
00: 31: 26		−9	29				
00: 31: 27		−8	30			HG	
00: 31: 28		−7	31				
00: 31: 29		−6	32				
00: 31: 30		−5	33		Mm.		
00: 31: 31		−4	34				+ I call im= imperative author.
00: 31: 32		−3	35				
00: 31: 33		−2	36	HN		HG	
00: 31: 34		−1	37			HN	
00: 31: 35	MBT_F	0	−112	HN/HG	Mm. Great. Yeah again I think just couple of couple of sentences +		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

Transcript B.2-23 Episode 3 in BJC2_MBT

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 27: 41		2	−165	SC/chin	+erm which is interesting but but again I think there're quite few concepts.+		
00: 27: 42		3	−164	HG			
00: 27: 43		4	−163				
00: 27: 44		5	−162	HG			
00: 27: 45		6	−161				
00: 27: 46		7	−160		+you know what I mean. There're few process going on back here on page five erm I mean you can add about this and this <$G?>.		
00: 27: 47		8	−159				
00: 27: 48		9	−158				
00: 27: 49		10	−157				
00: 27: 50		11	−156	HG			
00: 27: 51		12	−155				
00: 27: 52		13	−154				
00: 27: 53		14	−153				
00: 27: 54		15	−152				
00: 27: 55		16	−151				
00: 27: 56		17	−150				
00: 27: 57	Pause	−4	−149		<$E> pause </$E>		
00: 27: 58		−3	−148	SC/nose			
00: 27: 59		−2	−147				
00: 28: 00		−1	−146				
00: 28: 01	MBT_F	0	−145		I mean it is just a small example of general points really.+		

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking. SC/nose = self-comfort with nose

Transcript B.2-24　　Episode 6 in BJC2_MBT

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 26: 33		50	−25	HG	+and also what readers doing in terms of <$H> inacting </$H> emotion you know what that means? you know +		
00: 26: 34		51	−24				
00: 26: 35		52	−23	HG			
00: 26: 36		53	−22			HN	
00: 26: 37		54	−21	HG		HN	
00: 26: 38		55	−20			HN	
00: 26: 39		56	−19	HG	+ So I think that might be worth worth exploring the triangle concepts. You've got staging, <$H> inactment </$H> and participation yeah which +		
00: 26: 40		57	−18				
00: 26: 41		58	−17				
00: 26: 42		59	−16				
00: 26: 43		60	−15	HG			
00: 26: 44		61	−14				
00: 26: 45		62	−13	HG			
00: 26: 46		63	−12				
00: 26: 47		64	−11				
00: 26: 48		65	−10				
00: 26: 49		66	−9				
00: 26: 50		67	−8	HG			
00: 26: 51		68	−7	SC/mouth			
00: 26: 52		69	−6				
00: 26: 53		70	−5				
00: 26: 54	Pause	−3	−4		<$E> pause </$E>		
00: 26: 55		−2	−3				Mm.
00: 26: 56		−1	−2		yeah.		
00: 26: 57	MBT_F	0	−1	HN	What do you think?	HG	Yes I want
00: 26: 58	MJS_F	−5	0				yes I want to emphasise these words performative connotation.

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence,
MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

B.2.3.8 Turn-structural episodes and listenership in BJC2_MJS

Table B.2.3-7 Turn-structural episodes and listenership in BJC2_MJS

	BJC2_MJS Number	%	Mean of Leadtime (seconds)	SD	Variance
Episode 1	5	21%	44.20	63.95	4090.16
Episode 2	3	13%	66.33	43.61	1901.56
Episode 3	1	4%	4.00	0.00	0.00
Episode 4	0	0%	--	--	--
Episode 5	13	54%	94.38	70.00	4899.78
Episode 6	1	4%	3.00	0.00	0.00
Episode 7	0	0%	--	--	--
Unclassified	1	4%	--	--	--
Total	24	100%			

Transcript B.2-25 Episode 5 in BJC2_MJS

Timeline	Floor	BJC2_MBT_leadtime	BJC2_MJS_leadtime	BJC2_MBT_gesture	BJC2_MBT_Transcript	BJC2_MJS_gesture	BJC2_MJS_Transcript
00: 29: 52	MBT_F	0	-34		Erm and yeah I mean there're something interesting you thought about imperative of the language <$G?>.+		
00: 29: 53		1	-33				
00: 29: 54		2	-32				
00: 29: 55		3	-31	HG			
00: 29: 56		4	-30	SC/chin			
00: 29: 57		5	-29				
00: 29: 58		6	-28		+I mean deciding imperative is it's it's quite it's interesting I think comes back to nightmare as well but erm +		
00: 29: 59		7	-27	HG			
00: 30: 00		8	-26				
00: 30: 01		9	-25	HG			
00: 30: 02		10	-24				
00: 30: 03		11	-23				
00: 30: 04		12	-22	SC/chin			
00: 30: 05		13	-21		+ it's erm again it's quite <$G?> sensitive <$G?> not so much <$H> affair </$H> <$G?> of ethical response. Imperative language.		
00: 30: 06		14	-20				
00: 30: 07		15	-19				
00: 30: 08		16	-18				
00: 30: 09		17	-17	HG			
00: 30: 10		18	-16				
00: 30: 11		19	-15				
00: 30: 12		20	-14	HG			
00: 30: 13		21	-13				
00: 30: 14		22	-12	HG			
00: 30: 15		23	-11	SC/chin			
00: 30: 16		24	-10				
00: 30: 17		25	-9		Do you do you mean responsibility of the language?+		
00: 30: 18		26	-8	HG			
00: 30: 19		27	-7		+Or do you mean imperative to express? or what did you mean by imperative <$G?>?		
00: 30: 20		28	-6	HG			
00: 30: 21		29	-5				
00: 30: 22		30	-4				
00: 30: 23		31	-3				
00: 30: 24		32	-2				
00: 30: 25		33	-1				
00: 30: 26	MJS_F	-69	0				Erm. I mean by imperative +

Keys: HG= hand gestures, HN= head nods, "="= unfinished sentence, "+"= describe the continuous of the sentence, MJS_T= male Japanese student's floor-taking, MBT_F = male British tutor's floor-taking

B.3 Tables and figures from preferences in turn size and the placement of response tokens

B.3.1 Preferences in placement

Table B.3.1-1 Verbal response tokens in BBC1_MBT's episode 1

BBC1_MBT_episode1	
Yeah.	22
Right.	11
Yeah yeah.	6
Yeah yeah yeah.	3
Uh-huh.	2
+literally erm.	1
<$E> laugh </$E>.	1
Alright. Okay.	1
Mm.	1
Oh god. So+	1
Oh hell. Right+	1
Oh I see right.	1
Oh right oh okay.	1
Oh yeah.	1
Okay.	1
Right yeah yeah yeah.	1
Right yeah yeah.	1
Right. Oh right yeah.	1
Sure yeah yeah.	1
That's right.	1
Yeah that's right yeah.	1
Yeah that's right.	1
Yeah. Okay.	1
Total	62

Table B.3.1-2 Verbal response tokens in BBC1_MBT's episode 2

BBC1_MBT_episode2	
Yeah.	25
Right.	11
Yeah yeah.	10
Okay.	4
Right yeah.	2
+and then yeah yeah+	1
+though isn't it?	1
but I+	1
Er yeah just so you can start getting on with it.	1
Erm yeah+	1
Erm.	1
Mm.	1
Oh god that yeah.	1
Oh really? Oh right.	1
Oh right yeah.	1
Oh yeah.	1
On metaphors.	1
Right yeah yeah.	1
Right. Alright.	1
Right. Okay.	1
Uh-huh.	1
Well yeah.	1
yeah schemas.	1
Yeah that's right.	1
yeah yeah	1
Yeah yeah .Yeah yeah.	1
Yeah yeah yeah.	1
yeah you you you.	1
Yeah. Right.	1
合計	76

Table B.3.1-3 Verbal response tokens in BBC2_MBT's episode 2

BBC2_MBT_episode2	
Sure.	5
Yeah.	5
Mm.	2
<$H> Okay </$H>.	1
Absolutely absolutely.	1
Gosh.	1
I think=	1
Oh yeah.	1
Okay.	1
Okay. Right.	1
Okay?	1
Right.	1
Well I mean+	1
Well that's right. It+	1
Yeah well=	1
Yes.	1
Total	25

Appendix C Top 200 most frequent words lists

C.1 Frequent word lists

C.1.1 A concordance software: TextSTAT

Frequencies of words in each participant's utterance in the four conversations in the main study data have also been analysed. A free corpus analysis software system TextSTAT developed by Mattias Huning is employed for the word frequency analysis. Conversation transcripts formatted as plain text files are imported into the software database. With a few clicks, word lists showing frequency and concordances are easily obtained.

Figure C.1-1 TextSTAT: Frequent word List

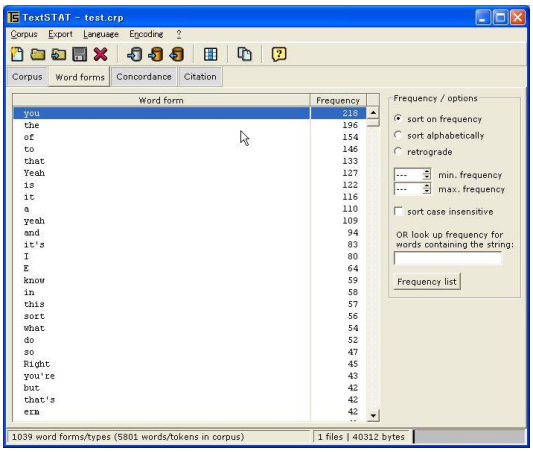
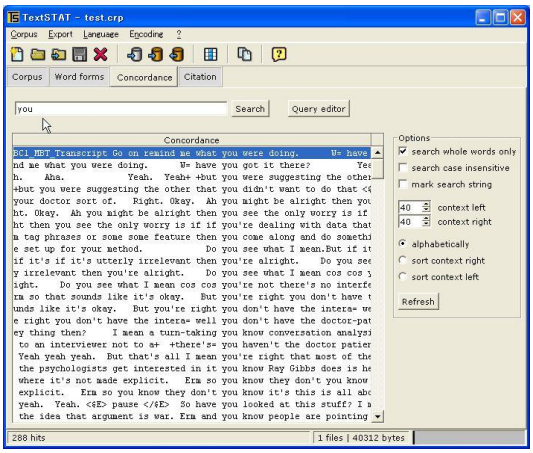


Figure C.1-2 TextSTAT: Concordance



C.1.2 Top 200 most frequent word lists from students' utterances

Table C.1.2-1 Top 1-50 most frequent words in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS		BJC2_MJS	
1	yeah	126	mm	220	mhm	118	the	49
2	i	101	and	95	yeah	40	i	37
3	the	66	the	86	i	16	to	21
4	erm	65	i	82	erm	15	and	18
5	of	63	erm	81	to	14	will	17
6	to	56	yeah	74	ah	11	of	17
7	it	51	a	57	have	10	erm	14
8	and	46	to	47	laugh	10	so	14
9	that	42	that	47	in	9	chapter	12
10	a	42	was	44	the	8	is	10
11	er	37	of	34	mm	8	ah	10
12	well	36	er	33	er	8	in	10
13	in	36	it	31	students	7	think	9
14	it's	32	yes	28	teacher	6	yeah	9
15	so	32	so	28	so	6	position	8
16	but	31	in	24	and	6	it	8
17	is	26	sure	22	is	5	be	8
18	just	24	on	21	maybe	5	want	7
19	think	24	well	20	of	5	or	7
20	quite	24	there	20	class	5	that	7
21	you	23	think	20	some	5	clarify	6
22	on	22	as	19	speak	5	yes	6
23	laugh	21	but	18	english	5	novels	6
24	i'm	21	be	18	japanese	4	theme	5
25	at	19	it's	17	a	4	a	5
26	right	18	some	16	if	3	for	5
27	no	18	which	16	don't	3	my	5
28	for	18	corpus	15	quite	3	three	5
29	do	17	at	15	but	3	very	5
30	okay	17	mhm	15	my	3	anticipation	4
31	metaphor	17	is	14	like	3	language	4
32	what	17	very	14	it's	3	next	4
33	kind	17	quite	14	student	3	write	4
34	was	16	laugh	13	do	3	it's	4
35	have	15	that's	13	how	3	but	4
36	about	15	have	12	no	2	have	4
37	they	15	with	12	then	2	two	4
38	not	14	about	12	between	2	er	4
39	with	14	just	12	question	2	do	4
40	this	14	talk	11	enter	2	authobiograp	4
41	as	13	for	11	want	2	not	4
42	be	13	you	11	interaction	2	writing	3
43	i've	13	i'm	10	large	2	more	3
44	like	13	because	10	think	2	reader	3
45	which	13	i've	10	that	2	refer	3
46	really	12	we	10	they	2	meaning	3
47	cos	12	right	10	this	2	should	3
48	how	12	what	10	okay	2	ethics	3
49	some	12	people	9	observation	2	biographical	3
50	can	12	this	9	well	2	arbitrary	3

Table C.1.2-2 Top 51-100 most frequent words in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS		BJC2_MJS	
51	going	12	relevant	9	teachers	2	emphasise	3
52	are	12	me	8	especially	2	like	3
53	if	11	when	8	specific	2	on	3
54	mean	11	interesting	8	can	2	this	3
55	there	10	there's	8	not	2	writer	3
56	my	10	one	8	something	2	text	3
57	more	9	were	8	opportunity	2	novel	3
58	or	9	good	8	afraid	1	then	3
59	data	9	health	7	y	1	some	3
60	them	9	no	7	school	1	read	3
61	then	8	also	7	actually	1	one	2
62	doing	8	an	7	individual	1	only	2
63	perspective	8	linguistics	7	with	1	mean	2
64	use	8	like	7	hypothesis	1	meeting	2
65	stuff	8	methodology	7	uh-huh	1	important	2
66	don't	8	up	6	vary	1	possible	2
67	there's	7	do	6	university	1	uh-huh	2
68	something	7	data	6	that's	1	if	2
69	would	7	will	6	frequent	1	vary	2
70	because	7	part	6	it	1	thank	2
71	find	7	are	6	opinion	1	about	2
72	particular	7	again	6	main	1	relationship	2
73	get	7	more	5	chance	1	mm	2
74	that's	7	would	5	expect	1	when	2
75	website	7	then	5	quietly	1	aspects	2
76	explaining	6	all	5	front	1	close	2
77	metaphors	6	or	5	situation	1	form	2
78	through	6	much	5	difficult	1	add	2
79	know	6	after	5	somewhere	1	i'm	2
80	where	6	can	5	got	1	much	2
81	look	6	been	5	fre	1	you	2
82	interesting	6	going	5	usually	1	side	2
83	sure	6	analysis	5	most	1	page	2
84	suppose	6	they	5	or	1	imperative	2
85	they're	6	definitely	4	just	1	kind	2
86	word	6	really	4	mention	1	future	2
87	actually	6	has	4	study	1	these	2
88	conceptual	6	still	4	i've	1	different	2
89	study	6	background	4	un	1	lessing	2
90	oh	6	got	4	book	1	certain	1
91	talk	5	don't	4	observe	1	word	1
92	you're	5	met	4	say	1	scientific	1
93	medical	5	suppose	4	out	1	afraid	1
94	used	5	not	4	normally	1	before	1
95	still	5	something	4	here	1	anticipate	1
96	concepts	5	had	4	opportunitie	1	that's	1
97	mm	5	lot	4	for	1	continue	1
98	i'll	5	he's	4	i'm	1	actually	1
99	things	5	how	4	recommend	1	connect	1
100	patient	5	argument	4	two	1	introductory	1

Table C.1.2-3 Top 101-150 most frequent words in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS		BJC2_MJS	
101	from	5	sarah	4	always	1	effect	1
102	an	5	out	4	use	1	really	1
103	saying	5	interested	4	there	1	with	1
104	probably	5	useful	4	remember	1	although	1
105	sort	5	paper	4	right	1	between	1
106	these	5	duplication	4	individually	1	each	1
107	especially	5	language	3	examination	1	history	1
108	been	5	him	3	talkative	1	return	1
109	looking	5	cos	3	confusing	1	readers	1
110	theory	5	know	3	essay	1	opinion	1
111	trying	5	review	3	old	1	reader's	1
112	up	4	vague	3	classes	1	sorry	1
113	health	4	where	3	choose	1	im	1
114	guess	4	take	3	experience	1	least	1
115	vague	4	big	3	activities	1	contents	1
116	interviews	4	emails	3	alright	1	thousand	1
117	off	4	did	3	high	1	linguistic	1
118	take	4	our	3	research	1	also	1
119	down	4	thirty	3	problem	1	by	1
120	pragmatics	4	go	3	point	1	her	1
121	also	4	from	3	kinds	1	has	1
122	got	4	nice	3	understand	1	tenth	1
123	could	4	richard	3	quiet	1	end	1
124	perhaps	4	mean	3	talk	1	extremely	1
125	much	4	few	3	clifton	1	response	1
126	pain	4	talking	3	oh	1	most	1
127	gonna	4	into	3	teenager	1	nature	1
128	see	4	looked	3	says	1	at	1
129	one	4	sessions	3	way	1	finished	1
130	anyway	4	words	3	need	1	rewrite	1
131	patients	4	audience	3			effort	1
132	lot	4	conferences	3			near	1
133	want	4	follow	3			sections	1
134	only	4	knowing	3			conclusion	1
135	out	4	um-hm	3			both	1
136	why	4	seems	3			book	1
137	conditions	4	he	3			example	1
138	embodied	3	back	3			discussing	1
139	language	3	several	2			roles	1
140	should	3	although	2			send	1
141	hard	3	three	2			days	1
142	understandir	3	writing	2			look	1
143	far	3	her	2			importance	1
144	largely	3	enjoyable	2			properly	1
145	hold	3	seemed	2			performative	1
146	linguistic	3	care	2			better	1
147	by	3	say	2			staging	1
148	start	3	presented	2			author	1
149	go	3	chris	2			positions	1
150	idea	3	actual	2			referred	1

Table C.1.2-4 Top 151-200 most frequent words in students

	BBC1_FBS		BBC2_MBS		BJC1_MJS	BJC2_MJS	
151	abstract	3	lots	2		theoretical	1
152	laughs	3	always	2		subjects	1
153	behind	3	count	2		august	1
154	can't	3	certainly	2		underline	1
155	mapping	3	keep	2		appreciate	1
156	social	3	sent	2		non	1
157	he's	3	future	2		social	1
158	reading	3	line	2		were	1
159	experience	3	intuition	2		basically	1
160	using	3	themes	2		from	1
161	ask	3	design	2		first	1
162	research	3	course	2		frame	1
163	read	3	fair	2		articles	1
164	way	3	already	2		they	1
165	people	3	if	2		objects	1
166	process	3	practitioners	2		would	1
167	moment	3	by	2		bring	1
168	we	3	small	2		are	1
169	chronic	3	finished	2		paragraphs	1
170	interested	3	could	2		bit	1
171	better	2	time	2		criticism	1
172	although	2	even	2		further	1
173	proposal	2	attachment	2		he	1
174	me	2	things	2		authobiograp	1
175	andrew	2	remember	2		written	1
176	essays	2	two	2		who	1
177	perceptions	2	approach	2		relation	1
178	terms	2	points	2		level	1
179	cue	2	empiricism	2		go	1
180	here	2	first	2		eight	1
181	concept	2	give	2		big	1
182	always	2	bit	2		words	1
183	count	2	weren't	2		useful	1
184	tagging	2	should	2		possibility	1
185	searches	2	looking	2		says	1
186	online	2	thing	2		discuss	1
187	analogy	2	initial	2		philosophy	1
188	explanations	2	touching	2		i'll	1
189	both	2	area	2		require	1
190	explain	2	f	2		connotation	1
191	thesis	2	w	2		call	1
192	already	2	corpora	2		again	1
193	edited	2	parallel	2		order	1
194	between	2	doing	2		formulation	1
195	thinking	2	shape	2		itself	1
196	difficult	2	his	2		age	1
197	yes	2	those	2		concern	1
198	time	2	done	2		overwhelm	1
199	book	2	stopped	2		recent	1
200	what's	2	home	2		event	1

C.1.3 Top 200 most frequent word lists from tutors' utterances

Table C.1.3-1 Top 1-50 most frequent words in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
1	yeah	228	you	138	the	115	you	232
2	you	215	of	113	is	86	erm	174
3	the	193	and	104	to	75	of	142
4	of	147	the	98	you	73	the	135
5	to	136	in	89	i	71	i	116
6	that	128	to	85	this	69	and	101
7	is	121	that	76	a	63	to	90
8	it	108	i	76	of	51	know	72
9	a	104	erm	54	and	50	yeah	72
10	and	91	a	51	teacher	40	that	71
11	it's	82	know	42	in	39	is	67
12	i	77	it	40	that	37	that's	59
13	so	73	think	40	think	36	just	55
14	erm	71	is	38	students	33	think	55
15	right	62	er	38	it	30	it's	53
16	this	60	your	37	okay	30	mean	51
17	know	58	so	32	are	30	in	51
18	but	56	there	30	not	29	it	49
19	what	55	what	28	with	26	about	47
20	in	52	you're	27	have	24	what	47
21	do	50	but	25	yeah	23	this	45
22	sort	50	be	25	so	23	mm	44
23	that's	45	it's	25	observation	22	do	41
24	or	43	some	25	for	22	so	40
25	you're	42	that's	24	they	22	be	35
26	as	39	or	23	about	20	or	31
27	well	36	yeah	23	erm	20	something	30
28	at	35	about	23	do	19	kind	30
29	er	33	are	23	what	19	a	30
30	about	33	at	22	be	18	sort	28
31	all	31	if	21	it's	18	on	27
32	doing	31	on	21	here	17	which	27
33	are	31	as	20	because	17	as	26
34	on	30	well	20	some	17	your	25
35	if	29	terms	19	but	16	how	24
36	there's	29	how	19	that's	16	we	23
37	mean	29	with	18	teaching	16	again	23
38	stuff	29	was	17	bit	15	but	21
39	not	29	right	16	individual	15	idea	21
40	for	27	for	16	don't	14	very	20
41	an	26	this	16	at	14	not	17
42	just	26	up	15	from	14	say	16
43	oh	26	corpus	15	your	14	really	15
44	i'm	24	need	15	oh	14	here	15
45	have	24	very	15	class	13	you've	15
46	your	24	because	15	need	13	because	15
47	then	23	have	14	all	13	ethics	15
48	you've	22	like	14	one	13	back	15
49	look	22	literature	13	good	13	got	14
50	okay	22	going	13	or	13	see	14

Table C.1.3-2 Top 51-100 most frequent words in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
51	cos	21	sure	12	if	12	one	14
52	be	21	relation	12	got	12	where	14
53	with	21	good	12	can	12	might	14
54	get	21	health	11	should	11	you're	13
55	like	20	go	11	teachers	11	have	13
56	can	20	sort	11	there	11	there	13
57	thing	20	kind	11	more	10	like	13
58	think	20	really	10	say	10	will	13
59	metaphor	19	you've	10	then	10	right	12
60	got	17	do	10	i'm	10	if	12
61	going	17	mean	10	mean	10	two	12
62	out	17	linguistics	10	we	10	there's	12
63	need	15	okay	10	kind	10	they	12
64	data	15	work	10	really	9	should	11
65	study	15	mental	10	me	9	could	11
66	way	15	approach	9	laugh	9	imperative	11
67	there	15	would	9	very	9	well	11
68	how	15	own	9	question	9	chapter	11
69	really	14	where	9	research	9	into	11
70	say	14	healthcare	9	others	8	some	11
71	from	14	language	8	large	8	can	11
72	why	14	other	8	interaction	8	quite	11
73	other	13	laugh	8	classes	8	er	11
74	me	13	an	8	japanese	8	up	10
75	don't	13	can	8	how	8	there're	10
76	something	13	just	8	interact	8	fine	10
77	gonna	13	out	8	actually	8	staging	10
78	where	13	when	7	out	8	i'm	10
79	here's	12	review	7	right	7	can't	10
80	no	12	debate	7	as	7	these	10
81	go	12	did	7	on	7	way	10
82	one	12	still	7	impossible	7	terms	9
83	want	12	not	7	way	7	did	9
84	dissertation	12	he's	7	will	7	great	9
85	up	11	were	7	which	7	at	9
86	different	10	will	7	may	7	mhm	9
87	even	10	conference	7	just	7	novels	9
88	was	10	research	7	their	7	with	9
89	they're	10	has	6	why	7	are	9
90	isn't	10	s	6	problems	6	seems	9
91	down	10	we've	6	help	6	absolutely	8
92	looking	10	y	6	quietly	6	need	8
93	whether	9	thesis	6	make	6	then	8
94	when	9	also	6	you're	6	also	8
95	them	9	things	6	come	6	things	8
96	into	9	want	6	frequency	6	much	8
97	see	9	these	6	mm	6	come	8
98	would	9	which	6	therefore	6	process	8
99	words	9	process	6	own	6	useful	8
100	people	9	stuff	6	like	6	me	7

Table C.1.3-3 Top 101-150 most frequent words in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
101	good	9	whole	6	read	6	through	7
102	metaphors	8	obviously	6	another	6	text	7
103	set	8	again	6	my	6	position	7
104	most	8	me	5	better	5	time	7
105	time	8	make	5	when	5	go	7
106	things	8	little	5	questions	5	was	7
107	two	8	got	5	where	5	theories	7
108	give	8	mm	5	you've	5	sense	7
109	though	8	don't	5	purpose	5	get	7
110	conceptual	8	yes	5	speak	5	going	7
111	might	8	probably	5	between	5	actually	7
112	trying	8	perhaps	5	two	5	good	7
113	always	7	from	5	much	5	yes	6
114	still	7	bit	5	something	5	whether	6
115	y	7	being	5	study	5	more	6
116	analysis	7	one	5	working	5	language	6
117	much	7	had	5	lot	5	other	6
118	talking	7	wouldn't	5	quite	5	make	6
119	those	7	study	5	ah	5	three	6
120	metonymy	7	his	5	almost	4	different	6
121	were	7	been	5	writing	4	memory	6
122	which	7	people	5	front	4	all	6
123	only	7	get	5	range	4	don't	6
124	actually	7	actually	5	answer	4	authobiograpl	6
125	by	7	now	5	harmer	4	from	6
126	now	7	interested	5	prefer	4	many	6
127	they	7	they	5	there're	4	end	6
128	big	6	words	4	little	4	find	6
129	ma	6	cos	4	grammar	4	interesting	6
130	we're	6	case	4	speaking	4	read	6
131	chop	6	ways	4	professional	4	thought	6
132	than	6	indicate	4	points	4	gibson	6
133	thousand	6	making	4	would	4	almost	5
134	else	6	then	4	doing	4	talk	5
135	stylistic	6	look	4	maybe	4	written	5
136	rather	6	all	4	opportunity	4	anticipation	5
137	phd	6	practitioners	4	level	4	form	5
138	done	6	sexuality	4	use	4	bit	5
139	framework	6	i'm	4	problem	4	relation	5
140	th	6	adolescence	4	arrangement	4	looking	5
141	question	6	there's	4	english	4	isn't	5
142	cognitive	6	points	4	going	4	those	5
143	worth	6	give	4	work	4	wanna	5
144	interested	6	another	4	sorry	4	said	5
145	again	6	into	4	haven't	4	events	5
146	talk	5	doing	4	might	4	metafiction	5
147	more	5	done	4	stated	4	out	5
148	domain	5	srikant	4	seating	4	he	5
149	pretty	5	we	4	opportunities	4	words	5
150	type	5	doesn't	4	now	4	put	4

Table C.1.3-4 Top 151-200 most frequent words in tutors

	BBC1_MBT		BBC2_MBT		BJC1_FBT		BJC2_MBT	
151	keep	5	analysis	4	back	4	debate	4
152	doctor	5	peter	4	talk	3	performative	4
153	wouldn't	5	side	4	tasks	3	precisely	4
154	mm	5	something	4	sitting	3	issue	4
155	could	5	back	4	blind	3	being	4
156	start	5	absolutely	3	link	3	concepts	4
157	what's	5	issue	3	brown	3	readers	4
158	interesting	5	more	3	different	3	next	4
159	very	5	everything	3	school	3	by	4
160	mapping	5	relevance	3	no	3	sections	4
161	come	5	always	3	around	3	even	4
162	ideas	5	big	3	fine	3	relate	4
163	ones	5	around	3	than	3	demands	4
164	alright	5	both	3	also	3	page	4
165	some	5	its	3	difficult	3	point	4
166	read	5	no	3	yes	3	topic	4
167	lot	5	within	3	following	3	forward	4
168	style	5	linguistic	3	start	3	anything	4
169	either	5	by	3	time	3	thousand	4
170	source	5	stockwell	3	book	3	i've	4
171	does	5	time	3	situation	3	doing	4
172	god	5	too	3	indicates	3	done	4
173	literary	5	discourse	3	go	3	four	4
174	idea	5	driven	3	silence	3	who	4
175	back	5	suppose	3	helping	3	that'll	4
176	war	4	people've	3	photocopy	3	couple	4
177	off	4	willing	3	well	3	lessing	4
178	try	4	looking	3	often	3	thing	4
179	having	4	see	3	see	3	complicated	4
180	analogies	4	aren't	3	brief	3	worth	4
181	journal	4	area	3	procedure	3	bring	4
182	steen	4	isn't	3	frequently	3	reader	4
183	weeks	4	many	3	talking	3	conceptual	4
184	using	4	needs	3	into	3	now	4
185	thinking	4	way	3	survey	3	whole	4
186	essentially	4	sense	3	i've	3	exploring	4
187	called	4	context	3	however	3	authobiograpl	3
188	first	4	c	3	sense	3	responsibility	3
189	realisations	4	over	3	want	3	theme	3
190	point	4	presenting	3	reason	3	mark	3
191	order	4	key	3	who	3	help	3
192	next	4	set	3	alright	3	what's	3
193	w	4	salience	3	these	3	nature	3
194	because	4	v	3	am	3	jennet	3
195	i've	4	part	3	er	3	when	3
196	linguistics	4	wasn't	3	spot	3	us	3
197	important	4	may	3	student	3	certainly	3
198	can't	4	particularly	3	does	3	talks	3
199	these	4	seem	3	idea	3	both	3
200	doctors	4	louise	3	vary	2	themes	3

Appendix D Consent form

Keiko Tsuchiya
Mphil in English Studies
University of Nottingham

Dear Participant

Letter of Information

This letter is an invitation to consider participating in a study I am conducting as part of my PhD degree in English Studies at University of Nottingham under the supervision of Dr. Svenja Adolphs. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

This study investigates the use of backchannels and gestures in conversation in English language. The aim of this research is to find out the patterns in linguistics choices of backchannels in conversation and to offer implications for English teaching and learning in Japan based on the research. In order to capture the features of spoken discourse in pedagogic context, supervisions in university are vide-recorded and transcribed. The data is used for academic research purposes only.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me at ***(07875798354)*** or by email at ***(aexkt@nottingham.ac.uk)***.

Yours Sincerely,

Letter of Consent

I agree to take part in a research study being conducted by Keiko Tsuchiya, Mphil in English Studies at The University of Nottingham.

I have read and understand the information of the study above. All the procedures, any risks and benefits have been explained to me. I have had the opportunity to ask any questions and to receive any additional details I wanted about the study. If I have questions later about the study, I can ask the researcher ***(Keiko Tsuchiya, Mphil in English Studies at The University of Nottingham, 07875798354).***

KATIE HEAD
Name of Participant

Katie Head
Signature of Participant

21 Feb 2007
Date

Keiko Tsuchiya
Mphil in English Studies
University of Nottingham

Dear Participant

Letter of Information

This letter is an invitation to consider participating in a study I am conducting as part of my PhD degree in English Studies at University of Nottingham under the supervision of Dr. Svenja Adolphs. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

This study investigates the use of backchannels and gestures in conversation in English language. The aim of this research is to find out the patterns in linguistics choices of backchannels in conversation and to offer implications for English teaching and learning in Japan based on the research. In order to capture the features of spoken discourse in pedagogic context, supervisions in university are vide-recorded and transcribed. The data is used for academic research purposes only.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me at ***(07875798354)*** or by email at ***(aexkt@nottingham.ac.uk)***.

Yours Sincerely,

Letter of Consent

I agree to take part in a research study being conducted by Keiko Tsuchiya, Mphil in English Studies at The University of Nottingham.

I have read and understand the information of the study above. All the procedures, any risks and benefits have been explained to me. I have had the opportunity to ask any questions and to receive any additional details I wanted about the study. If I have questions later about the study, I can ask the researcher ***(Keiko Tsuchiya, Mphil in English Studies at The University of Nottingham, 07875798354).***

Takuo Tomizawa
Name of Participant

Signature of Participant

11/04/07
Date

Keiko Tsuchiya
PhD in English Studies
University of Nottingham

Dear Participant

Letter of Information

This letter is an invitation to consider participating in a study I am conducting as part of my PhD degree in English Studies at University of Nottingham under the supervision of Dr. Svenja Adolphs. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

This study investigates the use of backchannels and gestures in conversation in English language. The aim of this research is to find out the patterns in linguistics choices of backchannels in conversation and to offer implications for English teaching and learning in Japan based on the research. In order to capture the features of spoken discourse in pedagogic context, supervisions in university are vide-recorded and transcribed. The data is used for academic research purposes only.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me by email at ***(aexkt@nottingham.ac.uk)***.

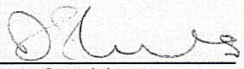
Yours Sincerely,

Letter of Consent

I agree to take part in a research study being conducted by Keiko Tsuchiya, PhD in English Studies at The University of Nottingham.

I have read and understand the information of the study above. All the procedures, any risks and benefits have been explained to me. I have had the opportunity to ask any questions and to receive any additional details I wanted about the study. If I have questions later about the study, I can ask the researcher ***(Keiko Tsuchiya, PhD in English Studies at The University of Nottingham, aexkt@nottingham.ac.uk)***.

Dr. David James
Name of Participant


Signature of Participant

27.07.2007
Date

Keiko Tsuchiya
PhD in English Studies
University of Nottingham

Dear Participant

Letter of Information

This letter is an invitation to consider participating in a study I am conducting as part of my PhD degree in English Studies at University of Nottingham under the supervision of Dr. Svenja Adolphs. I would like to provide you with more information about this project and what your involvement would entail if you decide to take part.

This study investigates the use of backchannels and gestures in conversation in English language. The aim of this research is to find out the patterns in linguistics choices of backchannels in conversation and to offer implications for English teaching and learning in Japan based on the research. In order to capture the features of spoken discourse in pedagogic context, supervisions in university are vide-recorded and transcribed. The data is used for academic research purposes only.

If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please contact me by email at ***(aexkt@nottingham.ac.uk)***.

Yours Sincerely,

Letter of Consent

I agree to take part in a research study being conducted by Keiko Tsuchiya, PhD in English Studies at The University of Nottingham.

I have read and understand the information of the study above. All the procedures, any risks and benefits have been explained to me. I have had the opportunity to ask any questions and to receive any additional details I wanted about the study. If I have questions later about the study, I can ask the researcher ***(Keiko Tsuchiya, PhD in English Studies at The University of Nottingham, aexkt@nottingham.ac.uk)***.

Yasuhiro Kondo
Name of Participant

Signature of Participant

27 [redacted] 2007
Date July

Appendix E Participants' information

E.1 Participants' information

E.1.1 Pilot study

Table E.1.1-1 Participants in the pilot study

	Participants				
				Age	Notes
British-British Conversation (C1)	C1_MBT	Male	British	40s	Professor in School of English Studies
	C1_FBS	Female	British	Mid 20s	MA student
British-Japanese Conversation (C2)	C2_FBT	Female	British	50s	Course leader of MA ELT
	C2_MJS	Male	Japanese	Mid 20s	MA student

E.1.2 Main study

Table E.1.2-1 Participants in the main study

	Participants				
				Age	Notes
British-British Conversation (BBC1)	BBC1_MBT	Male	British	40s	Professor in School of English Studies
	BBC1_FBS	Female	British	Mid 20s	MA student
British-British Conversation (BBC2)	BBC2_MBT	Male	British	50s	Professor in School of Nursing
	BBC2_MJS	Male	British	30s	PhD student, Part-time lecturer
British-Japanese Conversation 1(BJC1)	BJC1_FBT	Female	British	50s	Course leader of MA ELT
	BJC1_MJS	Male	Japanese	Mid 20s	MA student
British-Japanese Conversation 2 (BJC2)	BJC2_MBT	Male	British	30s	Lecturer in School of English Studies
	BJC2_MJS	Male	Japanese	Mid 20s	MA student

Notes: BBC1_MBT = C1_MBT in the pilot study, BBC1_FBS= C1_FBS in the pilot study, BJC1_FBT = C2_FBT in the pilot study, BJC1_MJS= C2_MJS in the pilot study

References

Adolphs, S. (2001). Linking Lexico-grammar and Speech Acts: a Corpus-based Approach. *School of English Studies.* Nottingham: The University of Nottingham.

Adolphs, S. (2006). *Introducing Electronic Text Analysis: A Practical Guide for Language and Literary Studies.* London: Routledge.

Adolphs, S. (2008). *Corpus and Context: Investigating Pragmatic Functions in Spoken Discourse.* Amsterdam: John Benjamins Publishing Company.

Adolphs, S., Brown, B., Carter, R., Crawford, P. & Sahota, O. (2004). Applying corpus linguistics in a health care context. *Journal of Applied Linguistics* **1:** 9-28.

Ager, M. (1994a). The intercultural frame. *International Journal of Intercultural Relations* **18:** 221-237.

Ager, M., H. (1994b). *Language Shock: Understanding the Culture of Conversation.* New York: Harper Collins Publishers.

Alred, G., Byram, M. & Fleming, M. (2003). *Intercultural Experience and Education.* Bristol: Multilingual Matters.

American Council on the Teaching of Foreign Languages (1999). *National Standards in Foreign Language Education.* http://www.actfl.org/i4a/pages/index.cfm?pageid=3392. New York: American Council on the Teaching of Foreign Languages, Inc. [Accessed: 9th May 2006].

Anderson, G. (1998). *Fundamentals of Educational Research.* London: The Falmer Press.

Antaki, C. (2002). "Lovely": Turn-initial high-grade assessments in telephone closings. *Discourse Studies* **4:** 5-24.

Armstrong-Warwick, S., Thompson, H., S., McKelvie, D. & Petitpierre, D. (1994). Data in Your Language: The ECI Multilingual Corpus 1. *HCRC Technical Report.* Edinburgh/Geneva: HCRC/Institute Dalle Molle.

Aston, G. & Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA.* Edinburgh: Edinburgh University Press.

Atkinson, D. (1999). TESOL and culture. *TESOL Quarterly* **33:** 625-654.

Austin, J. L. (1962). *How To Do Things With Words.* London: The Clarendon Press.

Baker, C., Emmison, M. & Firth, A. (2001). Discovering Order in Opening Sequences: Calls to Software Helpline. In McHoul, A. & M. Rapley (eds.) *How to Analyse Talk in Institutional Settings: A Casebook of Methods.* London: Continuum.

Bakhtin, M. M. (1956). *Speech Genre and Other Late Essays.* Austin: University of Texas Press.

Bhatia, V., K. (1993). *Analysing Genre: Language Use in Professional Settings.* Harlow: Longman.

Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers.* Amsterdam: John Benjamins Publishing Company.

Biber, D., Connor, U. & Upton, T., A (2007). *Discourse on the move: Using corpus analysis to describe discourse structure.* Amsterdam: John Benjamins Publishing Company.

Blommaert, J. (2005). *Discourse.* Cambridge: Cambridge University Press.
Brown, G. & Yule, G. (1983). Discourse Analysis. Cambridge: Cambridge University Press.
Brown, J. D. & Rodgers, T. S. (2002). *Doing Second Language Research.* Oxford: Oxford University Press.
Brown, P. & Levinson, S., C (1987). *Politeness: Some Universals in Language Usage.* Cambridge: Cambridge University Press.
Bryman, A. (1988). *Quantity and Quality in Social Research.* London Routledge.
Burton, D. (1981). Analysing Spoken Discourse. . In Coulthard, M. & M. Montgomery (eds.) *Studies in Discourse Analysis.* London: Routledge. 61-81.
Carroll, D. (2004). Restarts to Novice Turn Beginnings: Disfluencies or Interactional Achievements? In Gardner, R. & J. Wagner (eds.) *Second Language Conversations.* London: Continuum. 201-220.
Carroll, J. B. (1956). *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf.* Massachusetts: The MIT Press.
Carter, R. (2004). *Language and Creativity: The Art of Common Talk.* London: Routledge.
Carter, R. (2007). Response to Special Issue of Applied Linguistics devoted to Language Creativity in Everyday Contexts. *Applied Linguistics* **28:** 597-608.
Carter, R. & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics* **16:** 141-158.
Carter, R. & McCarthy, M. (1997). *Exploring Spoken English.* Cambridge: Cambridge University Press.
Carter, R. & McCarthy, M. (2006). *Cambridge Grammar of English: A Comprehensive Guide Spoken and Written English Grammar and Usage.* Cambridge: Cambridge University Press.
Chafe, W., L., Du Bois, J., W. & Thompson, S., A. (1991). Towards a New Corpus of Spoken American English. In Ajimer, K. & B. Altenberg (eds.) *English Corpus Linguistics: Studies in Honour of Jan Svartvik.* Harlow: Longman. 64-82.
Chiasson, C. A. & Hayes, L. J. (1993). The effects of subtle differences between listeners and speakers on the referential speech of college freshmen. *Psychological Record* **43:** 13-25.
Chomsky, N. (1971). *Syntactic Structures.* Hague: Mouton.
CILT The National Centre for Languages (2004). *The INCA (Intercultural Competence Assessment) Project.* http://www.incaproject.org/index.htm. London: INCA Project, CILT (The Chartered Institute of Logistics and Transports). [Accessed: 30th May 2006].
Clyne, M. (2003). *Dynamics of Language Contact.* Cambridge: Cambridge University Press.
Cohen, L., Manion, L. & Morrison, K. (2000). *Research Methods in Education.* London: Routledge Falmer.
Cook, G. (1994). *Discourse and Literature.* Oxford: Oxford University Press.
Coulthard, M. (1977). *An Introduction to Discourse Analysis.* Harlow: Longman.
Coupland, N. (1983). Patterns of encounter management: Further arguments for discourse variables. *Language in Society* **12:** 459-476.
Coupland, N. (2007). *Style: Language Variation and Identity*. Cambridge: Cambridge

University Press.
Couture, B. (1986). Effective Ideation in Written Text: A Functional Approach to Clarity and Exigence. In Couture, B. (ed.) *Functional Approaches to Writing Research Perspectives.* London: Frances Pinter. 69-92.
Davies, J., W & Vaks, S. (2001). A Perceptual User Interface for Recognizing Head Gesture Acknowledgements. *the 2001 ACM Workshop on Perceptive User Interfaces.* Orland.
Davies, M. (2004). *BYU-BNC: The British National Corpus.* http://corpus.byu.edu/bnc/. Birmingham: Birmingham Young University. [Accessed: 20th November 2006].
Davison, R. (1993). Building a million-word computer science corpus in Hong Kong. *ICAME Journal* **16:** 123-124.
Day, D. (1998). Being Ascribed, and Resisting Membership of an Ethnic Group. In Antaki, C. & S. Widdicombe (eds.) *Identities in Talk.* London: Sage publications. 151-170.
de Saussure, F. (1972[1983]). Course in General Linguistics. London: Gerald Duckworth.
Dodd, C. H. (1998). *Dynamics of Intercultural Communication.* Boston: McGraw-Hill Companies.
Drew, P. (1984). Speakers' Reportings in Invitation Sequences. In Atkinson, M., J. & J. Heritage (eds.) *Structures of Social Actions: Studies in Conversation Analysis.* Cambridge: Cambridge University Press. 129-147.
Drew, P. & Heritage, J. (1992). Analyzing Talk at Work: An Introduction. In Drew, P. & J. Heritage (eds.) *Talk at Work: Interaction in Institutional Settings.* Cambridge: Cambridge University Press. 3-65.
Drummond, K. & Hopper, R. (1993). Back channels revisited: Acknowledgment tokens and speakership incipiency. *Research on Language and Social Interaction* **26:** 157-177.
Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* **23:** 283-292.
Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society* **3:** 161-180.
Duncan, S. & Niederehe, G. (1974). On signalling that it's your turn to speak. *Journal of Experimental Social Psychology* **10:** 234-247.
Edmondson, W. (1981). *Spoken Discourse: a Model for Analysis.* London: Longman.
Eggins, S. & Slade, D. (1997). *Analysing Casual Conversation.* London: Equinox Publishing Ltd.
Erickson, F. & Shults, J. (1982). *The Counselor as Gatekeeper: Social Interaction in Interviews.* New York: Academic Press.
Evison, J. & McCarthy, M. (2010). How Do Turn-openers Relate to Turn Closers? *IVACS annual symposium.* Leeds Metropolitan University.
Faerch, C. & Kasper, G. (1987). Perspectives on language transfer. *Applied Linguistics* **8:** 111-136.
Fairclough, N. (2003). *Analysing Discourse: Textual Analysis for Social Research.* London: Routledge.
Fellegy, A. (1995). Patterns and functions of minimal response. *American Speech* **70:** 186-199.

Firth, J. R. (1934). *Papers In Linguistics.* London: Oxford University Press.
Francis, G. & Hunston, S. (1992). Analysing Everyday Conversation. In Coulthard, M. (ed.) *Advances in Spoken Discourse Analysis.* London: Routledge. 123-161.
French, A., Greenhalgh, C., Crabtree, A., Wright, M., Brundell, P., Hampshire, A. & Rodden, T. (2006). Software Replay Tools for Time-based Social Science Data. *2nd International Conference on e-Social Science.* Manchester.
Fujimura-Wilson, K. (2007). Japanese exact repetitions involving talk among friends. *Discourse Studies* **9:** 319-339.
Fung, L. & Carter, R. (2007). Discourse Markers and Spoken English: Native and Learner Use in Pedagogic Settings. *Applied Linguistics* **28:** 410-439.
Furo, H. (2001). *Turn-taking in English and Japanese Projectivety in Grammar, Intonation and Semantics.* New York: Routledge.
Gardner, R. (1998). Between speaking and listening, the vocalization of understandings. *Applied Linguistics* **19:** 204-244.
Gardner, R. (2002). *When Listeners Talk: Response Tokens and Listener Stance.* Amsterdam: John Benjamins.
Garside, R. & Leech, G. (1982). Grammatical Tagging of the LOB Corpus: General Survey. In Johansson, S. (ed.) *Computer Corpora in English Language Research.* Bergen: Norwegian Computing Centre for the Humanities.
Goffman, E. (1955). On face work: An analysis of ritual elements in social interaction. *Psychiatry* **18:** 213-231.
Goffman, E. (1959). *The Presentation of Self in Everyday Life.* Gardencity, NY: Doubleday Anchor Books.
Goffman, E. (1972). *Interaction Ritual: Essays on Face-to-Face Behaviour.* Hamondsworth: Penguin.
Goldin-Meadow, S. (1999). The role of gesture in communication and thinking *Trends in Cognitive Sciences* **3:** 419-429.
Graddol, D. (2006). *English Next.* London: British Council.
Granger, S. (1993). The International Corpus of Learner English. In Aarts, J., P. de Haan & N. Oostdijk (eds.) *English Language Corpora: Design, Analysis and Exploitation.* Amsterdam: Rodopi. 57-69.
Greenbaum, S. (1991). The Development of the International Corpus of English. In Ajimer, K. & B. Altenberg (eds.) *English Corpus Linguistics: Studies in Honour of Jan Svartvik.* Harlow: Longman. 83-94.
Greenbaum, S. & Svartvik, J. (1990). The London-Lund Corpus of Spoken English. In Svartvik, J. (ed.) *The London-Lund Corpus of Spoken English: Description and Research.* Lund: Lund University Press. 11-51.
Grice, H. P. (1975). Logic and Conversation. In Cole, P. & J. L. Morgan (eds.) *Syntax and Semantics volume 3: Speech Acts.* New York: Academic Press. 41-58.
Grice, H. P. (1989). *Studies in the Way of Words.* Cambridge, MT: Harvard University Press.
Gumperz, J. J. (1978). The Conversational Analysis of Interethnic Communication. In E, L. R. (ed.) *Interethnic Communication: Southern Anthropological Society Proceedings, No. 12.* Athens, GA: University of Georgia Press.
Gumperz, J. J. (1992). Contextualization and Understanding. In Duranti, A. & C. Goodwin (eds.) *Rethinking Context: Language as an Interactive Phenomenon.* Cambridge: Cambridge University Press.

Halliday, M. A. K. (1978). *Language as Social Semiotic.* London: Arnold.
Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English.* Harlow: Longman.
Halliday, M. A. K. & Hasan, R. (1985). *Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective.* Victoria: Deakin University Press.
Handford, J. A. M. (2007). The Genre of The Business Meeting: A Corpus Based Study. *School of English Studies.* Nottingham: the University of Nottingham.
Haugen, E. (1972). *The Ecology of Language: Essays by Einar Haugen.* Stanford: Stanford University Press.
Hayashi, M., Mori, J. & Takagi, T. (2002). Contingent Achievement of Co-Tellership in a Japanese Conversation: An Analysis of Talk, Gaze and Gesture. In Ford, C., B. Fox & S. Thompson (eds.) *The Language of Turn and Sequence.* Oxford: Oxford University Press. 81-122.
Heath, C. (1997). Analysing Face-to-face Interaction: Video, the Visual and Material. In Silverman, D. (ed.) *Qualitative Research: Theory, Method and Practice.* London: Sage Publications. 267-282.
Heritage, J. (1984a). A Change-of-state Token and Aspects of Its Sequential Placement. In Atkinson, M., J. & J. Heritage (eds.) *Structures of Social Actions: Studies in Conversation Analysis.* Cambridge: Cambridge University Press. 299-345.
Heritage, J. (1984b). *Garfinkel and Ethnomethodology.* Cambridge: Polity Press.
Heritage, J. (1997). Conversation Analysis and Institutional Talk. In Silverman, D. (ed.) *Qualitative Research: Theory, Method and Practice.* London: Sage publications. 222-265.
Hopkins, D. (2002). *A Teacher's Guide to Classroom Research.* Maidenhead: Open University Press.
Hosoda, Y. (2006). Repair and relevance of differential language expertise in second language conversations. *Applied Linguistics* **27:** 25-50.
Hymes, D. (1972). Models of the Interaction of Language and Social Life. In Gumperz, J., J. & D. Hymes (eds.) *Directions in Sociolinguistics: The Ethnography of Communication.* New York: Holt, Rinehart and Wiston, Inc. 35-71.
Hymes, D. (1974 [1989]). Ways of Speaking. . In Bauman, R. & J. Sherzer (eds.) *Explorations of in the Ethnography of Speaking.* Cambridge: Cambridge University Press. 433-451.
Ide, N. & Macleod, C. (2001). The American National Corpus: A Standardized Resource of American English. *Proceedings of Corpus Linguistics 2001.* Lancaster, UK.
Iwasaki, S. (1997). The Northridge earthquake conversations: The floor structure and the 'loop' sequence in Japanese conversation. *Journal of Pragmatics* **28:** 661-693.
Jandt, F. E. (1995). *Intercultural Communication.* Thousand Oaks: Sage Publications Inc.
Johansson, S. & Jahr, M.-C. (1982). Grammatical Tagging of the LOB Corpus: Predicting Word Class from Word Ending. In Johansson, S. (ed.) *Computer Corpora in English Language Research.* Bergen: Norwegian computing centre for the Humanities.
Johnstone, B. (2002). *Discourse Analysis.* Oxford: Blackwell.

Kasper, G. (1993). *Interlanguage Pragmatics.* Oxford: Oxford University Press.
Kendon, A. (1972). Some Relationships Between Body Motion and Speech: An Analysis of an Example. In Goldstein, A., P. (ed.) *Studies in Dyadic Communication.* New York: Pergamon Press. 177-210.
Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.
Kita, S. (1996). Aizuchi to unazuki kara mita nihonnjin no taimen kominikeishion (Japanese face to face communication in relation to response tokens and headnods). *nihongogaku(Japanese studies)* **15:** 58-66.
Kita, S. & Ide, S. (2007). Nodding, aizuchi, and final particles in Japanese conversation: How conversation reflects the ideology of communication and social relationships. *Journal of Pragmatics* **39:** 1242-1254.
Knight, D., Bayoumi, S., Mills, A., Crabtree, A., Adolphs, S., Pridmore, T. & Carter, R. (2006). Beyond the text: Construction and Analysis of Multi-Modal Linguistic Corpora. *2nd International Conference on e-Social Science.* Manchester.
Knight, D., Evans, D., Carter, R. & Adolphs, S. (2009). HeadTalk, HeadTalk and the corpus: towards a framework for multi-modal, multi-media corpus development. *Corpora* **4:** 1-32.
Koester, A. (2006). *Investigating Workplace Discourse.* London: Routledge.
Kogure, M. (2005). Cominikeishon tekiouriron to nihongo no aizuchi to unazhuki no hindo no danjosa/ba no kankei ni tsuiteno shohoteki kenkyu (A preliminary study on relationship between communication accommodation theory and differences in frequency of aizuchi and head nods in Japanese conversations between female and male participants). *The 16th conference of the Japanese association of sociolinguistics sciences.* Ryukoku University, Kyoto. 78-81.
Kogure, M. (2007). Nodding and smiling in silence during the loop sequence of backchannels in Japanese conversation. *Journal of Pragmatics* **39:** 1275-1289.
Kramsch, C. (1998). *Language and Culture.* Oxford: Oxford University Press.
Kramsch, C. (2008). Ecological perspectives on foreign language education. *Language Teaching* **41:** 389-408.
Kramsch, C. & Whiteside, A. (2008). Language ecology in multilingual settings: Towards a theory of symbolic competence. *Applied Linguistics* **29:** 645-671.
Kubota, R. (2006). Critical approaches to culture in English Language Teaching. *IATEFL annual conference.* Harrogate.
Kucera, H. & Francis, N., W (1967). *Computational Analysis: Present-day American English.* Providence: Brown University Press.
Kyto, M. & Rissanen, M. (1988). The Helsinki Corpus of English Texts: Classifying and Coding the Diachronic Part. In Kyto, M., O. Ihalainen & M. Rissanen (eds.) *Corpus Linguistics, Hard and Soft.* Amsterdam: Rodopi. 169-180.
Larsen-Freeman, D. (1997). Chaos/complexity science and second language. *Applied Linguistics* **18:** 141-165.
Leech, N. G. (1983). *Principles of Pragmatics.* London: Longman.
Lerner, G. (2002). Turn-sharing: the Choral Co-Production of Talk-in-Interaction. In Ford, C., E., B. Fox, A. & S. Thompson, A. (eds.) *The Language of Turn and Sequence.* Oxford: Oxford University Press. 225-256.
Levinson, S. C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Levinson, S. C. (1992). Activity Types and Language. In Drew, P. & J. Heritage (eds.) *Talk at Work: Interaction in Institutional Settings.* Cambridge: Cambridge University Press.

Limberg, H. (2007). Discourse structure of academic talk in university. *Discourse Studies* **9:** 176-193.

LoCastro, V. (1987). Aizuchi: A Japanese Conversation Routine. In Smith, L., E. (ed.) *Discourse Across Culture: Strategies in World Englishes.* Hemel Hempstead: Prentice Hall International Ltd. 101-113.

Malinowski, B. (1923). The Problem of Meaning in Primitive Languages. In Odgen, C. K. & I. A. Richards (eds.) *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of The Science of Symposium.* London: Routledge & Kegan Paul Ltd. 296-336.

Martin, J. R. (1997). Analysing Genre: Functional Parameters. In Christie, F. & J. R. Martin (eds.) *Genre and Institutions: Social processes in the Workplace and School.* London: Cassell. 3-40.

Maynard, S. (1989). *Japanese Conversation: Self-contextualization through Structure and Interactional Management.* Norwood, NJ: Ablex Publishing.

Maynard, S. (1990). Conversation management in contrast: listener response tokens in Japanese and American English. *Journal of Pragmatics* **14:** 397-412.

Maynard, S. (1993). *Discourse Modality: Subjectivity, Emotion and Voice in the Japanese Language.* Amsterdam: John Benjamins Publishing Company.

Maynard, S. (1997a). Analyzing interactional management in native/non-native English conversation: A case of listener response *IRAL* **35:** 36-60.

Maynard, S. (1997b). *Japanese Communication: Language and Thought in Context.* Honolulu: University of Hawaii Press.

Maynard, S. (2007). *Linguistic Creativity in Japanese Discourse.* Amsterdam: John Benjamins Publishing Company.

McCarthy, M. (1998). *Spoken Language & Applied Linguistics.* Cambridge: Cambridge University Press.

McCarthy, M. (2000). Mutually Captuve Audiences: Small Talk and the Genre of Close-contact Service Encounters. In Coupland, J. (ed.) *Small Talk.* Harlow: Longman. 84-109.

McCarthy, M. (2002). Good Listenership Made Plain: Non-minimal Response Tokens in British and American Spoken English. In Reppen, R., S. Fitzmaurice & D. Biber (eds.) *Using Corpora to Explore Linguistic Variation.* Amsterdam: John Benjamins. 49-72.

McCarthy, M. & Carter, R. (1994). *Language As Discourse: Perspectives for Language Teaching.* Harlow: Longman.

McCarthy, M., Matthiessen, C. & Slade, D. (2002). Discourse Analysis. In Schmitt, N. (ed.) *An Introduction to Applied Linguistics.* London: Arnold. 55-73.

McCarthy, M. & O'Keeffe, A. (2004). Research in the teaching of speaking. *Annual Review of Applied Linguistics* **24:** 26-43.

McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-Based Language Studies: an Advanced Resource Book.* London: Routledge.

McKey, M., Davis, M. & Fanning, P. (1995). *Message: The Communication Skills Book.* Oakland: New Harbinger Publications Inc.

McPherson, A. & Herxheimer, A. (2001). *DIPEx (Database of Individual Patient*

Experiences). http://www.healthtalkonline.org/. Oxford: The DIPEx Charity. [Accessed: 15 January 2010].
Mead, R. (1999). *International Management: Cross-cultural Dimensions.* Oxford: Blackwell.
Mizutani, N. (1983). Aizuchi to outou (Aizuchi and response). In Mizutani, O. (ed.) *Kouza Nihongo no Hyogen 3 (Lecture Series: Expressions in Japanese 3).* Tokyo: Chikuma Shobou. 37-44.
Mori, J. (2002). Task design, plan, and development of talk-in-Interaction: An analysis of a small group activity in a Japanese language classroom. *Applied Linguistics* **23:** 323-347.
Mori, J. (2006). The workings of the Japanese token *hee* in informing sequences: An analysis of sequential context, turn shape, and prosody. *Journal of Pragmatics* **38:** 1175-1205.
Nakane, I. (2006). Silence and politeness in intercultural communication in university seminars. *Journal of Pragmatics* **38:** 1811-1835.
Nesi, H. (2000). *The British Academic Spoken English (BASE) Corpus.* http://www.reading.ac.uk/AcaDepts/ll/base_corpus/. University of Warwick, University of Reading. [Accessed: 15 January 2010].
O'Donoghue, T., F. (1991). Taking a parsed corpus to the cleaners: The EPOW Corpus. *ICAME Journal* **15:** 55-62.
O'Keeffe, A. (2004). 'Like the Wise Virgins and All that Jazz': Using a Corpus to Examine Vague Categorisation and Shared Knowledge. In Connor, U. & T. Upton, A (eds.) *Applied Corpus Linguistics: A Multidimensional Perspective.* Amsterdam: Rodopi. 1-20.
O'Keeffe, A. (2006). *Investigating Media Discourse: Domains of Discourse.* London: Routledge.
O'Keeffe, A. & Adolphs, S. (2008). Response Tokens in British and Irish Discourse. In Schneider, K., P. & A. Barron (eds.) *Variational Pragmatics.* Amsterdam: John Benjamins Publishing Company. 69-98.
O'Keeffe, A. & Farr, F. (2003). Using language corpora in language teacher education: Pedagogic, linguistic and cultural insights. *TESOL Quarterly* **37:** 289-418.
O'Keeffe, A., McCarthy, M. & Carter, R. (2007). *From Corpus to Class room: Language Use and Language Teaching.* Cambridge: Cambridge University Press.
Ochs, E. (2005). Constructing Social Identity: A Language Socialization Perspective. In Kiesling, S. & C. Paulston, Bratt (eds.) *Intercultural Discourse and Communication: Essential Readings.* Oxford: Blackwell. 71-91.
Ohama, R. (2006). *Nihongo kaiwa ni okeru taan koutai to aizuchi ni kansuru kenkyu (Turn taking and response tokens in Japanese conversation).* Hiroshima: Keisuisha.
Pennycook, A. (2001). *Critical Applied Linguistics: A critical introduction.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.
Pennycook, A. (2006). The Myth of English as an International Language. In Makoni, S. & A. Pennycook (eds.) *Disinventing and Reconstituting Languages.* Clevedon: Multilingual Matters Limited. 90-115.
Pennycook, A. (2009). Plurilithic Englishes: Towards a 3D model. In Murata, K. & J.

Jenkins (eds.) *Global Englishes in Asian Contexts: Current and Future Debates.* Hampshire: Palgrave Macmillan.

Peters, P. (1986). *Manual to Accompany the Australian Corpus of English (ACE).* http://khnt.hit.uib.no/icame/manuals/ace/. Macquarie University. [Accessed: 15 January 2010].

Prodromou, L. (2005). 'You see, it's sort of tricky for the L2-user': the puzzle of idiomaticity in English as Lingua Franca. *PhD Thesis.* Nottingham: The University of Nottingham.

Renouf, A. (1984). Corpus Development at Birmingham University. In Aarts, J. & W. Meijs (eds.) *Corpus Linguistics: Recent Development in the Use of Computer Corpora in English Language Research.* Amsterdam: Rodopi. 3-41.

Sacks, H. (1987). On the Preferences for Agreement and Contiguity in Sequences in Conversation. In Button, G. & J. R. E. Lee (eds.) *Talk and Social Organisation.* Clevedon: Multilingual Matters Ltd. 54-69.

Sacks, H. (1992). *Lectures on Conversation, Volume 1.* Massachusetts: Blackwell Publishers.

Sacks, H., Schegloff, E., A & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* **50:** 696-735.

Saft, S. (2007). Conversation Analysis and the Study of Aizuchi: The Organization of Argument in Japanese Faculty Meetings In Ohara, Y. (ed.) *Nihongo disukousu no tayouna apurouchi (Various approaches in Japanese discouse).* Tokyo: Heibonsha. 38-70.

Schegloff, E., A (1984). On Some Gestures' Relation to Talk. In Atkinson, M., J. & J. Heritage (eds.) *Structures of Social Action: Studies in Conversation Analysis.* Cambridge: Cambridge University Press. 247-265.

Schegloff, E., A (1992). Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* **97:** 1295-1345.

Schegloff, E., A (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis, Volume 1.* Cambridge: Cambridge University Press.

Schegloff, E., A., Koshik, I., Jacoby, S. & Olsher, D. (2002). Conversation analysis and applied linguistics. *Annual Review of Applied Linguistics* **22:** 3-31.

Schegloff, E. A. (1987). Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation's Turn-taking Organisation. In Button, G. & J. R. E. Lee (eds.) *Talk and Social Organisation.* Clevedon: Multilingual Matters Ltd. 70-85.

Schiffrin, D. (1987). *Discourse Markers.* Cambridge: Cambridge University Press.

Schonell, F., Meddleton, I., Shaw, B., Routh, M., Popham, D., Gill, J., Mackerell, G. & Stephens, C. (1956). *A Study of the Oral Vocabulary of Adults: An Investigation into the Spoken Vocabulary of the Australian Worker.* Brisbane: University of Queensland Press.

Scollon, R. & Scollon, S. W. (1995). *Intercultural Communication: A Discourse Approach.* Oxford: Blackwell.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language.* Cambridge: Oxford University Press.

Sinclair, J. M. (1996). *Preliminary Recommendations on Corpus Typology.* Online documents: http://www.ilc.cnr.it/EAGLES/typology/node1.html. University of Birmingham. [Accessed: 8 January 2007].

Sinclair, J. M. & Coulthard, R. M. (1975). *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* Oxford: Oxford University Press.

Singer, M., R. (1998). *Perception and Identity in Intercultural Communication: A Abridged and Revised Edition of Intercultural Communication.* Yarmouth, ME: Intercultural.

Spencer-Oatey, H. (2000). *Culturally Speaking: Managing Rapport Through Talk Across Cultures.* London: Continuum.

Spencer-Oatey, H. (2002). Managing rapport in talk: Using rapport sensitive incidents to explore the motivational concerns underlying the management of relations. *Journal of Pragmatics* **34:** 529-545.

Stenstrom, B. (1990). Lexical Items Peculiar to Spoken Discourse. In Svartvik, J. (ed.) *The London-Lund Corpus of Spoken English: Description and Research.* Lund: Lund University Press. 137-177.

Stubbs, M. (1996). *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture.* Cambridge: Blackwell.

Stubbs, M., Robinson, B. & Twite, S. (1979). *Observing Classroom Language. Block 5.* Walton Hall: The Open University Press.

Swales, M. J. (1990). *Genre Analysis: English in Academic and Research Settings.* Cambridge: Cambridge University Press.

Tannen, D. (1984). *Conversational Style: Analyzing Talk Among Friends.* Norwood, NJ: Ablex Publishing Corporation.

Tannen, D. (1986). *That's Not What I Meant!* London: Virago Press.

Tao, H. (2003). Turn initiators in Spoken English: A Corpus Based Approach to Interaction and Grammar. In Meyer, C. & P. Leistyna (eds.) *Corpus Analysis: Language Structure and Language Use.* Amsterdam: Rodopi. 187-208.

ten Have, P. (2001). Applied Conversation Analysis. In McHoul, A. & M. Rapley (eds.) *How to Analyse Talk in Institutional Settings: A Casebook of Methods.* London: Continuum. 3-11.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics* **4:** 91-112.

Tono, Y. (2004). Learner Corpora and SLA Research.

Turner, J., M & Hiraga, M., K (1996). Elaborating Elaboration in Academic Tutorials: Changing cultural assumptions. In Coleman, H. & L. Cameron (eds.) *Change and Language: Papers from Annual Meeting of the British Association for Applied Linguistics held at the University of Leeds, September 1994* Clevedon: Multilingual Matters Ltd. 131-140.

van Lier, L. (1988). *The Classroom and the Language Learner : Ethnography and Second Language Classroom Research* London: Longman.

van Lier, L. (2004). *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective.* Norwell, MA: Kluwer Academic Publishers.

Vygotsky, L. S. (1962 [1934]). *Thought and Language.* Cambridge, MA: The MIT Press.

Walter, B. (1988). *The Jury Summation as Speech Genre.* Amsterdam: John Benjamins Publishing Company.

Ward, N. (1998). Some exotic discourse makers of spoken dialog. *ACL Anthology: A Digital Archive of Research Papers in Computational Linguistics.* 62-63.

Watanabe, S. (1993). Cultural Differences in Framing: American and Japanese Group Discussion. In Tannen, D. (ed.) *Framing in Discourse.* Oxford: Oxford

University Press. 176-208.
White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society* **18:** 59-76.
Yamada, H. (1997). *Different Games Different Rules: Why Americans and Japanese Misunderstand Each Other.* Oxford: Oxford University Press.
Yngve, V. H. (1970). On Getting a Word Edgewise. *Papers from the 6th regional meeting.* Chicago: Chicago Linguistics Society, April.
Zimmerman, D. H. (1998). Identity, Context and Interaction. In Antaki, C. & S. Widdicombe (eds.) *Identities in Talk.* London: Sage publications. 87-106.