A CORPUS-BASED ANALYSIS OF SIMULTANEOUS SPEECH IN ENGLISH CONVERSATION

by

HE ANPING

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the
requirements for the degree of
Doctor of Philosophy
in Applied Linguistics

Victoria University of Wellington

1996

ABSTRACT

This is a study of the distribution of 2011 instances of simultaneous speech in a 91,802-word subcorpus from the London-Lund Corpus of Spoken English. Five categories of simultaneous speech (successful and unsuccessful turn-bidding, successful and unsuccessful turn-competing, and backchannelling) were analysed in terms of:

(a) characteristics of the prosodic, lexical and grammatical context in which simultaneous speech occurs;

(b) linguistic devices and strategies in aspects of prosody, discourse and pragmatics which are frequently used to introduce simultaneous speech;

(c) variables such as speech domain, degree of familiarity between interlocutors, speakers' status and gender which may influence the frequency of simultaneous speech and affect the occurrence of the linguistic features and devices associated with simultaneous speech.

In a complementary case study, 288 instances of simultaneous speech in Chinese (Cantonese) were also analysed in a 10385-word sample of Chinese conversation, and compared with simultaneous speech in English.

The findings of the study show:

(a) Simultaneous speech is rule-governed and context-constrained. It is most likely to occur at a unit boundary which is prosodically, lexically and syntactically marked. It is often introduced and carried out by a number of prosodic devices, discourse items and repetition strategies. This is particularly the case in turn-bidding and turn-competition.

(b) Frequency of simultaneous speech seems to be strongly associated with degree of formality of speech domain and degree of familiarity between interlocutors,

but loosely related to speakers' status and gender. However, particular linguistic devices and strategies seem more preferred by interlocutors in a specific speech domain, or with a specific degree of familiarity, or having specific status or gender.

(c) Chinese and English simultaneous speech share many similarities in terms of pragmatic functions, and linguistic devices and strategies employed, though equivalents between the two languages are not always found. However social constraints on turn-bidding seem different in the two languages especially in terms of age, status and gender.

The descriptive findings of the study help explain why Chinese learners of English find it difficult to take a turn in English conversation, and especially to bid for a turn. Thus the study enhances our awareness of the linguistic features of English conversation and the factors which can affect Chinese students' pragmatic and discourse competence. Moreover, the computer corpus approach adopted in the research provides a way of obtaining rich input for teaching English discourse devices in terms of prosody, lexicon and syntax and suggests further applications of corpus-based research in the study of language teaching and learning.

ACKNOWLEDGMENTS

I authorise Victoria University of Wellington to lend this thesis to other institutions or individuals for the purpose of scholarly research.

He Anping

I further authorise Victoria University of Wellington to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

He Anping

Victoria University of Wellington requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

CONTENTS

LIST OF TABLES

Transcription Conventions

The following symbols from the LLC coding are used in the prosodic transcription of examples in the thesis. Additional coding was added for the thesis. (See p.xv.)

TYPE	EXAMPLE	EXPLANATION
mark-up of the text	S1	text in the corpus
	S3.2a	subtext
	S3.2a: 1130	printing line number in the text
	(laughs)	contextual comment
	((yes))	subaudible words
	((sylls))	incomprehensible words
	yes	simultaneous talk
tone unit	#	end of tone unit
	^	onset
	{yes}	subordinate tone unit
nuclear	y\es	falling tone
	y/es	rising tone
	y=es	level tone
	yVes	fall-rising tone
	y/\es	rise-falling tone
	y\es y/es	fall-plus-rise tone
	y/es y\es	rise-plus-fall tone
booster	yes	continuance
	:yes	higher than preceding syllable
	|yes	higher than preceding pitch-prominent
	||yes	very high
stress	`yes	normal stress
	``yes	heavy stress
pause	yes . yes	brief pause (of one light syllable)
	yes - yes	unit pause (of one stress unit or 'foot')
graphic	A:	surreptitious speaker
	a:	non-surreptitious speaker
	[@],[dhi:]	equal to [ ]and [i:] in normal orthography, the ':' after 'i' means a prolonged sound of 'i'

Note: For the reader's convenience, some of the prosodic markers in the original LLC have not been included here. Only those markers which are relevant to the analysis being made here are included.

Additional Coding of the Subcorpus

==

	Current Speaker's Side	Overlapper's Side
Category of simultaneous speech		
STB (successful turn-bidding)	%	& ($ at the end)
UTB (unsuccessful turn-bidding)	|	‖
BC (backchannelling)	>	>>
STC (successful turn-competing)		<<
UTC (unsuccessful turn-competing)		<

==

Identity of speakers		
speaker with higher status		H
speaker with lower status		L
male speaker		M
female speaker		F

==

Environmental features		
syntactic boundary	G	
word boundary	W	
address tag	A	
hesitation marker	H	
termination tag	T	
emphasizer	E	
connector	C	

==

Pragmatic strategy		
repetition		R

==

CHAPTER I

INTRODUCTION

In general conversation, a commonly accepted norm is that speakers should take turns to speak, one speaker at a time. Yet in reality, it frequently happens that more than one person attempts to speak at the same time resulting in overlap or simultaneous speaking. Some simultaneous speech occurs in the middle of the current speaker's turn when the hearer is giving feedback usually known as a backchannel; some simultaneous speech occurs near the end of the current speaker's turn when the next speaker is trying to bid for a turn; and some simultaneous speech occurs when two speakers compete for the floor after a prior speaker has finished. The present research studies all these kinds of simultaneous speaking phenomena in a 91802-word computerised corpus of native English speakers' conversation. The study has been undertaken in the belief that a conversation analysis on a database of this size will reveal the recurrent patterns and mechanisms of these different kinds of simultaneous speech and show strategies of listening and speaking. The result should contribute to the teaching and learning of English by non-native speakers in particular.

Many non-native English learners have experienced difficulties in English conversation, including how long one is expected to speak for, when one is obliged to be silent, and how one should take the floor, or give feedback or make an interruption. A study of 101 mainland Chinese students in 7 British universities and 37 of their British supervisors (Jin: 1992) reveals that many students "experienced problems with perceptions of turn-taking and pausing in spoken English which means that they felt unable to participate adequately in seminars or

tutorials" (Jin & Cortazzi, 1995). Similarly many native English-speaking teachers find their Chinese students not active enough in English conversation and lacking in positive listening, thus resulting in ineffective communication. It is not known whether this is due to a cultural difference in speech etiquette or whether it is due to a lack of an English language communicative strategy. As a Chinese who has been teaching English for about twenty years in Chinese primary schools, middle schools and at university level and is now living and studying in an English speaking country, I have taken the opportunity to study English conversations so as to find out:

(a) how often simultaneous speech occurs in native English speakers' conversation;

(b) how simultaneous speech is affected by different speech domains, by different degrees of familiarity between interlocutors, and by the relative status and gender of the interlocutors;

(c) when, where and how simultaneous speech is likely to occur;

(d) how findings on the above can be applied to English teaching for non-native English learners .

Previous studies of simultaneous speech have addressed these phenomena from different directions and by different approaches, particularly in studying the phenomenon of interruption. Researchers such as Duncan (1972, 1973), Jefferson (1973), Sacks, Schegloff & Jefferson (1974), Ferguson (1977), French & Local (1986), Schegloff (1987), Stenström (1994) have taken a linguistic approach to the formal aspects of interruption and studied the linguistic bases, the language features and discourse functions of interruption. Others such as Zimmerman & West (1977), Tannen (1984, 1989, 1990), Roger et al. (1988), Coates (1989), Ng et al. (1993) have taken a sociolinguistic or psycholinguistic approach to the functional aspects of interruption, such as the association between interruption and social context, interruption and gender, interruption and power or dominance. Other kinds of simultaneous speech such as backchannels or brief feedback have been studied by Crystal & Davy (1969), Duncan (1973), Kendon (1977), Tao &

Thompson (1991) who looked at form and function and ethnic and cultural differences in their use.

However, these previous studies of simultaneous speech fall either into the category of qualitative case studies, or into quantitative studies using only a few dozen examples, or focusing on only one kind of simultaneous speech. There is not yet a comprehensive study of all kinds of simultaneous speech in a large body of data, looking at potential strategies of positive and active listening and speaking. Neither is there a large scale study made of all kinds of simultaneous speech in terms of its multi-layer linguistic features --- prosodic, lexical, syntactic, and in terms of its association with different conversational domains and the speaker's age, status, gender and cultural background. The present study focuses on the linguistic aspect of simultaneous speech and its relation to some social factors such as speech domains, gender and status of participants, and its relevance for the teaching and learning of English. Furthermore, instead of being based entirely on experimental or case study data, a much larger corpus of naturally occurring speech is used in the present study for analysis by computer.

This study investigates simultaneous speech using a 91,802-word subcorpus from the London-Lund Corpus of Spoken English (LLC) (Svartvik & Quirk, 1980), and explores the linguistic environment and devices used for different kinds of simultaneous speech. Both a quantitative and qualitative analysis of the corpus is undertaken. This study takes a formal linguistic point of view in looking for the recurrent language patterns or features of different kinds of simultaneous speech, a view which can reveal insights into current English use and which can produce rich input for English teaching. It also takes a sociolinguistic view of how different social factors, which could contrast with Chinese traditional values, may influence the occurrence of each kind of simultaneous speech. No corpus of spoken Chinese currently exists to compare with the English data from the LLC. A case study

however, has been undertaken using Chinese conversational data available to the researcher and a comparison has been made between a Chinese sample and an English sample of casual conversation.

The description of conversational norms in this study is based on, and limited to, corpora of spoken British English and Cantonese, and conclusions drawn about them do not necessarily apply to other varieties of English or Chinese. One of the sources of data in this study is a 91,802-word corpus of British English conversation and the other is a 10,385-word corpus of Cantonese conversation. Cantonese does not represent the Chinese-speaking world any more than British English does the English-speaking world. However, these sources of data were the only ones currently available. Nevertheless, as one of the seven major "dialects" of Chinese, Cantonese is similar enough to the grammatical structure of Mandarin in its syntactic aspect to be used in the comparative study in Chapter V. The differences in the corpora are discussed in Chapter IV, and the interpretation of the results of the comparison of the two languages based on them in Chapter V needs to be seen in this light.

The potential significance of this study is that it provides a more comprehensive picture of the multi-layer linguistic features of the speech mechanisms in carrying out different kinds of conversation, particularly the aspect of active listening responses. Both quantitative and qualitative findings from this study could become valuable input for teaching English to either native or non-native English learners, particularly in training them to be active and positive participants in English conversation. Moreover, the case study of Chinese conversational data can throw further light on the nature and processes of simultaneous speech, and increase the mutual understanding between English teachers and learners who do not share the same cultural backgrounds and norms so that language teaching can be carried out

in a more cooperative and effective way. The present study is concerned with the following research questions:

Research question 1

How often does simultaneous speech occur in English conversation? How is it distributed in different kinds of simultaneous speech such as turn-bidding, turn-competing and backchannelling?

Research question 2

What are the characteristics of the linguistic environment (prosodic, lexical, grammatical) which are most likely to be associated with simultaneous speech as a whole and associated with different kinds of simultaneous speech in particular?

Research question 3

What are the most frequently used linguistic devices and strategies (prosodic, discoursal and pragmatic) for introducing different kinds of simultaneous speech and how do they function in a successful conversation?

Research question 4

How are the frequency of different kinds of simultaneous speech and their linguistic features associated with speech domain, degree of familiarity between interlocutors and speakers' status and gender?

Research question 5

> How do these characteristics of English simultaneous speech compare with simultaneous speech in Chinese conversation and how could any differences help account for difficulties experienced by Chinese learners of English?

Chapter II contains a literature review which addresses a number of key terms and issues in this study, the different approaches which have been used for studying simultaneous speech, and their findings as well.

Chapter III provides a detailed account of the methods and procedures designed for this study. It describes the methods used for both English and Chinese data collection, corpus construction and editing, and data retrieval.

Chapter IV presents the results of the analysis of the English spoken corpus in which the findings are interpreted and the first four research questions are discussed.

Chapter V presents the results of a case study of the Chinese spoken data in comparison with the English sample. The findings are interpreted and the fifth research question is discussed.

Chapter VI draws conclusions from the previous two chapters and discusses their implications for language description and English teaching, particularly in China.

CHAPTER II

LITERATURE REVIEW

2.1 APPROACHES TO THE STUDY OF CONVERSATION

Conversation is the most typical kind of verbal behaviour in human communication. Generally it is a speech interaction between two or more participants. A normal conversation is a sequenced alternation of speakers. That is, the participants take turns to speak. When a current speaker A is speaking and the other interlocutors, B and/or C are listening, A is said to be having a turn. When A comes to the end of the turn and the next two speakers, B and C, start speaking at the same time, they are competing for a turn. When only one speaker, B, takes up a turn after A has stopped speaking, he or she is claiming a turn. If the next speaker, B, starts speaking even before A stops speaking, B may be giving feedback to A or may be bidding for a turn by interrupting A's speaking. In studying how such a turn-taking system actually works in conversation, there have been two main approaches in the language sciences, namely conversation analysis and discourse analysis.

Conversation analysis is distinguished from discourse analysis by Levinson (1983: 286-287)) and Schiffrin (1994: 223) as being based on a rigorously empirical approach derived from ethnomethodology. It is interested in how participants in conversation construct systematic solutions to recurrent organisation problems such

as how to carry on turns, allocate turns, bid for turns, compete for turns, or support

on-going turns. The assumption is that:

> it is because actors succeed in using the sequential progression of
> interaction to display their understandings of its events and rules that the
> shared world that has been·jointly achieved is publicly available for
> analysis.
>
> (Taylor & Cameron, 1987: 104, cited by Schiffrin, 1994: 234)

Therefore conversation analysis, as practised by Sacks, Schegloff & Jefferson

(1974) and others, was undertaken by locating as many instances as possible of

particular turn-taking phenomena in a corpus of naturally occurring conversation,

and then searching for recurrent patterns across those records, (sequentially-based

distribution), showing that what are hypothesised to be sequential expectations are

actually carried out by participants. In other words, such analysis places emphasis

on the interactional and inferential consequences of the choices between alternative

utterances and thus inductively generalises a set of rules in a turn-taking system.

On the other hand, discourse analysis

> employs both the methodology and the kinds of theoretical principles
> and primitive concepts (eg. rule, well-formed formula) typical of
> linguistics. It is essentially a series of attempts to extend the techniques
> so successful in linguistics, beyond the unit of the sentence.
>
> (Levinson, 1983: 286)

Such an approach, as practised by Sinclair & Coulthard (1975), Coulthard & Brazil

(1979) and others, proceeds in a deductive way. It firstly isolates a set of basic

categories or units of discourse and formulates a set of concatenation rules over

those categories, delimiting well-formed sequences. Then it takes one or more texts

(often constructed by the analyst) and analyses them in depth for all the interesting

features of this limited domain, particularly in terms of syntactic description, and

thus reveals "what is really going on" in this piece of conversation discourse (see Levinson, ibid.).

The present study adopts both approaches, because the writer shares Schiffrin's view that "the link between knowledge and action has an important bearing on the study of language" (1994: 224). In this study the conversation analysis approach may start on a large database of naturally occurring English conversation on which all instances of simultaneous speech are sorted out, regardless of their speakers' identities, status or age, but with an emphasis on the authenticity of the research data. Then when analysing various features of simultaneous speech within the corpus, a discourse analysis approach is adopted to describe the form and function of those features in terms of prosody, lexicon, syntax and discourse coherence. Finally when exploring the association between the linguistic characteristics of simultaneous speech and their social dimensions, both approaches contribute again to the study. Before we go into the methods and procedures in detail, a definition of a number of key terms and a survey of literature related to simultaneous speech is necessary.

2.2 DEFINITION OF KEY TERMS

2.2.1 TURN

Any piece of conversation can be regarded as an example of turn-taking in which a turn is a basic unit. A turn can be defined in a technical way or be based on the participants' intuition. The latter approach means that having a turn is a state of mind in addition to the display of particular behaviour and it should be judged by the speaker's sense or intention (see Edelsky, 1981: 385). However, the former approach, ie., the formal description of a turn, seems to be used by more researchers in conversation analysis. Some major definitions have been summarised by Edelsky (ibid) as follows:

Jaffe & Feldstein (1970) were the first who advocated a technique of chronography to identify a turn. They defined a turn as a solo talk beginning the instant one person starts to talk and ending prior to the instant someone else begins to talk alone. Obviously any overlaps are not considered part of anyone's turn.

Duncan (1973) looked for signals that themselves organise the stream of talk and that thus accomplish the smooth exchange of turns. Therefore he defined a turn as a unit of interaction with an end boundary marked by a turn-claiming response from the auditor.

Sacks, Schegloff & Jefferson (1974) emphasised the technical, non-intuitive characterisation of turns and said what occupies a turn slot is a turn unit containing one or more (turn) transition-relevance places (ie. a possible completion point) and that any change of speaker constitutes a trade of turn.

In this study, a turn is also defined in a technical way. It adopts Stenström's more simple and direct definition, saying that a turn is "everything the current speaker says before the next speaker takes over" (Stenström, 1994: 33). eg.

[1]		TURN
	B: ...I think they've got quite a good opinion of him#-	1
	A: well [@] I I have too#	2
	B: m#	3
	A: [@:m] -- ((well I mean)) . the way these chaps go#	4
	B: [m] [hm]#	5

(cited by Stenstöm, 1994:34)

Note:
For an explanation of all prosodic markers in the examples used in this thesis, see page xiv.

In [1], there are five turns with four turn-shifts involved. All the utterances, regardless of length, constitute turns in their own right, even B's "m" and "mhm" in this case. Let us consider another example:

[2]   TURN

     B: ...I've got three or four years more of examining which makes *a*   1

     A: *m*

     B: spot of money for me# ---

     A: oh well I don't think that anybody suggests for one moment   2

                       (cited by Stenström, 1994: 35)

Note:

The asterisk "*" in the examples throughout the thesis indicates the beginning and the end of simultaneous speech. For the convenience of printing, the overlapped speech of two speakers is not normally displayed in parallel.

In [2], there is only one turn shift. A's "m" starts and ends within B's turn indicating A is only giving a brief feedback to B, but not taking over the turn. Therefore a turn-taking presupposes a shift of speaker only when the speaker takes over the floor. Any utterance produced while the other party goes on speaking can consequently not be regarded as a turn. The advantage of such a definition is that, firstly, analysis can be carried out by machine, thus surveying more data without depending on the participants' personal judgement; secondly, it can make use of the linguistic forms used by speaker and hearer to reveal certain features of synchronised speaker change which are the focus of this study.

2.2.2   SIMULTANEOUS SPEECH

Simultaneous speech, in which two (or more) speakers talk at the same time, is a frequent phenomenon in conversation. The present study is interested in the following aspects of simultaneous speech:

(a) where does it occur?

(b) between whom does it occur?

(c) how long does it last?

(d) how is it actually realised?

(e) what function does it play in conversation?

To answer these questions, a further classification of simultaneous speech is necessary. In previous studies, some researchers classified simultaneous speech by means of functional criteria, but more researchers categorised simultaneous speech with formal criteria.

In terms of functional criteria, West and Zimmerman (1983) classified simultaneous speech by distinguishing "overlap" from "interruption". This is made by the researcher's subjective judgement, ie. to see whether simultaneous speech is facilitative or disruptive. They said:

> overlaps are (1) events occurring in the immediate vicinity of possible turn-transition places; and (2) those brief utterances (eg. "yeah", "right") or longer incursions (eg. "saying the same thing at the same time") which have some facilitative warrant. ... (While) interruptions, in contrast, have no such facilitative warrant. Instead, these incursions have the potential to disrupt turns at talk, disorganise the ongoing construction of conversation topics, and violate the current speaker's right to be engaged in speaking.
>
> (West & Zimmerman, 1983: 105)

Based on West & Zimmerman's distinction between overlap and interruption, Kennedy & Camden (1983) made a further classification of interruptions in terms of their functions. They identified interruptions in terms of "Clarification" (ie. an attempt to understand the current speaker's message), "Agreement" (ie. a demonstration of agreement, support, concurrence or compliance with the current speaker's idea), "Disagreement" (ie. a demonstration of rejection, disagreement, challenge or contradiction of the current speaker's communication), "Tangentialization" (ie. a speech which reflects an awareness of the current speaker's statement or in some way minimises it), "Subject change" (ie. no awareness of the current speaker's statement or an issue of a new theme irrelevant

to the current speaker's) and "Others" (ie. any speech not appropriate to the above categories) (see Kennedy & Camden, 1983: 51).

Tannen, on the other hand, categorised interruption criteria according to whether it is intended by the interruption maker and whether it is perceived as being disruptive by the person whose speech is interrupted. In other words, the researcher has to know a lot about both speakers and the situation, ie. what the current speaker has been actually doing and what the second speaker is actually trying to do. Therefore "it is a matter of individual perceptions of rights and obligations, as they grow out of individual habits and expectations" (Tannen, 1990: 192). Based on this, she classified several types such as interruption without overlap, overlap without interruption, successful cooperative overlapping and unsuccessful cooperative overlapping (see ibid: 192-200).

As the present study is carried out on a large computerised corpus, it is impossible to go into each simultaneous speech case in detail, nor can one retrospectively know the speakers' actual perceptions, feelings and purpose when they were overlapping others' speech or when their own speech was overlapped. Therefore in using a computerised corpus as database, it is necessary to use formal criteria as a starting point, as Sacks (1989: 29) said: "naturally occurring social activities are subjectable to formal description" and such description can "permit us to see non-trivial ways that actual activities in their details are simple." (cited by Psathas, 1995: 14). In terms of formal criteria, the following three models have been the most frequently used by previous researchers. They are the three formal classifications of simultaneous speech made by Sacks et al. (1974), Duncan (1973) and Ferguson (1977) respectively.

First, in the model of turn-taking proposed by Sacks et al. (1974), all simultaneous speech is called overlap, among which three kinds of overlaps are further identified:

(a) self-selectors' competition, ie. two or more new speakers compete for a new turn after the current speaker's turn is finished, see [3];

(b) overlap, ie. a speaker starts to speak at a transition-relevance place of the current speaker's turn, see [4];

(c) interruption, ie. a speaker starts to speak at a non-transition-relevance place of the current speaker's turn, see [5].

[3]   Mike: I know who d' guy is -
        Vic: * he's ba:::d*
        James: *you know* the gu:y

                                    (cited by Sacks et al., 1974: 707)

[4]   Desk: is it a stretcher patient *ma'am*
        Caller: *it's *- uh yes he is

                                    (cited by Jefferson, 1973: 47)

[5]   A: and Wednesdays I go back in the evening -I *take the car so I*
        B: *oh you come home?*
        A: leave about five o'clock.

                                    (Schegloff, 1987: 76)

Here in [3], Mike yields his turn without particularly selecting the next speaker, therefore Vic and James both start to talk at the same time to compete for a turn. In [4], the Caller starts to speak when the Desk says "ma'am" which is an address term occurring in "tag" position as the last particle of an otherwise possibly complete utterance, so it is an overlap. However in [5], B interrupts A after A has just started a new utterance. "I..." is obviously not a possible completion point; therefore it is an interruption. In summary, the categorisation proposed by Sacks et al. reiterates two points: firstly, whether the simultaneous speech occurs at a place where the current speaker finishes the turn without selecting the next speaker and secondly, whether the simultaneous speech occurs at a transition-relevance place of the current speaker's turn, ie. a possible completion point in the current speaker's

utterance. Yet this model does not account for those short utterances like "yeah" or "uhuh' which do not cause a turn-shift but are just made by the hearer simultaneously accompanying the current speaker's turn. As a remedy, Schegloff (1981) later called those discourse items "continuers" in conversation.

Second, Duncan's categorisation of simultaneous speech (1972: 286) takes full account of auditor response occurring simultaneously with the current speaker's speech, no matter whether it is verbal or non-verbal and no matter whether it causes a turn-shift or not. He classified them into two major kinds:

(a) simultaneous talk, ie the auditor gives a brief feedback to the current speaker without constructing a claim for the turn, see [6] below;

(b) simultaneous turn, ie the auditor claims the turn when the current speaker's turn has not finished, see [5] above.

[6]   B: I've got three or four years more of examining which makes *a* spot of
        A: *m*

                                          (LLC, S1.1: 411-414)

Here in [6], A's "m" is a brief response to B's speech and does not constitute a claim for the turn. However in [5], B, the auditor attempts to take his turn while the original speaker continues with his. Thus both participants simultaneously claim the speaking turn. Duncan's analysis emphasises the auditors' turn-taking attempts, ie. whether they are claiming the speaker's turn or retaining their auditor's role. His "simultaneous talk" corresponds to "backchannel behaviour" (Yngve: 1970), or an "accompaniment signal" (Kendon: 1967), or a "supportive minimal response" (Stubbe: 1991) or a "continuer" (Schegloff: 1981).   And his "simultaneous turn" is similar to Sacks et al.'s "interruption". However as Duncan considers "the 'turn-taking mechanism' is not designated to explain how the state of simultaneous turns is resolved" (1972: 286), his model does not address the result

of simultaneous speech, ie. which one of the two speakers will continue and which one will fall silent after the overlap ends.

Third, Ferguson's categorisation (1977: 296-297) emphasises the result of simultaneous speech. Her criteria in classifying simultaneous speech and interruption involve three conditions: first, whether there is simultaneous speech between two participants; second, whether such simultaneous speech causes a break in verbal continuity in one speaker's output; and third, whether the second speaker takes the floor after simultaneous speech. As a result, she classifies four kinds of simultaneous speech:

(a) simple interruption, ie. the second speaker's interruption breaks the continuity of the current speaker and gains the floor, see [7].

(b) butting-in interruption, ie. the second speaker's interruption breaks off before it is completed, see [8].

(c) overlap, ie. the second speaker takes the floor after a brief overlap but such simultaneous speech does not cause an apparent break in the current speaker's turn, see [9].

(d) interjection remarks, ie. the auditor's brief response which does not cause a speaker switch, see [9] again.

[7]   (Simple interruption)
        A: ... and this bit about him being bankrupt and having no money I just
               don't see how it's possible be***cause***
        B: ***I hav***en't heard that.

                                                            (Ferguson, 1977: 296)

[8]   (Butting-in interruption)
        A: I don't know ***I've got mix***ed feelings I think it would be nice to
               have a baby
        B: ***I think I - ***

                                                            (ibid: 297)

[9]   (overlap, interjection)
       A:...I expect you would like to go **with him**
       B: **well I'**d prefer **it** yeah - but then he would want me to go to a
           Ranger's football match ...
       A: **yes**   (ibid: 296)

According to Ferguson's description, in [7], B breaks the continuity of A's turn and takes the floor. In [8], B interrupts A but fails to complete his own utterance, because A continues to speak after the simultaneous speech. In [9], B gains a turn but in contrast with [7], there is no apparent break off in A's turn. Again in [9], A's "yes" is regarded as an interjection response. It does not lead to a turn-switch, because B continues to speak after that. In Ferguson's category, "simple interruption" seems to correspond to Mishler & Waxler's (1968) "successful interruption", while "butting-in interruption" corresponds to "unsuccessful interruption". The case of "overlap" here is quite similar to Sacks et al.'s second type of overlap while the case of "interjection remarks" is like a backchannel in Duncan's model. Although Ferguson accounts for the result of interruption, her criterion of "breaks in verbal continuity of one speaker's output" (Ferguson, 1977: 297) seems to address both interrupter and interruptee's output, and this could cause confusion in judging who is actually being interrupted, because in [8], according to her description, it is B, the interrupter's verbal continuity being broken off.

Benefiting from the above three models, the present study maintains a formal rather than functional classification of simultaneous speech. It adopts the formal criteria set out above but with some modifications. It also subcategorises some formally classified simultaneous speech according to functional criteria. Simultaneous speech is regarded as a conversational strategy because it can display hearers' active participation in conversation and their strategies in taking turns with a smaller gap between the current speaker and the next speaker. Moreover, as this study was

carried out on a computerised corpus, it was necessary to develop a model with a set of definitions and classification criteria as unambiguous and explicit as possible so that the data could be sorted out by means of the computer. Here "unambiguous" means that the model can clearly separate the kind of simultaneous speech which occurs when a speaker is trying to take over the current speaker's turn, from the other kind of simultaneous speech which occurs when a speaker is just offering a brief feedback (usually supportive and agreeable) to the current speaker's words. By "explicit", we mean that the model should have the power of classifying different types of simultaneous speech by their distinguishable formal appearance, ie. represented by different formal features. It is for this reason that the present study takes the recurrent formal patterns as primary criteria to classify different categories of simultaneous speech and then uses the functional criteria as complementary only when ambiguity occurs in the formal patterns.

In terms of formal position, simultaneous speech can occur as follows:

   (a) at the beginning of two turn-competitors' utterances:

      A:.... *

      B: *..... *

      C: *..... *.....

This is the formula of a turn-competing case (TC) in which A is the prior speaker who ends the turn without selecting the next speaker. Therefore, B and C self-select themselves to compete for the next turn. Here B is regarded as an unsuccessful turn-competing case (UTC) while C is a successful one (STC). Whether one can continue to speak after simultaneous speech is a key criterion in judging whether one has taken the turn or gained the floor to speak. (See [3]).

(b) near the end of the current speaker's turn:

A:.... *..... *

B: *..... *........

This is the formula of a typical successful turn-bidding case (STB) in which B is regarded as a successful turn-bidder, because B continues to speak after the overlap (see [7])

(c) in the middle of the current speaker's turn:

A:.... *..... *......

B: *..... *

This is the formula of either a turn-bidding case or a backchannelling case, because B can either be a turn-bidder or a backchannel maker. This needs further categorisation according to its content and function.

In terms of functional meaning, the third category of simultaneous speech above can be subdivided into:

(c1) backchannelling (BC), if it is composed of a single sound like "mm, mhm, uhuh, aha, yes, yeah, ok, right" or short phrases like "I see", "That's right" etc.(see [6])

(c2) unsuccessful turn-bidding (UTB), if it is not composed of the words in (c1) but by longer utterances whose meaning is commenting, or challenging, or questioning the current speaker's utterance or trying to start a new topic (see [8]).

Finally the five categories can be summarized as in the following diagram. More detail is given in Chapter II.

Five Categories of Simultaneous Speech

```
                                    successful (STB)      A:....  *.....  *
                                                          B:      *.....  *....
                  turn bidding (TB)
                                    unsuccessful (UTB)    A:....  *.....  *....
                                                          B:      *.....  *

Simultaneous                                              A:....
Speech            turn competing(TC)  successful (STC)    B:      *.....  *....
                                      unsuccessful(UTC)   C:      *.....  *

                  backchannelling(BC)   ----------------> A:....  *.....  *....
                                                          B:      *.....  *
```
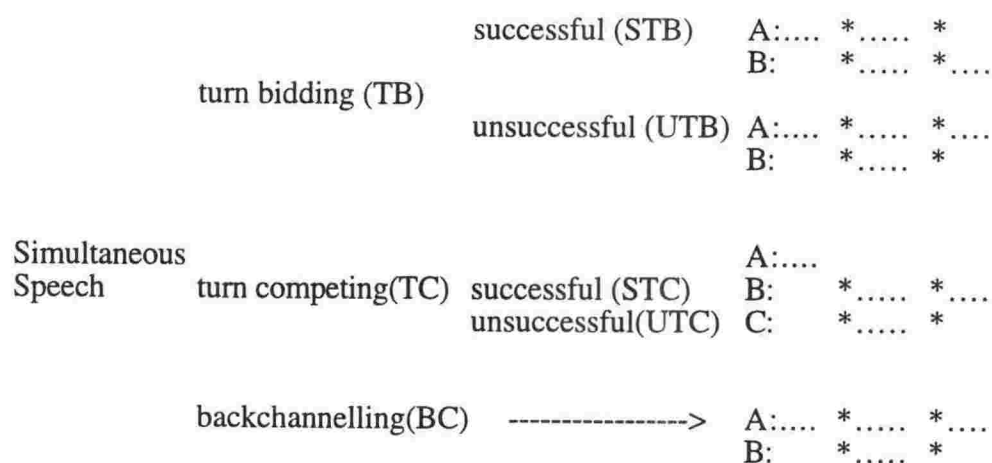
In this diagram, all the abbreviations of simultaneous speech such as TB and BC refer to the process of making that type of speech act. TB corresponds to Sacks et al.'s "interruption" or to Duncan's "simultaneous turn". It is carried out by a hearer who starts speaking when the current speaker has already started speaking (no matter how long he or she has spoken) thus causing some verbal overlap between the two (no matter how long that overlap lasts). STB corresponds to Ferguson's "simple interruption" while UTB corresponds to her "butting-in interruption". In other words, STB is a successful turn-bidding case, because the second speaker continues to speak after the overlap ends (no matter how long he or she continues). But UTB is an unsuccessful case, because the second speaker does not continue to speak after the overlap ends.

TC is what Sacks et al call "self-selectors" competing for the turn in which two (or more) subsequent speakers strive for a turn after the current speaker's turn is finished. It is further categorised into successful turn-competing (STC) and unsuccessful turn-competing (UTC) using the same criterion which were used to distinguish STB from UTB.

The term "backchannelling" (BC) is adopted from Yngve (1970). It is mainly composed of single words like "mm", "mhm", "uhuh", "aha", "oh", "yeah", "yep", "yes", "ok", "right", "good", but it does not extend backchannel activity (as Duncan (1972) does in his model) to include sentence completions, requests for clarification, brief statements and non-verbal responses.

The rationale for the model used in this thesis is:

(a) It uses the more neutral terms "turn-bidding" and "turn-competing" instead of "interruption" and "overlap", because bidding for a turn does not necessarily mean the turn-bidder is deliberately wishing to interrupt, thus depriving the current speaker of the right to speak, nor does it mean that the current speaker actually feels interrupted by the turn-bidder. The word "interruption" often has a negative connotation in social behaviour and the word "overlap" is too general and not sensitive to different kinds of simultaneous speech. Moreover the words "turn-bidding" and "turn-competing" are more closely related to hearers' turn-taking or turn-responding strategies in conversation, which is the focus of the present study.

(b) It uses a formal classification as a starting point, which is easier to use in identifying and annotating the three general categories of simultaneous speech (ie. TB, TC and BC) in a large tagged computerised corpus. Formal classification need not go into the detailed content of the corpus, which may be complicated and unreliable (due to the lack of audio and video recording of the presently available LLC data). Then, based on the formally identified simultaneous speech, we may, with the help of functional criteria, go into the verbal meaning and function of the third category of simultaneous speech, so as to distinguish UTB from BC by looking at its verbal and contextual meaning.

(c) It emphasises the outcome of simultaneous speech by subcategorising the successful and non-successful turn bidders or turn competitors (ie. it subdivides TB into STB vs. UTB, and TC into STC vs. UTC), because by exploring the linguistic environment and devices associated with these two categories, we can obtain insight into the mechanisms of discourse. In addition, possible linguistic input for helping develop language learners' communicative skills may be obtained.

2.3   PREVIOUS RESEARCH ON SIMULTANEOUS SPEECH

After setting up the classification of five categories of simultaneous speech, we can explore this phenomenon in terms of:

(a) why simultaneous speech occurs --- its pragmatic foundation

(b) where or when it is likely to occur --- its linguistic environment

(c) how it is actually realised --- the linguistic devices and strategies which are used to initiate simultaneous speech

(d) how it is associated with social context --- the sociolinguistic features which influence simultaneous speech

The following are the major findings related to these four aspects in previous research.

2.3.1   THE PRAGMATIC FOUNDATION OF SIMULTANEOUS SPEECH

In addressing the pragmatic foundation of simultaneous speech, three different approaches should be mentioned. They are (a) Sacks et al's sequence-producing

approach, (b) Duncan's signalling approach and (c) Stenström's discourse approach.

(1) Sequence Producing Approach

Sacks et al. (1974) consider that turn-allocating and turn-taking is a rule-governed system in which two basic patterns are posited:

Pattern 1: One party talks at a time, which is achieved by
(a) the current speaker selecting the next speaker when he or she yields a turn by addressing him or her verbally or non-verbally,
(b) one subsequent speaker self-selecting to start speaking after the current speaker ends the turn without selecting the next speaker.

Pattern 2: More than one party talks at a time, which is achieved by
a) more than one "next speaker" self-selecting themselves (in multi-party conversation) to start speaking after the current speaker ends the turn without selecting the next. (see [3])
b) one subsequent speaker starting to talk even before the current speaker actually yields the turn.

Obviously it is in the situation of Pattern 2 that simultaneous speech occurs. As described by Sacks et al., in the case of Pattern 2 (a), when the current speaker ends the turn without selecting the next , he or she in fact

in allocating a turn to that self-selector who starts first, encourages (the) earliest possible start for each self-selector. It thereby provides for overlap by competing self-selectors for a next turn, when each projects his start to be (the) earliest possible start at some possible transition-relevance place, producing simultaneous starts.

(Sacks et al, 1974: 706-707)

However in the case of Pattern 2 (b), there exists another basis of overlap which

derives from the projectability of possible completion or transition-relevance places. Variation in the articulation of the projected last part of a projectably last component of a turn's talk, which is in fact a consequential locus of articulatory variation, will expectably produce overlap between a current turn and a next.

(ibid)

[10]   A: well if you know my argument why did you bother to a:*sk*
       B: *be*cause I'd like to defend my argument

(cited by Sacks et al. 1974: 707)

In [10], B may regard A's "why did you bother to ..." as a projected last part of the turn. Therefore he takes it as a transition-relevance place and starts speaking, resulting in overlapped speech.

(2)   Signalling Approach

Duncan's signalling approach considers that the turn-taking mechanism which involves turn-yielding, claim-suppressing and within-turn segmentation signals is essential in organising and directing the stream of conversation. Duncan is interested in the cooperation between speakers' turn signals and auditors' responses. His research reveals that:

Simultaneous turns may be caused by a violation of the system on the part of either participant. The previous speaker may fail to relinquish his turn after displaying a turn signal and the auditor's subsequent claim of the turn; or the previous auditor may suddenly interrupt by claiming the turn in the absence of the speaker's turn signal.

(Duncan, 1973: 33)

[11]   B: ... it was the ^day \after# - *^on ((my 1 syll)) [@]*
       A: *^m\ay have been# - ^m\ay have been#*
       B: ^when I r=ang# ((and)) we and we ^fixed up to meet ((in our h\ouse#))

(LLC, S1.2: 90-94)

In [11], B, the current speaker has produced, before the overlap, some signals which could be regarded as turn-yielding signals in Duncan's terms, such as a phonemic clause boundary "#", a falling nuclear tone on "after", and a pause after it. Yet B does not relinquish his turn after giving such signals, and therefore, A's subsequent claim of the turn causes a simultaneous turn. This is a case of the current speaker's violation of the turn-taking system in Duncan's sense. An example of the second cause of a simultaneous turn is as follows:

[12]   B: ^I suppose I ^mean *. pre^sVumably#*
        A: *you ^have some "p\ull* {with the ^m\anagement 'do you#}#
                                (LLC, S.2.10: 775-76)

Here A starts his turn after B's "I mean" which is by no means a turn-yielding signal in Duncan's terms. A's sudden interruption also causes a simultaneous turn and this is a case of current auditor's violation of the turn system in Duncan's sense.

(3) Discourse Approach

Stenström's discourse approach summarises three reasons for the occurrence of interruption in conversation, which are also relevant to this study. According to her observation, when B is listening to A's speech, B may interrupt A in the following three situations:

(a) B has got the impression that A has nothing more to say;
(b) B thinks that she or he has got the message and that there is no need for A to elaborate;
(c) B wants to speak up at a particular point in the on-going talk, before it is too late.

                                (Stenström, 1994: 73-74)

[13]   A: ... well his manner appealed to me
        B: m
        A: y\es *((I . I))*
        B: *I can see that* if you d\idn't . get \on with it

                                            (LLC, S1. 5: 126-134)

[14]   Female: but I uh asked my physics professor if I couldn't chan*ge that*
        Male: *don't* touch that
        Female: what
        Male: I've got everything jus'how I want it in that notebook, you'll screw
                   it up leafin' through it like that

                                            (West & Zimmerman, 1983: 105)

Here in [13], B's "I can see that" overlaps with "I . I" in A's preceding turn.
Maybe B has got A's message before it is spoken and therefore manages to silence
A and take the turn. In [14], B interrupts A all of a sudden because he wants to
stop B from spoiling his notebook before it is too late.

As to the causes of other kinds of simultaneous speech such as backchannels,
Hatch (1992) stresses the speakers' psychological need of backchannels. She says

> When we begin a conversation, we expect that others wish to converse
> with us and will value what we have to say. In part, we judge this
> willingness in terms of backchannel signals. ... (Such) overlaps will let
> the speaker know that he or she is not talking to the wall --- everyone is
> participating.

                                            (Hatch, 1992: 49-53)

In summary, Sacks et al.'s approach is explained by the projectability of speech
which is based on those so-called "transition-relevance places"; Duncan's approach
emphasises the speakers' turn-signals which are composed of a set of linguistic and
non-linguistic cues; and Stenström's approach is more interested in the value of
information developing in a conversational discourse. All of them explain the
causes of simultaneous speech to some extent. Yet the problem is how such

transition-relevance places are signalled, represented by the speaker and recognised by the hearer, or how the hearer can know that the current speaker has no more to say. This problem leads to further examination of the linguistic context in which simultaneous speech is likely to occur.

2.3.2 THE LINGUISTIC ENVIRONMENT ASSOCIATED WITH SIMULTANEOUS SPEECH

With regard to the linguistic environment associated with simultaneous speech, previous research is mainly concerned with linguistic contexts in which interruption is likely to occur, yet the findings are interesting enough to extend to other kinds of simultaneous speech.

Sacks et al.'s study (1974) includes the syntactic context of overlap. It assumes that every speaking turn is realised by one or more turn-constructional units and the boundaries of these units are defined as having transition relevance. Constructional units are syntactic, where the range of unit types is demarcated by grammatical categorisation. They "turn out to be 'possible completion points' of sentences, clauses, phrases, and one-word constructions" (ibid: 721). In other words, it is the completion of any syntactic construction that provides hearers with an opportunity to take the turn. eg.

[15]   Old man: Th' Funfair changed it 'n *ah*ful lot **+didn't it+**
       Parky: ***Th-***   **+That -+**

                                                    (cited by Sacks et al., 1974:721)
Note:
The marker "+" here indicates a second case of simultaneous speaking which occurs immediately after the first case marked with "*".

[16]   A: we just came in from Alexandria, just got home and ***these winds***
        B: ***mm hm***
        A: were so bad we're gettin scared again heh
        B: no, ***we doh -***
        A:***and we*** wondered whether we should go to a motel or something
        B: no, you stay right where you are ...

                                   (cited by Sacks et al. 1974: 721)

Here in [15], the next two turn-starts made by Parky come at the first possible transition-relevance place (ie. a phrasal boundary) and the next possible transition-place (ie. a tag question boundary). In [16], A finishes two turn-constructional units before B comes in, A comes back in after a first-lexical-unit (ie. after B's word "no"), and B comes back in at the first possible completion of a first-sentential-unit (see Sacks et al., 1974: 721). Yet Sacks et al. do not describe and demonstrate those syntactic boundaries in detail; they do not define the concept of "possible completion point", neither do they take full account of those overlaps which happen to start at non-unit boundaries.

An adjacency pair is another syntactic construction which could elicit overlap in Sacks et al.'s turn-taking system. They considered that

> the turn-taking machinery includes as one component a set of procedures for organising the selection of 'next speakers', and, as another, a set of procedures for locating the occasions on which transition to a next speaker may or should occur. ... It is within any current utterance that possible next speaker selection is accomplished, and upon possible completion of any current utterance that such selection takes effect and transition to a next speaker becomes relevant.
> (Schegloff & Sacks, 1973: 293)

That being the case, adjacency pairs such as an exchange of "Bye-bye", the overall structural organisation such as a greeting exchange, a terminating exchange or asking a question; and the pre-closing markers such as "well, ok, so" (with

downward intonation contours) can all initiate a closing section or can be a transition-relevance place. This is the mechanism of turn-shifting on the current speaker's side.

Correspondingly on the next speaker's side,

> a recipient of some ongoing talk has the technical capacity to produce his talk with precision in relation to that ongoing talk, because a current speaker may position some objects within his utterance by reference to the recipient's capacity for and orientation to precision placement of his responsive talk.
>
> (Jefferson, 1973: 49)

[17]   (at the end of Jean and Mel's conversation)
Jean: thank you *Mel*
Mel: *Thank you* +Jean+
Jean: +Bye+*bye*
Mel: *Bye*

(cited by Jefferson, 1973: 49)

This is an example of an adjacency pair. The first two turns form a pair and the next two form another pair. When Jean says "thank you", she is producing the first part of a pair which Mel can take as a possible completing point and expect that Jean is going to stop after finishing it. So Mel overlaps Jean's turn and starts hers, which turns out to be another adjacency pair being overlapped by Jean in the same way. Jefferson calls this a case of precision timing in tag-position addressing terms. Let us consider another example.

[18]   Louise: No a Soshe is someone who *is a carbon copy of their
            friend*
       Roger: *drinks Peps*

(cited by Jefferson, 1973: 50)

This is an example of "collaborative sentence" in which Roger finds just the place to start up with a "completion" of Louise's as yet incomplete utterance, because

> a recipient of some sort of ongoing utterance has the technical capacity
> to select a precise spot to start his own talk "no later" than the exact
> appropriate moment to place it so that it will sound like a "continuation"
> of the prior/ongoing talk.
>
> (Jefferson, 1973: 51)

Here Jefferson emphasises the hearer's technical capacity of precise timing, but she does not go further to explore why hearers have such a capacity and how they actually make use of it.

Instead of examining the syntactic environment, Duncan and his colleagues (1972, 1973, 1974, 1977, 1981, 1985) are more interested in the vocal signals and gestures of the current speaker issued before the auditor's response. He identified a number of turn signals at the end of phonemic clauses such as :

(a) using any pitch level --- terminal juncture combination other than intermediate pitch level at the end of a phonemic clause;

(b) drawing out the final syllable or on the stressed syllable of a phonemic clause;

(c) using sociocentric sequences such as one of several stereotyped expressions like "but uh", "or something" or "you know" to tail off the turn;

(d) dropping in paralinguistic pitch and/or loudness in conjunction with one of the sociocentric sequences described above;

(e) completing a grammatical clause, involving a subject-predicate combination.

Duncan's studies show a very high correlation (.987) between the number of speaking turn signal displays and the auditors' response. Duncan mainly addresses the normal and smooth turn-taking signals. Further research is needed to explore the relation between speakers' turn-yielding signals and hearers' turn-bidding signals in particular, and this is a primary focus of the present research.

As to the placement of backchannels in conversation, Bublitz (1988: 183) suggests that it is subject to very few restrictions in terms of placement. It may be given during a "planning" and "breathing" pause of the current speaker or simultaneously with the current speaker's utterance. That is, it may occur at any position in an utterance. However, it occurs less frequently at the initial words in a speaker construction. According to Duncan (1972), the current speaker's within-turn signal may precede the auditor's backchannel. Such within-turn signals are either linguistic or paralinguistic: (a) a syntactic signal, ie. the completion of a grammatical clause. (b) a head direction, ie. the current speaker turns his or her head towards the auditor, from a previously "away" position. Although a body-language feature cannot be analysed in the present computer corpus, further research is needed to see if there are particular lexical or prosodic signals which could elicit backchannels.

In summary, the present study accepts Sacks et al.'s concept of a unit-type boundary, but goes further by extending it to a different level of language. It assumes there should be at least three levels of unit-type boundary --- prosodic, lexical (at a discoursal level) and grammatical --- existing in the current speaker's turn. Each of them displays signals or cues which serve as transition-relevance places for the hearers to start simultaneous speech. It is also assumed that the three levels of boundary markers combine together so as to create a stronger signal eliciting the hearers' response. The present study also addresses the problem of why some cases of simultaneous speech occur at non-unit-boundary places. It is expected that the concept of projectability and the complementary distribution of some boundary markers will help to answer this question. All this leads to a further exploration of linguistic features within simultaneous speech.

2.3.3   THE LINGUISTIC DEVICES AND STRATEGIES
ASSOCIATED WITH SIMULTANEOUS SPEECH

Previous studies of the linguistic devices and strategies used in simultaneous speech are mainly related to turn-bidding or turn-competing phenomena. In terms of prosodic devices, French & Local (1986: 158-178) address more basic questions of how interruptions are produced and responded to. They extend an analytic consideration to the ways that participants make use of prosodic features in the design of their overlapped speech and find that high pitch and high volume provide a powerful resource for the achievement of turn-competitive interruptions. eg.

[19]   A: ... we're trying to adapt it to **make it better***
        B: *plus:::::*** to keep her ...
            <h........h>
            <f........f>

                                        (cited by French & Local, 1986: 173)

[20]   P: I can have a fascinating conversation I mean it's a bonus I'm re**garded
              as***
        S: *I wonder - I wonder*** how they start these things up ...
            <h..............................h>
            <f..............................f>

                                        (cited by French & Local, 1986: 173)
Note:
<h...h> indicates a rise of pitch in the interrupter's overlap; <f...f> indicates an increase of loudness in the similar part of the overlap; ":::::" indicates the prolonged sound of [s] in "plus".

Similar findings also occur in studies by Duncan & Fiske (1977), Roger (1989), Ng et al. (1993) which all indicate that turn-competitive interruptions are prosodically marked, and that features like rapid speech rate and higher vocal amplitude are associated with successful interruption, because such features make turn transitions highly salient to participants and hence gaining the floor to speak is made easier.

In terms of lexical devices, attention has been paid to the initial words or phrases for interrupting speech. Schegloff (1987: 74) identifies some interruption markers such as "wait a minute" or "oh" which are used to announce that an interruption has thereby started. He also identifies some pre-placed appositionals, which he calls an "initial-to-the-turn". These include "well", "but", "so" "y' know", "yeah" etc. which can be placed at the beginning of a turn, when the second speaker begins to overlap so as to absorb the overlap with prior turns without impairing an actual turn beginning.

[21]   Ken: you know the new fad in uh in about seven years will be women
                smoking cigars you - because before it used to be all men were
                smoking cigarettes *and they* -
        Al: *well so*   +you'll be smoking a cigar in seven years+
        Ken: +waita minute+
        Al: *I don't care*
        Ken: *waita minute* they had a -

                                        (cited by Jefferson, 1973: 80)

Here Al's "well so" and Ken's "waita minute" are all lexical signals for a start of interruption. Although Schegloff considers these lexical signals as devices for minimising the gap and overlaps in conversation (because he regards overlaps and interruptions as abnormal), we can, from another angle, investigate those items as lexical or discourse devices which are used for turn-bidding.

From a discoursal point of view, Stenström (1994: 73-74) pointed out two primary speech acts frequently occurring in cases of taking a turn by interrupting. "Alerts" initiate interruption with words like "hey", "listen", "look" (usually in a separate tone unit and in a falling tone). "Metacomments" start by saying "Can I just tell ...", or "Could I halt you there..." or "Let me just ...", etc. All these can be discourse markers for initiating a turn-bidding speech act.

French & Local found that repetition is an important pragmatic strategy for making successful turn-competition. They said:

> there are ... non-prosodic features of talk which recur---more variably and sporadically - in the in-overlap portions of interruptions. In particular, it appears that interrupters utilise resources of sound sustention and repetitious incomplete syntax in displaying their turn-requirements from within overlap.
>
> (French & Local, 1986: 173)

If we examine [19] and [20] again, the interrupter sustains the final sound segment of the initial in-overlap component "plus:::::" and repeats the pattern of "I wonder- I wonder", which all indicate that he has something to say on the matter being addressed by the turn-occupant and that the saying of it is attendant upon the turn-occupant's completion. Such features will last until the interrupter has gained sole occupation of the floor. This is similar to Sacks et al.'s "repair" mechanism in dealing with abnormal turn-exchange cases such as mistiming the next attempt to start a turn or interruption. That is, after the simultaneous speech, either the interrupter or the interruptee may partially repeat their incomplete sentence structure so as to restart the turn. However, the repetition mentioned above only relates to those cases occurring in a self-repetition or self-repair and those occurring within or after the overlap. It is assumed in the present study that even those repetitions occurring at the initial part of the overlap, ie. the hearer's repetition of the current speaker's word or sentence pattern, can be significant in revealing the turn-taking mechanism.

As for the linguistic features of backchannels (or "hearer signals" in Bublitz's terms), Bublitz (1988: 184) has argued that:

(a) Prosodically, it is a characteristic feature of a hearer signal that its nuclear tone is lower than the nuclear tone of the preceding or simultaneously uttered syllable and that it is usually spoken with either falling or level intonation contour. Moreover, it is always spoken with reduced loudness, ie. at a lower volume than the preceding or simultaneous utterance and never heavily stressed.

(b) Lexically, hearer signals consist of short forms, usually monosyllables which constitute a closed paradigm consisting of "yes" and a number of variants like "yeah", "yep", "uhuh", "mhm" or "mm" or a few short forms such as "I see" or "Oh, really" (with falling nuclear tone).

[22]   B: b=e away from home thVen until at \any rate the \end of *-* about the
            end of \August --
       A: *=m*

                                              (cited by Bublitz, 1988: 176)

In [22], A is backchannelling during a pause made by B. The single syllable sound "m" is at level tone which is lower than the preceding stressed and falling-toned word "end" in B's utterance.

All the above findings are significant for further studies of simultaneous speech. They raise new issues such as how far such linguistic devices and strategies can extend to a larger scale data base of English conversation in different domains; how often they are used by participants of different status and gender; or if there are other important prosodic, discoursal features and pragmatic strategies used in simultaneous speech and how they coexist in conversation.

2.4 SOCIOLINGUISTIC FEATURES OF SIMULTANEOUS SPEECH

Some kinds of simultaneous speech such as interruptions have long been a disputed issue in social communication research. There have been many studies of

sociolinguistic and psycholinguistic aspects, particularly in terms of the association of simultaneous speech with gender, age, participants' relationship and culture.

2.4.1   INTERRUPTION AND GENDER

According to a comprehensive survey of the studies of interruption related to gender (James & Clarke, 1993: 230-280), the traditional assumption that males make more interruptions than females in mixed-gender conversation (in order to achieve male dominance and control) has been strongly challenged. This survey refers to research in which the great majority of observed interruptions have no distributional difference associating to gender. In some cases, they even occur more in female than in male conversation (see Kennedy & Camden: 1983, Tannen: 1984, 1989, Stubbe: 1991). These surveys also show that females' interruptions function differently from males'. The former seem more involved, supportive and cooperative. Yet there has not been any particular study made to find out whether male or female speakers use different linguistic devices and strategies in making interruptions and other kinds of simultaneous speech.

2.4.2   INTERRUPTION AND CONTEXT

It is normally accepted that the occurrence of interruption is associated with communicative settings. Edelsky (1981) studied interruption in a series of faculty committee meetings and found that task-based conversation (ie. report talk) is less likely to invite interruption than free conversation (ie. rapport talk). Coates (1989) showed that interruption occurs more in private conversation than in conversation in the public domain, which is similar to the findings of Tannen's research (1990) on family conversations. However, such comparative studies have not as yet extended to domains such as non-face-to-face conversation and public radio discussion, in which the stylistic differences of interruption could be significant.

2.4.3   INTERRUPTION AND POWER

The status or power of a person can be displayed by his or her dominance or control in speaking, and interruption is usually regarded as a reliable and objective indicator. Roger & Schumacher (1983: 700-705) tested 18 pairs of unacquainted speaker dyads in three conditions designed to reflect the complementarity of the dominance predispositions of partners: "high-low" (complementary) and "high-high" and "low-low" (noncomplementary). They found that the significantly higher rates of interruption for the "high-high" dyads emerged only after an initial period during which the three dyadic conditions did not differ from one another. That is, when speakers are still strangers to each other, there are constraints on interrupting, thus indicating that shared social norms governing turn-taking do indeed limit potentially offensive behaviour, such as interrupting in noncomplementary pairs during the initial stages of acquaintanceship. Ng et al. (1993) studied the interruption distribution in group conversation among unacquainted participants and found that speakers who have a greater proportion of aggregate speaking turns tend to have more interruption-turns in speaking, thus indicating that the manner in which speakership is gained contributes to an impression of influence. Moreover, Goldberg (1990: 900) identified a relationally neutral interruption outside the traditionally defined power-carrying continuum. He designated interruption as relational acts of neutrality, power, or rapport, and claims that interruptions are indicative of interpersonal relations other than dominance. His research encourages further study of interruption distribution among participants with different status and age.

2.4.4 SIMULTANEOUS SPEECH AND ETHNIC OR CULTURAL BACKGROUND

In the past few years, a number of cross-cultural studies have addressed the use of conversational mechanisms related to speakers with different cultural background. Their interest seems to centre on the contrast between western and oriental cultures. Results from three studies of conversation between native American English speakers and native Japanese speakers using English (Maynard: 1986, White: 1989, Murata: 1994) show that native Japanese speakers seem to make fewer interruptions than native English speakers, especially in terms of intrusive interruption. However, the fact that Japanese speakers make many more backchannels than English speakers may be due to: (a) the greater frequency of certain Japanese linguistic environments (eg. more short clause boundaries) that favour backchannels, (b) the Japanese honorific system (eg. more consideration in social status and intimacy of relationship), (c) the Japanese linguistic system (eg. less tolerance of redundancy), and (d) the lack of fluency in English of the Japanese participants. The study made by Tao and Thompson (1991), which compares backchannels in conversations between English-dominant Chinese and Mandarin-dominant Chinese, shows that the former type of Chinese speakers tend to make more backchannels (both in English and in Mandarin) than the latter type of Chinese speaker, thus indicating interference from the second (now dominant) language on the first.

The above findings reveal an interesting issue concerning the relationship between culture and language. Although some conversational mechanisms are universal, their actual applications are culturally specific. This is particularly the case when people are communicating in a foreign language. The present study explores this phenomenon by comparing the conversation of native English speakers' with that of native Chinese speakers in order to find out how the speakers' cultural norms or

the speakers' foreign language proficiency are transferring and interfering, either positively or negatively, in their use of simultaneous speaking devices in conversation.

2.5 SUMMARY

This chapter has reviewed a number of issues arising from earlier research. They include the classification of simultaneous speech, the pragmatic foundation, the linguistic environment and language devices and strategies associated with simultaneous speech. It also reviewed the literature about the association between simultaneous speech and its communicative context and the interlocutor's gender, status and cultural background. The present research will go further to see whether the findings of the previous research can be supported by quantitative data from a substantial corpus of English speech and whether such a large corpus can reveal new information about the simultaneous speech phenomenon as a whole and of the five categories of simultaneous speech in particular.

CHAPTER III

METHODS AND PROCEDURES

After reviewing formal and functional approaches to simultaneous speech in Chapter II, the present chapter gives an account of the methods and procedures designed for this study. First, the reason why the present research used a computer corpus as the source of data is explained. Then the English corpus used as the basis for linguistic analysis, the London-Lund Corpus of Spoken English (LLC) is introduced. Thirdly, the method used for compiling a subcorpus from the LLC, the coding system used to annotate this subcorpus, and the procedures used for data retrieval are outlined. Finally a complementary case study of Chinese conversation for comparison with the English corpus is described.

3.1 SOURCES OF DATA

3.1.1 A COMPUTER CORPUS AS A SOURCE OF DATA

Much of the research reviewed in Chapter 2 has been limited, either in terms of the data sources, or in terms of the method of analysis. Some qualitative studies (eg. French & Local: 1986, Tannen: 1990) have made detailed observations of a number of examples of interruption or focused on interruption in a single case study. Their findings have been based on authentic language data but are limited in size and lack quantitative results. Other experimental studies (eg. Duncan: 1972, Ng et al.,

1993) have obtained results supported by statistical analyses, but their findings are based on non-natural language data. That is, their data comes from recorded speech produced spontaneously, but the situation in which it was elicited was experimentally contrived and designed especially for producing the particular data the researchers wanted. Although quantitative studies are extremely valuable as a means of empirically testing a number of stereotypes and hypotheses and providing consistent evidence of clear trends, they can fail to capture adequately the complexities inherent in any analysis of interactional data. On the other hand, although qualitative analysis can provide a basis for developing appropriate classification systems, adding a valuable perspective to the interpretation of results, it is usually based on limited data.

Moreover, the studies mentioned above have usually concentrated on one kind of simultaneous speech (eg. interruption) or on one dimension of its linguistic features (eg. the amplification of the speaker's voice) or on one aspect of its function (eg. supportive and non-supportive interruption). All of this is necessary but there appears to be a lack of a comprehensive global study of simultaneous speech which regards it as an important part of the interlocutors' strategic behaviour. Nor have there been studies which explore the association between the teaching of foreign languages and the training of non-native language learners in effective communication. From an applied linguistic point of view, a study of a particular language phenomenon first of all needs to elicit a true and comprehensive picture of this phenomenon, then describe and generalise its patterns or features, explore their functions, and finally draw out implications for language teaching and learning. In this sense, a computerised corpus which can provide a potentially large source of data for analysis is particularly appropriate as a basis for research.

Nowadays, computerised corpora are playing a more and more important role in modern linguistic research because of their ability to provide huge amounts of data

for analysis with accuracy and speed. With the availability of computerised corpora since the 1960s, they have been recognised and explored as a source of systematically retrievable data, a testbed for linguistic hypotheses and a useful tool in applied linguistics. As Aijmer & Altenberg (1991: 3) point out, "corpus linguistics has developed into an important framework where description, model-building and practical application prosper side by side." The main reasons lie behind the use of a computer corpus in this study include its value as a source of data for linguistic description, the application arising from the description and its contribution to research methodology.

(1)   The Descriptive Function of Corpus Analysis

The advantages of using computer corpora for the description of a language are twofold: (a) to provide a large and possibly representative data source; (b) to describe a language phenomenon in a comprehensive way. In other words, in dealing with a language phenomenon, or a language item, a computer corpus analysis not only tells us what it is like and how it is constructed, but more importantly, it can also tell us how often it is used and how it is actually used in different contexts by different people.

Firstly, a systematic collection of authentic language data can provide a more realistic foundation for the study of language than earlier types of material, a fact which has given new impetus to descriptive studies of the English language in terms of prosody, lexis, syntax and discourse. It also provides a fruitful basis for comparing linguistic variation and stylistic properties of texts and genres. Since the availability of the two earliest computerised corpora, the Brown Corpus (Francis & Kucera: 1964) and the LOB Corpus (Johansson et al.:1978) collected respectively one million words of American and British written texts published in 1961, covering 15 different genres of written English, there have been many studies undertaken. In the 1980s, the first large publicly available computerised corpus of

spoken English, the London-Lund Corpus (Svartvik & Quirk: 1980), provided 500,000 prosodically transcribed words of spoken British English from the 1960s to 1980s, whose varieties ranged from formal to informal, private to public, monologue to multi-party discussion. In the 1990s an even larger spoken corpus has been compiled as part of the British National Corpus (BNC). The International Corpus of English (ICE), consisting of varieties of English from 19 major English speaking countries or regions in the world, with one million words of spoken and written texts in each variety, is under development. When finally completed, it will be extremely valuable for studies which compare the English used in different genres or in different regional or sociolinguistic varieties and will facilitate research on English across cultures as well (see Greenbaum: 1991; Kennedy: 1996).

Secondly, a computer corpus offers a chance to describe a language in many dimensions. As language is a system with different structural layers --- prosodic, lexical, syntactic, semantic, pragmatic and so on --- each layer can contribute its role in fulfilling a certain function. A comprehensive description of the features of different layers can reveal how these features are coordinated in fulfilling a language function. For this reason, any study with only one layer of language features can be relatively limited. However, computer corpora, especially those grammatically tagged or prosodically annotated, have opened up possibilities to address this problem. For example, in studying a number of lexical items peculiar to spoken English, Stenström (1990: 137-176) studied a discourse marker "really" as a starting point. She first looked for its text frequency in a large corpus both in spoken and written texts and found that a remarkable difference in distributional frequency of "really" makes it a salient feature of conversation. Then to explore its ecology in both corpora, she identified five functions of "really", some grammatical and some discoursal. Furthermore a concordance analysis revealed its different positions in the utterances or sentences and its different prosodic realisations associated with each of the grammatical or discoursal functions. All this contributes

to a full picture of "really", showing its multi-functions and multi-dimensional features in real use, thus deepening our understanding of the working of the language. Such descriptions have various applications including improving the design and content of the curriculum for language teaching purposes (see Sinclair & Renouf: 1988, Kjellmer: 1992 and Kennedy: 1992).

(2)   The Effect of Corpus Analysis on Research Methodology

The availability and use of computerised corpora have led to the development of more sophisticated research methodology. Many tasks which previously had to be done by hand can now be achieved automatically or semi-automatically by means of computer software, because no corpus, however large in size, could be of any use to a researcher until it could be accessed through a search and retrieval facility. In return, the development of such computer programs and software will contribute to the creation of probabilistic models of language and to the testing of theoretically motivated linguistic models.

As summarised by Lancashire (1995: 95):

> Researchers opt to use a corpus for the purpose of generalising about language behaviour. Unlike introspection or elicitation, linguistic corpora shared among researchers make it possible for them in public (and without having to be a native speaker) to verify all results, to turn to the same data source repeatedly for many kinds of language features, to compare studies of different features, and to analyse language across time and across register, tasks not well served by other methods.

Because of advantages such as these, a computer corpus approach has become important for research in modern linguistic study. According to the comprehensive ICAME bibliography from the 1960s to 1990s made by Altenberg (1991, 1993), over 1000 corpus-based studies have been undertaken on the bases of the published and unpublished computer corpora, covering various aspects of language such as

prosody, lexicon, syntax and discourse. To take the London-Lund Corpus of Spoken English (LLC) as an example, since it was first published in 1980 and supplemented in 1990, over 150 studies have been made with it. Some of the major ones have included Aijmer's (1984) research on some discourse items of spoken English, Biber's (1986) comparative study of spoken and written language, Altenberg's (1987) study on prosodic patterns and boosters in discourse and Stenström's (1994) latest published work on spoken interaction. Other studies have been made on such features as intonation, speech rate and turn-taking. Achievements in this area have greatly enriched our understanding of spoken language and enlarged the scope of research on aspects of discourse analysis, stylistics and grammar. Yet, in comparison with studies on written corpora, studies on spoken corpora are very small in number. There is still a great deal of room for further exploration. The present study analyses simultaneous speech in a subcorpus of the LLC on a larger scale than previously attempted. The methods and procedures used in data coding, retrieval and analysis will be discussed in the later part of this chapter.

3.1.2   AN INTRODUCTION TO THE LONDON-LUND CORPUS (LLC)

The present study has chosen the London-Lund Corpus of Spoken English (LLC) as the linguistic database. The LLC derives from two projects. The first is the Survey of English Usage (SEU) at University College London, launched in 1959 by Randolph Quirk. The second project is the Survey of Spoken English (SSE), which was started by Jan Svartvik at Lund University in 1975 as a sister project of the London Survey. Early in 1980, the first copies of the computerised version of the spoken part of the SEU became available as the London-Lund Corpus (LLC:o), consisting of 87 texts totalling some 435,000 words. Early in 1990, the completed version (LLC:c), with 13 more texts of the SEU added to LLC:o became publicly

available. The complete LLC consists of approximately 500,000 words of spoken British English recorded from 1953 to 1987, divided into 100 texts with approximately 5000 running words for each. The overall structure of the LLC is displayed in **Table 1**.

Table   1   Structure of the London-Lund Corpus

Dialogue (76 texts)	conversation (54 texts 6 non-surreptitious)	face to face (42 texts 118 male 68 female)	35 between equals 7 between disparates
		telephone (12 texts 100 male 202 female)	7 between equals 5 between disparates
	public discussion (22 texts all non-surreptitious 66 males 28 female)		13 between equals 9 between disparates
Monologue (24 texts)	spontaneous (17 texts)	all non-surreptitious 119 male 8 female	
	prepared (7 texts)		

Text: 100 texts in total, involving 242 extracts of conversation;

Setting of Recording: 52 non-surreptitious texts, 48 surreptitious texts;

Speakers' relationships: 55 conversation between equals, 21 between disparates;

Participants' gender: 403 male, 306 female, 709 speakers in total.

The reasons for using the LLC as the database for the present study are as follows:

(a) In the LLC different varieties of spoken British English are included from different speech domains and communicative settings. The collection involves monologues, dialogues and multi-party conversations and includes political speeches, sports commentaries, lectures, interviews, public discussions and private conversations. The settings involve face-to-face talk and non-face-to-face talk, both formal and informal. All these varieties provide a picture of spoken British English in use, which is broad enough for a wide range of studies. The more recent corpora such as the British National Corpus (BNC) use wider sociolinguistic coverage but were not yet available for the present research. When this study began, the LLC was quite simply the biggest available computerised corpus of prosodically analysed spoken English on which to undertake research on simultaneous speech.

(b) Conversations between equal, intimate speakers are distinguished from those between disparate, distant speakers, because "the gradient between intimacy and distance was found to affect very strikingly the kind of language used in conversation" (Svartvik & Quirk, 1980: 10-11). Moreover, every speaker's identity is marked out in terms of his or her age, gender, occupation or status, thus making it possible to involve sociolinguistic factors in research based on the corpus.

(c) Surreptitiously recorded and non-surreptitiously recorded data are differentiated, because "the presence of a microphone is important to move the speaker in the direction of careful formal speech, whether or not the matter was being broadcast or on a domestic machine" (Svartvik & Quirk, 1980: 11). Although the method of recording speech without the speakers being aware of it was acceptable in the 1960s and early 70s, such a method of surreptitious recording

is now rarely considered to conform to contemporary ethical standards in research on human subjects.

(d) Prosodic transcription is a particularly valuable characteristic of the LLC. About 100 prosodic and paralinguistic features have been marked in the full transcription version of the LLC (see Svartvik, 1990: 15); while the reduced transcription of the computerised LLC:c, ie. the version of LLC being used in the present study, still retains more than 50 prosodic and paralinguistic features, such as tone units (including the subdivision where necessary into subordinate tone units), onsets (the first prominent syllable in a tone unit), location of nuclei, direction of nuclear tones (falls, rises, levels, fall-rises, etc), boosters (ie. relative pitch levels), two degrees of pause (brief and unit pauses alone or in combination) and two degrees of stress (normal and heavy). Also indicated are speaker identity, simultaneous talk, contextual comments ("laughs", "coughs", "telephone rings", etc) and incomprehensible words (ie. where it is uncertain what is said in the recording). All these markers are particularly relevant to the present study in describing the multi-layer linguistic features of different categories of simultaneous speech.

The LLC was the biggest prosodically transcribed corpus until the mid 1990s and has been an unrivalled resource for the study of spoken English. Yet for the specific purpose of the present study, a further selection from this database was necessary.

3.1.3   A SELECTION OF A SUBCORPUS FROM THE LLC

As the present study concentrated on simultaneous speech, some data from the LLC was not relevant. Therefore a selection of texts had to be made to form a subcorpus subject to the following considerations:

(a) The subcorpus had to exclude all the monologues of the LLC but include dyadic conversations and multi-party conversations, because only in these conversations is it possible for simultaneous speech to occur.

(b) The dyadic conversations and multi-party conversations had to be further restricted to private conversation, group discussion, radio discussion and telephone conversation. Public interviews such as radio interviews had to be excluded, because such interviews are usually prepared or planned (at least on the side of the interviewer) and the "speaking turns" are likely to be allocated by the person who is in charge of the interview. An interaction in which the role of an interviewer is to ask while the role of an interviewee is to answer is less likely to invite naturalistic interruption.

(c) In order to obtain more authentic data and data which were fully prosodically transcribed, the subcorpus included in the present study used as much as possible those conversations which were surreptitiously recorded and prosodically transcribed. Yet authenticity is a matter of degree. The speaker's awareness of being recorded could reduce the degree of authenticity of the data. According to the compilers' description (see Svartvik & Quirk, 1980: 26), in conversations which involved both surreptitious and non-surreptitious speakers, the non-surreptitious speakers usually had been given the task of keeping the conversation going. Therefore, their utterances were not prosodically transcribed because of this special role, and their spoken words were not included by the original compilers in the total word count of the text. But in conversations such as radio discussion, in which all speakers were non-surreptitiously recorded, all of their utterances were still prosodically transcribed and their spoken words counted. Therefore the present subcorpus also included some radio discussions though they were non-surreptitiously recorded. In identifying simultaneous

speech, only those surreptitiously recorded or prosodically transcribed instances were taken into account so that we can analyse the prosodic devices used by the turn-bidders, turn-competitors and backchannel makers. However, there are still some instances in which the overlapped speech on the current speaker's side has not been prosodically transcribed because he or she is a non-surreptitious speaker. Therefore, when analysing the environmental features of simultaneous speech in terms of prosody, only those prosodically transcribed instances were taken into account.

(d) This study also addresses the distribution of simultaneous speech in different speech domains and among speakers with different degrees of familiarity, different status and gender. The subcorpus consists of 18 texts which are evenly distributed in three speech domains: casual conversation, public discussion and telephone conversation. Each domain consists of three texts from equal conversation and three other texts from disparate conversation. Within the category of disparate conversation, interlocutors are further classified as speakers with higher status and speakers with lower status. A further description of these classifications is in **3.3.4 (1)**. As to gender, although the LLC has more male speakers than female speakers, (402 male vs. 306 female), an attempt was made in the subcorpus selected for the present study to minimise such an imbalance to some degree. It includes the speech of 113 females and 80 males, among which 57 females are involved in female-to-female conversation (F/F); 26 males are involved in male-to-male conversation (M/M); 56 females and 54 males are involved in male-to-female conversation (M/F). As it was difficult to collect exactly equivalent samples for single- and mixed-gender conversation, the research counted the total word count of conversation among females (F/F), among males (M/M) and between males and females (M/F) respectively and used them for further analysis.

(e) In order to obtain data that is representative the subcorpus included speakers from as wide a range as was possible, not only in terms of age, gender and status, but also in terms of professions. As a result, the 193 speakers have various professions and occupations. They include university lecturers, secretaries, administrators, undergraduates, merchant bankers, housewives, research workers, publishers, politicians, journalists, editors, counsellors and social workers. Most of them are middle class professionals. This could be a limitation in the data source in the study, yet it is imposed by the constraints of the LLC itself.

With all the considerations above, the subcorpus which was extracted from the LLC was expected to be relevant, representative and suitable for the present research. A detailed description of the subcorpus used in the present study is presented in **Table 2.**

According to the LLC compilers' description (see Svartvik & Quirk, 1980: 11), each text in the corpus contains approximately 5000 words. In the present study, the total word count of each text was checked by computer so as to get accurate totals of words in terms of social variables such as speech domain, familiarity between interlocutors and speakers' status and gender. This was done for the further analysis of the distributional frequency of different categories of simultaneous speech in relation to different social variables (for detail see **3.3.4**). In the present study the total number of words only refers to the transcribed spoken words, including those subaudible words such as ((yes)), but excluding those untranscribable syllables such as ((sylls)), the text comments such as (laughs), and the sounds like "@" or "m" in [@m] or \ @m ! in the original corpus.

Table 2   <u>Subcorpus of the LLC Used in the Present Study</u>

==

Domain		Text	Words	Speakers	Males	Females
		S1. 2	5641	7	7	0
	equal	S1. 9	4968	3	2	1
		S2.10	5489	3	2	1
casual		3	16098	13	11	2
conversation		S3. 1	5080	6	3	3
	disparate	S3. 2	5006	6	4	2
		S3. 5	4950	4	4	0
		3	15036	16	11	5
subtotal		6	31134	29	22	7
		S5. 3	5358	3	3	0
	equal	S5. 4	6545	5	4	1
		S5. 6	4849	4	4	0
public		3	16752	12	11	1
discussion		S6. 4	4926	5	2	3
	disparate	S6. 5	5397	5	2	3
		S6. 8	4752	4	1	3
		3	15075	14	5	9
subtotal		6	31827	26	16	10
		S7. 1	4801	33	10	23
	equal	S7. 2	4816	24	7	17
		S8. 2	4756	21	3	18
telephone		3	14373	78	20	58
conversation		S9. 1	4916	29	10	19
	disparate	S9. 2	4621	25	8	17
		S9. 4	4931	6	4	2
		3	14468	60	22	38
subtotal		6	28841	138	42	96
TOTAL		18	91802	193	80	113

==

In summary, the 91802-word corpus shown in **Table 2** is a relatively large body
of text for the study of prosody. It is domain-specific and broad enough to

represent various linguistic features and patterns of different categories of simultaneous speech. The limitations of the database are as follows:

(a) The speakers come from quite a narrow range in the population, because most of them have an academic background and belong to the middle class in British society.

(b) Its classification of equal and disparate conversation is not very precise.

(c) Its identification of interlocutor's status, which is based mainly on professional ranking and age, could be rather subjective.

(d) It does not contain equal-sized samples of male and female speech, whether in single-gender or mixed-gender conversation.

However, these limitations are outweighed by its merits, being based on the fact that the LLC was the largest available corpus of spoken English and the only large one with the necessary prosodic transcription. The subcorpus has proved to be relevant to the purpose of this study. Findings based on it can be valuable, both for improving our understanding of the nature of English conversation and for English teaching and learning. In order to cope with the limitations of the corpus, an appropriate statistical method was chosen to interpret the resulting figures and a cautious interpretation was made of the results.

3.2 DATA CODIFICATION

Computer software makes possible very rapid searching and retrieval of items from a large corpus, but to make sorting and counting meaningful a great deal of manual analysis is still necessary, or as Leech (1991: 15) put it, "A successful analysis depends on a division of labour between the corpus and the human mind." In this study the interaction between the human brain and the machine was a repeated process of annotating the corpus and retrieving data. As a computer can only

recognise a coded form instead of its actual meaning, the coding system must be clearly defined so that the retrieved data can be valid for further analysis.

This section discusses a set of criteria for identifying different categories of simultaneous speech and for sorting out their environmental features, linguistic devices/strategies and sociolinguistic features. In addition to those prosodic markers which had already been marked in the original corpus (as shown on page xiv), it was necessary to provide additional coding for the present subcorpus with the system shown as follows, inserted at the initial point of the overlapped speech. For convenience, this additional coding is also given on page xv.

Additional Coding of the Subcorpus

	Current Speaker's Side	Overlapper's Side
Category of simultaneous speech		
STB (successful turn-bidding)	%	& ($ at the end)
UTB (unsuccessful turn-bidding)	|	||
BC (backchannelling)	>	>>
STC (successful turn-competing)		<<
UTC (unsuccessful turn-competing)		<
Identity of speakers		
speaker with higher status		H
speaker with lower status		L
male speaker		M
female speaker		F
Environmental features		
syntactic boundary	G	
word boundary	W	
address tag	A	
hesitation marker	H	
termination tag	T	
emphasizer	E	
connector	C	
Pragmatic strategy		
repetition		R

3.2.1   IDENTIFYING SIMULTANEOUS SPEECH

In the original LLC, all instances of simultaneous speech (or overlap) are marked by a pair of asterisks "*". They include overlapped spoken words as well as paralinguistic signals like laughs.or silent pauses. As there was no access to the original recording of the LLC, the present study only concentrates on those verbal or at least partially verbal cases of simultaneous speech. It excludes the following non-verbal overlapped cases as follows:

(a) non-verbal on both sides:

      A: ....*(laughs)*......
      B: *(laughs)*

[1]   d: I I mean [@:m] on our own *(- laughs)*
       c: *(- laughs)*

                         (LLC, S2.10: 943-944)

Here, B and A's overlap is filled with laughter which does not reveal any linguistic feature. Therefore it is a non-verbal overlap and not involved in this study.

(b) non-verbal on one side:

      A:..............*......*
      B: *(laughs)*......

[2a]   C: thank you-*for the sherry*-
       A: *(laughs -)* a great pleasure Dai.
                        (LLC, S1.9: 547-549)

Here, A overlaps C's speech with laughs. Although A continues to speak after the overlap, it is not clear whether he is making turn-bidding by laughing or whether he just starts his verbal turn after C has finished his. Hence it is not included as verbal simultaneous speech.

```
A:...............*(laughs)*
B:              *..........*......
```

[2b] C: may I borrow your floor (. laughs) if you bring it back*(- laughs)*
      B: *yes . bring it back* in good nick

(LLC, S2.10: 903-905)

Here, B starts speaking and gains his turn when C is laughing, but it is not clear whether B has interrupted C's turn. Maybe C has already finished his turn before he laughs. Therefore, this case is not identified as verbal simultaneous speech because of its formal appearance. However if the overlap (no matter on which side) involves partially spoken words and partially paralinguistic signals, it is still regarded as an example of simultaneous speech. eg.

[2c] B: ... about that . *that has come about has it*
      A: *(laughs -) ((so#))* oh yes you you know ...

(LLC, S1.9: 586-587)

Here, A's overlapped speech is initiated by laughter but is followed by a spoken word "so", indicating A's turn-bidding overlap is partially paralinguistic and partially verbal.

As mentioned in Chapter II, this research classified all verbal simultaneous speech into five categories.

Simultaneous Speech
1. successful turn-bidding (STB)
2. unsuccessful turn-bidding (UTB)
3. successful turn-competing (STC)
4. unsuccessful turn-competing (UTC)
5. backchannelling (BC)

Here STB, UTB and BC all involve the next speaker overlapping the prior speaker. Therefore in this study, all prior speakers are called the current speakers while all the next speakers are called the overlappers. In cases of STC and UTC, there are two turn-competitors overlapping each other, hence they are both called overlappers. The terms "current speaker" and "overlapper" are used in the rest of the thesis, distinguishing the two sides (or two interlocutors) involved in simultaneous speech.

Based on all verbal cases of simultaneous speech identified from the subcorpus, additional annotation of each category of simultaneous speech was carried out. Formal criteria were mostly used but the annotation also took account of the content of overlap in cases of ambiguity caused by formal similarities.

3.2.2 CODING SUCCESSFUL TURN-BIDDING (STB)

STB is identified by three criteria:

  (a) The overlapper starts speaking when the current speaker has already spoken, no matter how long he or she has spoken.

  (b) The overlapper has a verbal overlap with the current speaker, no matter how long it lasts.

  (c) The overlapper continues to speak after the overlap, no matter how long he or she continues.

The canonical formula of STB is:

    A:.... *..... *
    B: *..... *.....

[3]   B: ...so I gathered from Alec on the phone this morning *and [@:]*
       A: *^and [@]* they ^they more or less . ^they more or less conc\eded...
                                               (LLC, S1.2: 154-161)

Here, A starts speaking while B is speaking and thus causes a verbal overlap with B. After the overlap, A continues while B stops in a typical example of STB. However, if further simultaneous speech occurs immediately after the first overlap ends, it does not show that the turn-bidder is successful because he or she has not actually gained a single speaking turn (see the definition of turn in **2.2.1**) eg:

```
A:.... *..... *
B:    *..... *+.... +
C:           +.... +.....
```

[4]   M: I ^know you're `very anxious ``n\ot to dis`cuss *`stages ton/ight*
       D: *n\o# but it ^isn't* +a `question of th\at +
       B: +((it's ^simply `coming `back to that we're))+ not c\ompetent
              to `judge

                                        (LLC, S5.6: 1115-1118)

Here, D has broken off M's speech but he himself cannot actually gain a single speaking turn, because he is immediately overlapped by B, the third speaker or the second overlapper. Hence D's overlap is not a successful turn-bidding case, but B's overlap can be regarded as STB.

Three non-alphabetical markers "**&, %, $**" were used to code all the instances of STB after they were identified in the subcorpus. (See **Figure 3**) For example, [3] above is further coded as:

[5]   B: so I gathered from Alec on the phone this morning %*and [@:]*
       A: &*^and [@]*$ they ^they more or less . ^they more or less conc\eded
                                        (LLC, S1.2: 154-161)

Here "%" is put at the beginning of the current speaker's overlap, indicating the start of being overlapped; "&" is put at the beginning of the overlapper's speech, indicating the start of STB; "$" is put at the end of the overlap, indicating the end

of simultaneous speech by the overlapper, who, of course then continues speaking on his own.

3.2.3   CODING UNSUCCESSFUL TURN-BIDDING (UTB)

UTB shares the first two formal criteria of STB and only differs in the last criterion, ie. after the simultaneous speech, the overlapper stops while the current speaker continues to speak. The canonical pattern of UTB is:

A:.... *..... *.....
B: *..... *

[6]   B: ...the way that Mallet pres\ented them with ``^every [posib] *you
           ^kn=ow#   when ^Mallet gets the*bit between his . t\eeth# .he
           ^brings in `every ...
       A: *((^/\oh ^this is just 2 to 3 sylls))*

                                          (LLC, S1.2: 686-691)

Here A's speech overlaps current speaker B's speech, but A cannot gain the floor because it is B who continues to speak after the overlap. Therefore it is a typical instance of UTB. If the current speaker only has a silent pause during the overlap, but it is obviously a pause within the speaking turn, it is still categorised as UTB even though the overlapped part on the side of the current speaker is non-verbal. eg:

[7]   B: ((^whereas [m m m m])) . *. * ^my ^my point of v/iew# ((5 sylls..
       A: *((^well))*

                                          (LLC, S1.2: 569-570)

Here, A overlaps B's brief pause with a stressed "well" which is usually a turn-initiating marker. This indicates that A wants to say something. But B does not

relinquish the floor after the overlap and continues the turn with a repeated and emphatic use of "my point"; thus A's turn-bidding is regarded as UTB.

The non-alphabetical markers "ǁ" and "ǀ" were used to code all instances of UTB in the corpus, for example, [6] above is coded as:

[8]   B: ...the way that Mallet pres\ented them with``^every [posib] ǀ**you**
          ^kn=ow# when ^Mallet gets the* bit between his . t\eeth# . he
          ^brings in ``every ...
     A: ǁ***((^/\oh ^this is just 2 to 3 sylls))***

                                   (LLC, S1.2: 686-691)

Here, "ǀ" is put at the beginning of the current speaker's overlap, indicating the start of being overlapped while "ǁ" is put at the beginning of the overlapper's speech, indicating the start of an unsuccessful turn-bidding.

3.2.4   CODING BACKCHANNELLING (BC)

BC shares the same formal criteria as UTB and differs only in that the content of the overlapper's speech is always composed of a single sound such as "m", "mhm", "uhuh", "aha", "yes", "yeah", "yep", "ok" or very short phrases like "I see", "All right". eg:

[9]   B: [@] ^what [dhi dhi] the **final {\inwardness}*** \is# ^with ((with)) ...
     A: ***^[\m]# - ^[\m]#***

                                   (LLC, S1.2: 637-641)

Here, A's two "m"s are given as backchannels to B's turn, indicating A understands or agrees with B's speech, but is not interrupting B's speech. It is therefore an instance of BC. As has been mentioned in **3.2.3**, if the current speaker's overlapped part is composed of a silent pause but is obviously a short pause within the speaking turn, a brief feedback given by the overlapper in this

interval is still categorised as BC, even though the overlapped part on the current speaker's side is non-verbal. eg:

[10]   A: that ^meeting of the executive comm\ittee# ((3 to 4 sylls)) * - * and I
              ^r\ang you# ** - ** on the ^way to the /airport# - and ...
       B: *^y\es#*
       B: **^y\es#**

                                    (LLC, S1.2: 2-8)

Sometimes, a case of BC may appear as follows:
[11]   a: I'll try to get through them before the *((sylls))*
       C: *^y\es# that's ^r/ight# ^y\es#*
       B: I . ^wVould I `think# if I ^weren't ex``\aminer `here# ((6 to 8 sylls))
                                    (LLC, S1.9: 110-116)

Here neither the current speaker "a" nor the overlapper C continues to speak after the overlap, but C's words are obviously a backchannel to speaker "a", and are therefore still categorised as BC.

The coding of BC uses the non-alphabetical markers ">>" and ">" in the present study. For example, [9] above is further coded as:

[12]   B: [@] ^what [dhi dhi] the >*final {\inwardness}* \is# ^with ((with)) .
       A: >>*^[\m]# - ^[\m]#*
                                    (LLC, S1.2: 637-640)

Here, ">" is put at the beginning of the current speaker's overlap, indicating the start of being overlapped; ">>" is put at the beginning of the overlapper's speech, indicating the start of BC.

3.2.5 CODING SUCCESSFUL AND UNSUCCESSFUL TURN-COMPETING (STC & UTC)

STC is identified according to three criteria:

(a) The self-selected next speaker starts speaking after the prior speaker has finished his or her turn.

(b) The self-selected next speaker has a verbal overlap with another self-selected next speaker, no matter how long it lasts. In other words, there are two speakers starting to compete for the next turn at the same time.

(c) The self-selected speaker continues to speak after the overlap ends, no matter how long he or she continues.

UTC shares the first two formal criteria of STC and only differs in the last criterion, ie. after the overlap, the unsuccessful turn-competitor stops speaking, while the successful competitor continues.

The canonical pattern of STC and UTC is:

    A: ....

    B: *..........*

    C: *..........*............

[13]   C: I can recommend you a book
       B: *^it's a ``form of an``\analysis my ``dear#*
       A: *^if you were . ^if your were* if ^you were if ^you were sort of
             un\edu`cated#

                    (LLC, S2.10: 351-353)

Here, both B and A start speaking after the current speaker C has finished his turn. They are competing for the next turn at the same time. Yet A is identified as a successful turn-competitor because she continues to speak after the overlap ends while B is identified as an unsuccessful turn-competitor because he stops after the overlap ends. Here, the content of the UTC overlap is not taken into account,

because even when the UTC ends at a completed utterance boundary, it is difficult to judge whether the unsuccessful turn-bidder has actually finished his or her turn during the overlap or whether he or she still wants to speak after the overlap.

Turn-competing usually occurs in multi-party conversation, but occasionally it may occur in a two-party conversation, provided that the prior speaker has obviously finished the speaking turn and there is an obvious pause before the next turn starts. eg:

[14]   (B: ...but I ^haven't `brought it wVith me to`day#. / (- * - * laughs)
        A: *^o\k#* -)
        B: +^so+
        A: +^w\ell#+ . / ^I'll look `forward to `seeing you a`bout s\ix# ...
                                  (LLC, S7.3d: 208-213)

Here, both speakers B and A start speaking at the same time after they finish their prior turns respectively and then B laughs and A has a unit pause. In other words, B and A are competing for a new turn after the laughs and pause. Therefore it is also regarded as an instance of TC in which B's overlap is UTC, while A's overlap is STC.

The non-alphabetical coding for the beginning of STC and UTC are "<<" and "<". For example, [13] above is further coded as follows:

[15]   C: I can recommend you a book
        B: <*^it's   a ``form of an``\alysis my ``dear#*
        A: <<*^if you were . ^if your were* if ^you were if ^you were sort
            of - un\edu`cated#
                                  (LLC, S2.10: 351-353)

Here "<" in B's speech signals the start of UTC while "<<" in A's speech signals the start of STC.

In summary, the identification and codification of different categories of simultaneous speech were mainly based on formal criteria. The rationale for adopting these formal criteria was that firstly, it made both annotation and retrieval easier to carry out on a large linguistic corpus by computer. The present study is a quantitative analysis of a corpus of over 90,000 words which involves over two thousand cases of simultaneous speech. Even if there are a few instances with alternative interpretations or classifications possible, the majority of instances indicate a tendency for speakers to use some features and devices more than others. Secondly, as the researcher did not have access to the recording tapes of the corpus and there was no possibility of recovering the intentions of the original speakers, it is difficult to judge whether a turn is actually finished, or whether one speaker is actually intending to interrupt the other. Some previous researchers such as Roger et al. (1988) have tried to distinguish a successful interruptor from an unsuccessful interruptor by examining whether or not a speech ends in a complete clause boundary. However, the present researcher considers that even a complete clause cannot indicate the completion of a turn if there is no audio or video evidence. In such a situation, it is better to adopt a formal criterion to distinguish a successful turn-bidder from an unsuccessful turn-bidder. That is, to see if he or she continues to speak after the overlap, and to regard this as a sign of gaining the floor successfully.

Codification of the subcorpus using a set of non-alphabetical markers (eg. "&, %, $, |, ||, >, >>, <, <<") to code different categories of simultaneous speech was designed to obviate confusion in computer retrieval. Such markers highlight the beginning and ending of each simultaneous speech instance, thus serving as sign posts. They turned out to be very useful in retrieving the linguistic features in later analysis. More detailed discussion on the data retrieval is in the following section.

3.3 DATA RETRIEVAL

After the subcorpus had been coded, concordances, wordlists and statistics were produced by the corpus analysis software --- the Oxford Concordance Program (OCP). The data retrieval went through four stages:

(a) retrieval of overall distribution of different categories of simultaneous speech.

(b) retrieval of environmental features which occur immediately before instances of simultaneous speech in each category.

(c) retrieval of linguistic devices and strategies which are used by the overlappers in different categories of simultaneous speech.

(d) retrieval of sociolinguistic features associated with different categories of simultaneous speech.

Within each stage, the procedures were: first, identification of relevant features; second, coding the features; and third, sorting the features by computer.

3.3.1 THE RETRIEVAL OF INSTANCES OF SIMULTANEOUS SPEECH

First, all instances of the five categories of simultaneous speech in the subcorpus were retrieved by the computer as they had been coded with non-alphabetical markers (see **3.2.2** to **3.2.5**). For example, by searching for "**&, ||, >>, <<**" as respective key words, the computer can retrieve all tokens of STB, UTB, BC, STC in concordance files, thus displaying all examples of each category of simultaneous speech on the overlapper's side, including the linguistic context before and after the overlap. Similarly, by searching for "**%, |, >, <**" as respective key words, the computer can retrieve all tokens of the overlapped part on the current speaker's side (or the other turn-competitor's side in a case of turn-

competing), including the linguistic context before and after the overlap. The concordance files present each instance of simultaneous speech on a horizontal line as shown in [16] below.

[16]
1 2   720 B   ^since I was th\ere ((then# / %*^/is it#))*   // &*^that's*$ ^y\es# / that's...

[16] is an example of the output from this search procedure. This is a token of STB in context, as obtained in a concordance for "&". The first six columns, ie. "1 2 720 B", contain the text reference. The rest of the line is a concordance text of the STB. In the reference part, "1   2" means the data is from Sample 1, Text 2 of the LLC. "720" is the tone unit number in which the STB occurs. "B" is the turn-bidder, ie. the speaker who utters those words immediately after "&". In the text part, the speech is divided by a "//" in the middle. On its right is B, the overlapper's speech. On its left is A, the current speaker's speech, which is the linguistic context in which B's turn-bidding actually occurs.

When all simultaneous speech instances are displayed in such a format, a further analysis can be made on the re-edited files for various linguistic features. To take [16] as an example, if using "%" as a key word for retrieval, OCP can sort out the environmental features in the part just before "%" (see "^since I was th\ere ((then# /" in [16]), such as the last tone unit structure, the last lexical word or phrase and the last grammatical structure immediately before "%". If using "%" and "&" as key words for retrieval, OCP can sort out linguistic features within a text starting from "%" and ending at "&" (see "%*^/is it#))* // &*" in [16]). This part of text includes the linguistic features in the overlapped part on the current speaker's side such as the occurrence of unclearly-heard cases (which may indicate the diminishing amplitude) and the last tone unit structure before "&". Similarly, if using "&" and "$" as key words for retrieval, OCP can sort out features within a

text starting from "&" and ending in "$" (see "&*^that's*$" in [16]) which include the linguistic features in the overlappers' simultaneous speech such as the prosodic features and the lexical words immediately after "&", the last tone unit structure immediately before "$" and the word and sentence pattern immediately after "$". Briefly, OCP can produce wordlists or concordances or statistics for any defined items.

However, not all features of simultaneous speech can be retrieved in the concordance format as in the STB case above. In STB, the turn-bidder's beginning word always overlaps at least the last word of the current speaker's turn, therefore the concordance file can represent the chronological order of the STB context as it is in the original LLC. But in some instances of UTB or BC, their original format in the corpus is as follows:

[17a]   B: ^all this >*[@] -* ((4 to 5 sylls)) . we'd
        A: >>*^y\es# ^y\es#*
        B: ((^already)) h/\ad the meeting#

$$\text{(LLC, S1.2: 139-141)}$$

[18a]   B: ^what was she d\oing / ^w\/orking -- /
        B: >>*^\oh# / ^quite g\ood#*
        c: being nanny - >*to - . *an English family who ...

$$\text{(LLC, S2.10: 878-882)}$$

If using ">>" as a key word, the computer can only make a concordance of the above two examples as follows:

[17b]
1 2   150   A   ^all this >*[@] -* ((4 to 5 sylls)) . we'd //   >>*^y\es#
^y\es#*/ ((^already)) h^ad the meeting#

[18b]

2 10   882   B ^what was she d\oing# / ^w\/orking#--// >>*^\oh# / ^quite
g\ood#* / being nanny >*to - *an English family who ...

In order to retrieve the end feature of the overlapped part on the current speaker's

side in [17b] and [18b], OCP defines ">>" as the key word in the concordance,

including words just before ">>", ie. "((4 to 5 sylls))" in [17b] and "^what was

she d\oing# / ^w\/orking#--" in [18b]. Yet these are not valid data, because they are

not the exact words at the end of the overlapped part on the current speaker's side.

Therefore, a further editing was needed in such BC instances so that the overlapped

speech of the current speaker is aligned to the overlapper's initial word. With the

help of the pre-coded marker ">", the overlapped part on the current speaker's side

had to be moved to the front of the overlapper's speech so that the format was

changed as follows:

[17c]

1 2   150 A ^all this >*[@] -* // >>*^y\es# ^y\es#* / ((^already)) h/\ad
the meeting#

[18c]

2 10   882 B being nanny - >*to - .* // >>*^\oh# / ^quite g\ood#* /

Only when all the simultaneous speech instances were displayed in this way was it

possible for the computer to retrieve the end feature of the overlapped part in the

current speaker's speech. Therefore, some further editing had to be undertaken at

this stage before retrieving the environmental features and linguistic devices

associated with UTC and BC.

3.3.2 THE RETRIEVAL OF ENVIRONMENTAL FEATURES ASSOCIATED WITH SIMULTANEOUS SPEECH

As part of the study, the linguistic features or patterns before the overlapped part on the current speaker's side were analysed. They are the features occurring immediately before "%, I, >", forming a linguistic environment in which STB, UTB and BC occur. In the case of turn-competing (ie., STC and UTC) it always occurs after the prior speaker finishes a turn and therefore the overlapped speech is just between two turn-competitors but not with the prior speaker's. Therefore the present study does not analyse the environmental features associated with STC and UTC as it does for the other three categories of simultaneous speech such as STB, UTB and BC, because turn-competing is not inserting a speech act into the current speaker's turn and the context in which turn-competing occurs is not the same as that for turn-bidding and backchannelling.

In analysing the linguistic environment in which simultaneous speech is likely to occur, what Sacks et al. called "the possible completion points of the turn-constructional unit" is a key concept. Sacks et al. (1974: 721) have described such a "unit type" from a syntactic aspect. Duncan (1973: 32) has described the prosodic boundary signals of a "phonemic clause" which is relevant to a "turn-constructional unit" in Sacks et al.'s terms. In this study, the identification of a turn-constructional unit boundary follows Sacks et al.'s syntactic criteria and Duncan's prosodic criteria, but goes into more details to seek significant boundary signals. It also takes account of Stenström's (1990) model of discourse markers as criteria to identify lexical signals of the unit boundaries.

(1) **Prosodic Context**

Here the analysis of the prosodic context for simultaneous speech focuses on the tone unit (TU). As Ford and Thompson (1992: 27) point out:

Intonational completion almost always involves grammatical completion and semantic completion; hence intonation units are a major component of convergence points, and therefore of the turn-taking system itself. Projecting when a new turn could start must centrally involve the perception of intonation units and pitch peaks within intonation units.

Duncan (1973: 32) described a TU as a phonemic clause which is marked by a drop of pitch or a decrease of loudness and density, or by any nuclear patterns other than level tone, or by a drawl on the last syllable or on the last stressed syllable of a clause. In the present study, the prosodic boundary markers can be:

(a) a tone unit (TU) ending marker "#",

(b) an appearance of a nuclear tone marker, no matter whether it is falling (\,
Λ,) or rising (/, V), or level (=),

(c) a silent pause marker (such as ".", " -", "--", " ---")

(d) a reduction in amplitude, ie. speech sounds marked with "(( ))".

As all of these features had been marked in the original LLC version, no further coding was necessary. OCP was used to identify the following items occurring immediately before or almost immediately before "%", "|" and ">". They include :

(a) all words marked with "#";

(b) all words marked with "\" (including "Λ") or "/"(including "V") or " = ";

(c) all silent pauses such as ".", "-" and " --";

(d) all the cases which are marked by double brackets such as **((words))** or **((sylls))**.

The following are some examples of STB:
[19]
3 1c 13950 A: {I ^l\iked that}very **m\uch#** / %*((5 to 6 sylls))* // &*apart

[20]
1 2   13060 A: ...what committee are you **t\alking** %*about# ((^/Arthur#))* //

[21]

3 1c 10690 A:.../ ^w\orking in the d/ay# . / %*and ^trying to `study* //&*

[22]

1 2 710 A: ..../^since I was th\ere ((**then**# / %*^/is it#))* // &*^that's * ...

[23]

1 2 1220 B:...ch/um((s))# / this is ((**2 to 3 sylls** / %*4 to 5 sylls))* //&*^...

Here in [19], the prosodic context immediately before the STB, ie. before "%" is the word "much" which has a tone unit (TU) boundary marker "#", indicating that STB occurs at a TU boundary. In [20] the word "talking" has a falling nuclear marker "\", which is in the last TU just before "%", indicating that STB occurs at a TU with a nucleus. [21] has the word "day" followed by a brief pause "." before "%", indicating that STB occurs as the current speaker is making a silent pause. [22] has a bracketed "((then))" before "%", indicating that STB occurs when the current speaker's words happen to be unclear. [23] has a case of "((sylls))" before "%", indicating that STB occurs when the current speaker's volume declines and the words become incomprehensible. If the above markers occur immediately **after** "%" or just **before** "&", they indicate that the overlapped part on the current speaker's side is **initiated** or **ended** by the same features.

(2)   **Lexical Context** .

Lexically, words or phrases like "well", "you see", "you know", "ok", which are described by Stenström (1990, 1994) as discourse devices for organising a conversation, can also be signals for shifting a turn or developing a topic and may occur at unit boundaries. In this study, the discourse markers related to simultaneous speech are classified into five sets as follows:

(a) Address tags, including tag question, naming and addressing phrases such as "..., isn't it", "..., are you", "..., didn't he", "you know", "you see";

(b) Hesitation markers, such as "@:", "@m", "m", "well", or a repeated word like "it it it";

(c) Termination tags, such as "... so on and so forth", "... things like that", "..., please", "..., thank you", "'... as well", "... or so";

(d) Emphasizers, such as "yes", "no", "right", "mhm" and "that's right" (all in falling tone);

(e) Connectors, such as "and", "or", "but", "since", "because", "so".

Discourse items which occur immediately before simultaneous speech were identified and coded with the capital letters A (set a), H (set b), T (set c), E (set d) and C (set e) (the use of capital letters for coding could make such features salient and thus easier for computer sorting). Then all such cases in different categories of simultaneous speech were retrieved. Here are some examples from the STB file.

[24]
8 2a 4900 B: [tsh] tom\orrow `then# ^/is it# - A%*[@:] - ^yeah* // &*a...

[25]
2 10 13080 C: very hVard# / ^\isn't it# / ^[=m]# - / H%*((2 to 3 sylls))*

[26]
5 3 6310 w: ...conditions and `so on and s\o forth# / T%*[b@]* //&*but

[27]
8 2a 6730 C: ... can than `we can ``s/\ave# / ^y/eah# . / E%*^that's tr/ue...

[28]
3 2a 4580 A: ...actually# / ^y\ah# . / [@:m] - ^=and# . C%*^what ((you'...

These examples show that STB may occur just after the current speaker produces a tag question such as "...isn't it" in [24]; or a filled pause such as "m" in [25]; or a

terminating tag such as "...so on and so forth" in [26]; or an emphatic marker such as "yeah" in [27]; or a connecting word such as "and" in [28].

(3)   **Grammatical Context**

Grammatically, unit boundaries can be defined as " 'possible completion points' of sentences, clauses, phrases, and one-word constructions." (Sacks et al., 1974: 721) As Sacks et al. do not give a detailed demonstration of such complete structures in terms of grammar, the present study classified such grammatical boundaries into two major categories: syntactic boundaries and word boundaries. The purpose was to test whether simultaneous speech is actually occurring at complete syntactic boundaries as was argued by Sacks et al. Here a syntactic boundary signals an independent syntactic unit. It is defined as occurring at the beginning or end of :

(a) a sentence or a clause,

(b) a prepositional, an adjectival, an adverbial or a noun phrase,

(c) a subject, a predicate or an object,

(d) a one-word unit utterance such as "Yes", "No" and "OK".

A word boundary signals an incomplete syntactic unit. It is defined as occurring after an item such as:

(a) a determiner, an adjective, or a preposition before a noun,

(b) an adverb or an auxiliary before a verb,

(c) a syllable which is not the last one within a word.

Using these boundary marking criteria, a further coding was made on the output files which involve the linguistic environment of simultaneous speech. Focus was laid on the structural pattern immediately before the overlapped part on the current speakers' side. A capitalised "G" was coded on all instances ending with a syntactic boundary and a capitalised "W" was coded on all instances ending with a

word boundary. Then the syntactic and word boundary cases coded in different categories of simultaneous speech were sorted out by computer. The following are some examples from the output file of STB:

[29]
1 9   40   A: **I've been** ``^l\onging to `see you# / G%*how ^\are you* //

[30]
1 2   710   A: .../ **^since I was th\ere** ((then# / G%*^/is it#))* //&*^that's.

[31]
1 9   1310   A: yes yes / and then **^in the /end G%*** of course# / ^people...

[32]
1 9   7820   A:...`know# /at the ^m\/oment# /**^y\es# / G%***I ^mean* //&*...

[33]
1 2   840   B: ...enough on that occ/asion# / **the W%*person*** // &*^Steven*...

[34]
1 2   1980   B: ...^y\es# . / ^th\at's true ((**in W%***/England#))* // &*^when...

[35]
2 10   9720   B: ...Vestion# / I mean ``^I ``**w\ould W%***((get the `urge to...

[36]
3 1b   9780   A: ... ^publi`cations **dep\art W%***ment# / ^[/mhm]#* ...

From [29] to [32] "G" signals that the instances of STB all happen to occur at a syntactic boundary of the current speaker's turn. They include a sentence boundary such as "I've been longing to see you" in [29], or a clause boundary such as "since I was there then" in [30], or a phrase boundary such as "in the end" in [31], or a one word unit such as "yes" in [32]. From [33] to [36] are instances of STB occurring at word boundaries marked by "W". In [33] there is a boundary between an article "the" and a noun "person". In [34] there is a boundary between a preposition "in" and a noun "England". In [35] there is a boundary between an auxiliary "would" and a verb "get". In [36] there is a boundary within a three-syllable word "department".

3.3.3 THE RETRIEVAL OF LINGUISTIC DEVICES AND STRATEGIES USED BY OVERLAPPERS

At this stage, the linguistic features and patterns associated with the overlapper's simultaneous speech were analysed, covering five categories of simultaneous speech: STB, UTB, BC, STC and UTC. The rationale for including STC and UTC at this stage is that in a turn-competing case, any self-selected next speaker is regarded as an overlapper of the other respective speaker (see **3.2.1**). Therefore it is appropriate that the linguistic features of both STC and UTC should be studied. The analysis of linguistic devices and strategies used by overlappers in simultaneous speech was undertaken in terms of prosodic devices, discourse items and repetition strategies.

(1) Prosodic Devices

The analysis of prosodic devices focused on the prosodic features occurring at the beginning, the middle and the end of the overlapper's simultaneous speech.

(a) In the beginning of the simultaneous speech, the prosodic features of the initial word are examined. They include stress pattern, nuclear pattern and the single-word TU structure. A single-word TU is a tone unit which consists of only one word with both a stress and a nuclear tone pattern, eg. / ^y\es /.

(b) During the simultaneous speech, the speech rates between the current speaker and the overlapper during the simultaneous speech are compared. The speech rate within the overlapped speaking time is evaluated by counting the average number of words spoken by each side in the overlap. It is obtained by dividing the total number of spoken words within all the overlaps by the total tokens of simultaneous speech. Here the spoken words include both transcribed words and inaudible syllables. In dealing with the inaudible syllables (which are marked by "((sylls))"), the present study assumes that the average length of each word is about two syllables because the majority of the commonly used words

are monosyllables. (see Zipf: 1935, cited by Miller, 1963: 89) This is particularly the case in spoken English. As was reported in a study of word frequency in telephone conversation (see French, Carter & Koenig: 1930, cited by Miller, 1963: 91), more than 80% of the 800 most commonly used words are one or two syllable structures.. Therefore, two inaudible syllables such as "((2 sylls))" in the overlap are counted as one word which is added to the total word count of the overlaps. The assumption is that within a given time, ie. within the short interval from the beginning to the end of simultaneous speech, the one who speaks more words is considered to have a faster speech rate than the other who speaks fewer words. If the pauses occurring during the interval are also taken into account, it is further assumed that the one who speaks more words and has more pauses has an even faster speech rate than the other who speaks fewer words and has fewer pauses. Here the pauses include both silent pauses and paralinguistic sounds. Two brief silent pauses are counted as one unit pause, because according to the LLC compilers' description (see Svartvik & Quirk, 1980: 22; Crystal & Davy, 1969: 39), a unit pause "-" is about a stressed unit's length, ie. about the length of a long vowel, while a brief pause "." is equal to a short vowel, ie. half the length of a unit pause. The same method is used in counting the length of paralinguistic sounds like laughs and coughs whose length is also marked by "." or "-" in the LLC. Measuring speech rate is difficult if there is no actual recording accessible. Counting the number of syllables in a given time can be another way, yet syllables with different vowels are different in length as well. Therefore the present study counts the number of words instead of syllables in overlaps. This is only an approximate way of measuring speech rate, but when it is applied to more than two thousand cases of overlapped speech, it could reveal any tendency in speech rate on each side of the overlap.

(c) During the simultaneous speech, the amplitude of the current speaker and the overlapper is compared by counting the unclearly-heard cases initiating and

ending the overlaps on both sides. An unclearly-heard case means those uncertain word(s) and those incomprehensible syllable(s) which are shown by means of "(( ))" in the original LLC such as ((words)) and ((sylls)). It is assumed that amplitude is associated with the clarity of voice and those who have fewer unclearly-heard cases in the overlap are likely to display a higher degree of clarity and vice versa.

(d) At the end of the simultaneous speech, the last TU boundary structure and the nuclear pattern are examined on the overlapper's side.

As there was no need to re-edit or code the above prosodic features on the output file, the computer retrieval was the same as is described in **3.3.2**. The following are some of the retrieved examples of STB:

[37]
1 9 130 A: ...of the %*other department* // **&*^y\/es#** /^both*$ the...

[38]
3 5b 6830 A: %*- - - does he like / people* // **&*((^y=es#))** - - - / [@:m]$

[39]
6 8 610 a: ...%*((several sylls . laugh))* // **&*oh yes I ((1 to 2 sylls))*$**...

[40]
1 9 2080 A: ... /\always a `problem#))* // **& *yes it's ^h\orrible /actually#$**...

Here in [37], the bold type word "yes" shows that the initial word "yes" in the overlapper's simultaneous speech (marked by "&") has a stress, a falling-rising nuclear tone and forms a single-word TU as well, thus displaying three prosodic features within one word. In counting the total number of spoken words and unit pauses within the overlap on both sides in [38], the current speaker's side (which is marked by "%*......*") has three unit pauses ("- - -") and four words ("does",

"he", "like", "people") in total while the overlapper's side (which is marked by "&*......$*") has two words("yes", "@m") and three unit pauses ("- - -") in total. In sorting the unclearly-heard cases, [38] illustrates one instance "((yes))" at the beginning of the overlap while [39] displays another instance "((1 to 2 sylls))" at the end of the overlap. In counting the number of words of this inaudible case, "1 to 2 sylls" is counted as one word. In [40], the last word of the overlap "actually" is marked with both "#" and "$", indicating that this STB ends at a complete tone unit boundary. If the last TU before "$" has neither a nuclear tone pattern nor a TU boundary marker, it indicates that this STB ends with an incomplete tone unit or that this STB has an incomplete prosodic ending.

(2) Discourse Items

Here the analysis is concentrated on the most frequently occurring discourse items at the beginning of the overlapper's simultaneous speech. First, the computer software retrieved wordlists which contains a rank ordering of the frequency of all initial words, ie. all the words marked with "&", "ǁ", ">>", "<<" and "<" in respective output files obtained in **3.3.1**. Those words at the top of the list, ie. the most frequently occurring lexical items, were selected for further study in terms of their discourse position, their semantic meaning, their pragmatic function and prosodic features. In an initial analysis of the LLC subcorpus, the most frequently occurring initials were classified into five sets of discourse items. They are:

(a) initials of agreement such as "yes", "mhm";

(b) initials of disagreement such as "no";

(c) initials of hesitation such as "well", "@:";

(d) initials of exclamation such as "oh", "ah";

(e) initials of continuity such as "and", "but".

Detailed description of this classification is contained in **4.3.2**.

(3)   **Repetition Strategies**

Beyond the use of individual words as discourse items, what is notable in the literature and in the present data is the use of repetition of sequences of words by the overlappers in simultaneous speech. Here the term "repetition" mainly follows Persson's definition of "sequential repetition" which refers to:

> immediate repetitions of one or several identical lexical items produced by one speaker or writer in a continuous spoken or written sequence.
>
> (Persson, 1974: 11, cited by Bublitz, 1988: 228)

The present analysis of repetition does not only cover the cases in which the overlapper duplicates his or her own lexical words and those of the current speaker, but also includes the cases in which the overlapper repeats the grammatical form of the current speaker, ie. "the repetitious incomplete syntax" in French and Local's terms (1986: 173). The latter kind of repetition can be regarded as a grammatical device associated with simultaneous speech, but we prefer to take it as a pragmatic strategy because we will focus more on its pragmatic functions than on its forms. In the initial analysis of the LLC subcorpus mentioned above, repetitions were identified at the beginning of the overlap, during the overlap and immediately after the overlap. Therefore the present research classified repetitions into three types in terms of these three positions in the overlapped speech. They are:

(a) Initial repetition, in which the overlapper starts speaking by repeating the current speaker's last words or syntactic patterns which occur just before simultaneous speech. It also includes the case in which the overlapper starts speaking by continuing the current speaker's unfinished sentence pattern (see [41] [42] below).

(b) Mid repetition, in which the overlapper repeats his or her own words or syntactic pattern during the overlap (see [43]).

(c) Post repetition, in which the overlapper repeats his or her own words or
syntactic pattern immediately after the overlap ([see [44]).

In terms of the producer of the part being repeated, the first type can be called
"other-repetition" while the other two types can be called "self-repetition". Before
retrieval, these three types of repetition were all coded with a capitalised "R" and
reclassified either as "other-repetition" or "self-repetition" after they had been sorted
out by computer. The following are some examples from the STB file:

[41]
1 2   13920   A: [dhi]^**that's** [**dhi**] %*((.^wh\at do you [m] `call it#))* //
R&*^that's the "g\auleiters#* ...

[42]
7 3f   4860   A: ...how did `you . `**get on** - - %*sk\iing#* // **R&*^skViing#*$**
-/.^skiing was good...

[43]
1 9   13560   a: ...^walking/ `stick#* // **R&***yes **this is this is** * Meak to ...

[44]
2 10   5660   B: ...*((2 to 3 sylls))* // **R&*^oh I `thought*** ^**I thought** it...

Here, "**R**" signals a repetition in the overlappers' speech. [41] and [42] were
identified as cases of "other-repetition". In [41], overlapper A repeats current
speaker B's incomplete utterance "that's the", using it as a start for STB. In [42],
overlapper A does not actually repeat the current speaker's words or pattern, but
continues the current speaker's unfinished pattern "get on (skiing)". In a sense she
starts her turn by repeating the current speaker's syntactic frame. However, [43]
and [44] were identified as cases of "self-repetition". In [43] overlapper "a" repeats
his own words "this is" within the overlap while in [44] overlapper B repeats her
own words "I thought" immediately after the overlap.

3.3.4 THE RETRIEVAL OF SOCIOLINGUISTIC INFORMATION FROM THE CORPUS

In the previous stages of data retrieval (see from **3.3.1** to **3.3.3**), the frequency distribution of simultaneous speech, the environmental features and the linguistic devices and strategies associated with five categories of simultaneous speech (ie. STB, UTB, STC, UTC and BC) had been sorted out respectively. At this stage a further retrieval was made on each of these output files so as to see how these results were associated with social variables such as speech domain, degree of familiarity between interlocutors, and speakers' status and gender. Here the five categories of simultaneous speech were generalised into two major ones:

(a) category TB (turn-bidding) includes STB, UTB, STC and UTC, which are all aiming at bidding for a turn;

(b) category BC includes only backchannelling behaviour, which is not aiming at bidding for a turn.

The present researcher regarded these two categories of speech acts as the two major listener response strategies in terms of simultaneous speech. It was expected that a comparison of the distributional frequency, environmental features and linguistic devices and strategies used for turn-bidding and backchannelling, in terms of the above social variables, might provide more evidence of the social and cultural constraints on simultaneous speech in English conversation.

Firstly the frequency distribution of TB and BC in terms of social variables were retrieved. The social variables involve:

(a) three speech domains: casual conversation vs. public discussion vs. telephone conversation;

(b) two degrees of familiarity between interlocutors: equal conversation vs. disparate conversation;

(c) two degrees of relative status within the disparate conversation: speakers with higher social status vs. speakers with lower social status;

(d) two genders: male speakers vs. female speakers.

Then, the environmental features were investigated so as to find out if the above social variables could affect the hearers in choosing a unit boundary to make TB and BC. The major environmental features involved were:

(a) prosodic boundaries, ie. the TU which ends with a TU boundary marker and which occurs immediately before the overlap (see **3.3.2 (1)**).

(b) lexical boundaries, ie. the word or phrase which is one of the five sets of discourse items occurring immediately before the overlap (see **3.3.2 (2)**).

(c) syntactic boundaries, ie. the syntactic structure which is a complete sentence or a clause or a phrase or a one-word utterance and which occurs immediately before the overlap (see **3.3.2 (3)**).

Finally, the linguistic devices and strategies were examined so as to find out if the different social variables would affect the overlapper's preference for a particular linguistic device or strategy. The major devices and strategies involved were:

(a) prosodic devices, including stressed initial sounds at the beginning of the overlap, unclearly-heard cases during the overlap and incomplete tone unit structure at the end of the overlap (see **3.3.3 (1)**).

(b) discourse items, ie. the initial item of the overlap which is one of the five sets of discourse items (see **3.3.3 (2)**).

(c) repetition strategies, ie. the repetition occurring at the beginning, middle and end positions of the overlap (see **3.3.3 (3)**).

(1)   <u>**Coding and Retrieving TB and BC in Terms of Social Variables**</u>

The classification of speech domain, degree of familiarity between interlocutors, speaker's status and gender is based on the LLC compilers' descriptions in the original corpus.

In terms of the three speech domains, casual conversation refers to a face-to-face dialogue held in a private setting when the participants can see each other and can observe each other's reactions. Public discussion refers to a dialogue that is heard by an audience that does not participate in the dialogue; including radio discussions or panel discussions that have been broadcast. Telephone conversation is also privately held but the participants are not in the same place (see Greenbaum & Svartvik, 1990: 12).

In terms of two degrees of familiarity between interlocutors, the LLC compilers distinguished equal conversation from disparate conversation based on the perceived relationship between interlocutors. In equal conversation, participants regarded each other as on an intimate, equal footing while in disparate conversation, participants regarded each other as on a more distant footing (see Svartvik & Quirk, 1980: 10-11). For example, Text 1.2 in the LLC was classified by the compilers under the category of conversation between equals. Its participants were described as follows:

[45]   S1.2.   A: male academic, age c. 43
              B: male academic, age c. 42

(Svartvik & Quirk, 1980: 26)

Correspondingly, Text 3.1. was classified by the LLC compilers under the category of conversation between disparates and the participants were described as follows:

[46] S3.1 a: male academic, age c.40

A: female prospective undergraduate, age c. 20

B: male academic, age c.40

(ibid: 30)

As the LLC compilers have already categorised all dialogues in terms of face-to-face conversation, public discussion and telephone conversation and also subcategorised each dialogue as either "between equals" or "between disparates" (see Greenbaum & Svartvik, 1990: 20-40), the computer programme can firstly sort out tokens of simultaneous speech occurring in each of the three domains and then based on the output of each domain, it can further sort out tokens of simultaneous speech in equal conversation and disparate conversation within this domain. For example, retrieving the TB tokens from S1.2, S1.9, S2.10, S3.1, S3.2 and S3.5 gave the distribution of TB in the domain of casual conversation. Similarly, retrieving the TB tokens from only S1.2, S1.9 and S2.10 obtained its distribution in equal conversation. The rest of the TB tokens in this domain were distributed in disparate conversation (see **Table 2 in 3.1.2**).

The identification of interlocutor's status is based on the LLC compilers' description of each participant in the category of disparate conversation. In a dialogue between disparates, the one who is in a position with higher rank or power or occupation or age is identified as a high status speaker, the other interlocutor being thus identified as a low status speaker. For example, in [46] above, speaker "a" and B were further coded as speakers with higher status. A capitalised "H" was coded in front of the overlapped speech of speakers "a" and "B" respectively, assuming that they are higher in status and older in age in comparison with speaker A. Speaker A was marked as a lower status speaker by adding a capitalised "L" in front of her overlap, assuming that she is lower in status and younger in age in comparison with speakers "a" and "B". But this is only a subjective judgement, which should not be considered as absolute in the interpretation of results and

discussion, because one's dominance in conversation is not always associated with his or her status. The criteria for assigning a speaker as H or L may have involved other factors such as the personal relationship between the two speakers, the topic involved and the purpose of the talk. For example, in a specific speaking context, a speaker with higher status may be patient enough to listen to someone who is lower in status and seldom interrupt because he or she is aware of that speaker's embarrassment or shyness when speaking in front of a H speaker. On another occasion, an L speaker may become dominant in speaking and interrupt a H speaker quite often, because in that particular situation, he or she is in a better position to have a say than the H speaker is. For example, a young publisher may have power to approve or refuse to publish a book written by a distinguished scholar and therefore, although the academic may have higher status, that may not count for much in the particular interaction.

In the present research, the analysis of status was not carried out on those equal conversations, because according to the LLC compilers' description, interlocutors in this category do not have an obvious difference in their perceived relationship as do those in disparate conversation, ie., they are not greatly different in terms of professional ranks, social power, or age.

Finally the coding of an interlocutor's gender is also based on the LLC compilers' description. For example, in [46] above, speaker A, a 20-year-old girl is coded with a capitalised "F" (indicating female) while speaker B, the 40-year-old academic is coded with a capitalised "M" (indicating male).

After each speaker has been coded in terms of status and gender, the computer program can help the researcher to analyse the association between various features of simultaneous speech and these social variables.

(2) Measuring the Frequency of TB and BC in Terms of Social Variables

After the coding, the computer program retrieved all instances of TB and BC in terms of different social variables. Then a frequency count per thousand words was obtained. In previous research, the frequency distribution of a certain language phenomenon such as interruption was either obtained by counting its occurrences within a certain period of time or within a certain number of words spoken (see Duncan: 1972, West & Zimmerman: 1983; Ng et al.: 1993). In the present research, there was no access to the tape-recordings, and therefore the frequency of occurrence of TB and BC in a time unit could not be measured. However, the frequency could be obtained by dividing the total number of tokens of each category of simultaneous speech by the total word count, then converting this to a frequency per thousand words. For example, where the total word count in the domain of casual conversation is 31134 and the total number of tokens of TB in this domain is 557, the frequency per thousand words of TB in casual conversation is counted as follows:

$$557 \div 31134 \times 1000 = 17.9 \ /1000 \text{ words}$$

The same thing is done to obtain the TB frequency in the equal conversations and disparate conversations respectively.

However, to measure the frequency of TB by a particular kind of speaker within a conversation --- such as females, males, higher status speakers (H speakers) and lower status speakers (L speakers) --- an alternative method had to be adopted. West & Zimmerman's method (1983: 104) involved recording mixed-gender conversations for a defined period of time and then counting the total occurrences of female interruption and male interruption respectively. They assume that within a given time of conversation, those who make more interruptions will naturally have a higher frequency than those who make fewer interruptions in that given time.

However, Stubbe (1991: 66) regards such measurement as unreliable, because she considers that the frequency of one kind of speaker's interruptions should be obtained on the basis of the total number of words produced by another speaker whose speech is being interrupted. It should not be counted on the basis of the elapsed time or the total word count of both kinds of speakers in the joint conversation. Therefore, Stubbe divided the total occurrences of female interruption by the total word count of males in mixed-gender conversation, hence the female interruption frequency. Her method of measuring the interruption frequency accounts for one speaker's interruption frequency based on the total word count of another speaker whose speech is interrupted. Yet it does not account for one speaker's interruption frequency in relation to his or her own total word count, ie. how often his or her words are used for the purpose of interrupting the speech of another speaker.

The present study tried two methods for counting the TB and BC frequency in terms of status and gender. Method 1 follows Stubbe's method of measuring. It accounts for the TB and BC frequency per thousand words in relation to the total word count of the speakers whose speech has been overlapped. For example, where the total number of tokens of TB produced by H speakers is 205 and the word count of L speakers is 16232, the frequency per thousand words of TB by H speakers is counted as follows:

$$205 \div 16232 \times 1000 = 12.6/1000 \text{ words}$$

The same thing is done to obtain the BC frequency of H speakers. Here the focus is laid on the receiver of TB or BC to see how often the current speaker's turn is overlapped.

However, method 2 lays its focus on the producer of TB and BC, ie. to see how often the hearer is making overlaps on the current speaker's turn. It assumes that if

both kinds of speakers have the same word counts in the joint conversation, which indicates that they have similar amount of time of occupying the floor, they should have equal opportunities to make TB or BC and thus obtain approximately the same frequency of TB or BC as well. In other words, those who have a larger word count are likely to have a lower frequency of TB or BC, because they have already had more time to occupy the floor and thus do not necessarily make that frequent turn-bidding to strive for the floor. It may therefore be that the speaker with a smaller word count in the conversation would be likely to make more turn-bidding. Therefore, in method 2 the TB and BC frequency per thousand words in terms of status and gender was obtained by dividing the total tokens of TB and BC produced by one kind of speaker by the total word count of this kind of speaker. For example, where the total tokens of TB produced by H speakers are 205, and the total word count of H speakers is 28394, the frequency per thousand words of TB by H speakers is counted as follows:

$$205 \div 28394 \times 1000 = 7.2 \; / \; 1000 \; \text{words}$$

The same method is used to obtain the TB and BC frequency of different kinds of speakers in terms of status and gender. The results obtained by the two different methods of counting the frequency of TB and BC are compared in Chapter IV (see **4.4.3** and **4.4.4**) and the methodological implications are discussed.

In counting TB and BC frequency in terms of gender, the subcorpus was first separated in terms of single-gender and mixed-gender conversation. In single-gender conversations, the frequency of females' TB and BC was obtained by dividing the total tokens of females' TB and BC by the total word count of female-to-female conversation. For example, where the total word count of female-to-female conversation is 8058 and the total tokens of females' TB are 93, the frequency per thousand words of females' TB is counted as follows:

$$93 \div 8058 \times 1000 = 11.5 / 1000 \text{ words}$$

The same method was applied to male-to-male conversation in order to obtain the males' TB frequency and BC frequency. However, in mixed-gender conversation, the TB and BC frequency per thousand words in terms of gender was obtained by the two methods described above which were used to count the TB frequency of high and low status speakers.

(3) Comparing Environmental Features Associated with TB and BC in Terms of Social Variables

The purpose of comparing the features of the linguistic environment associated with TB and BC in terms of their social variables was to find out if there are any social constraints on hearers when they choose a time or place to make TB or BC. The present study investigated the prosodic, lexical and syntactic boundary markers occurring just before TB and BC in terms of the above social variables to see if there is any significant difference in frequency of occurrence. For example, if the results show that there is a significantly higher frequency of syntactic boundary markers occurring before the simultaneous speech in formal conversations rather than in informal conversations, it may suggest that a higher degree of formality in a speech domain can tend to lead hearers to choose more frequently a syntactic boundary to initiate TB and BC.

(4) Comparing the Linguistic Devices and Strategies Used by Overlappers in Terms of Social Variables

The purpose of comparing the linguistic devices and strategies used by overlappers in terms different social variables was to find out how the application of a particular device or strategy is constrained by the above social factors. Therefore, the retrieval of different linguistic devices and strategies was made directly from the instances of simultaneous speech which are associated with a particular social

variable. For example, to analyse the TB initial features associated with females, the following features were sorted out: the frequency of stressed initials, nuclear patterned initials, single-word TU initials and the five sets of initial discourse items which were marked with a capitalised "F" in the previous coding stage (no matter whether they occurred in female-to-female conversation or in female-to-male conversation). These were then compared with the same features in the corresponding males' TB.

3.3.5   SUMMARY

So far we have described the methods and procedures for analysing the LLC subcorpus to identify different categories of simultaneous speech, coding and retrieving various linguistic features. Although the annotation process was arduous and time consuming, the computer retrieval afterwards was speedy and accurate. Moreover, methods of formatting or ordering an analysis by means of a concordance helped stimulate further exploration of any undiscovered patterns or features which the present researcher had not earlier anticipated. However the laborious manual annotation of the subcorpus needed for the present study suggests that an improvement in corpus annotation and the development of more sophisticated software relevant to corpus analysis is still urgently needed.

In summary, the methodology described in the first stage (**3.3.1**) would address the first research question: how often do different categories of simultaneous speech occur. In the second stage (**3.3.2**), it would address the second research question: when and where are these simultaneous speech acts likely to occur. In the third stage (**3.3.3**), it would address the third research question: how does the person who is seeking to bid for a turn or to make a hearer's response actually behave. In the fourth stage (**3.3.4**), it would address the fourth research question: how are English simultaneous speech features and devices associated with social variables

such as speech domain, degree of familiarity between interlocutors, and speakers' status and gender.

3.4   CONVERSATION DATA FROM SPEAKERS OF CHINESE

One of the initial motivations of the present study was to answer the question: why do Chinese learners of English find it difficult to take a turn or to bid for a turn in English conversation.   It would therefore have been desirable to have a corresponding corpus of Chinese conversation so that a comparative analysis of the simultaneous speech phenomena in both languages was possible, but no such computerised spoken corpus of Chinese is yet available.   However, the present researcher has had access to the transcription of 12 hours of conversation recorded in Wellington by speakers of Chinese (Cantonese).   Although this data was collected for research into "Intergenerational Communication and Stereotypes of the Elderly", it included many examples of simultaneous speech which are significant for the present research.

3.4.1   DESCRIPTION OF THE CHINESE DATA

The recordings are cross-generational family conversations made by Chinese living in Wellington.   Four extracts of conversation from four different families were selected for the present study.   They are all casual conversations which occurred at home among elderly grandparents, middle-aged parents and young grandchildren. Each extract is ten minutes long and the total word count is 10385 words.   The data was transcribed orthographically with pausing and overlapping indicated but without a full prosodic transcription.   Therefore, it is in no way comparable with the LLC in terms of the total word count, the variety of speech domains,

participants' relationships, status and gender. For example, the Chinese data can all be categorised into one domain of casual conversation instead of covering other domains such as public discussion and telephone conversation; they are all equal conversations among closely related family members instead of being held among disparates; their participants consist of both adults and children and the latter were not included in the speakers in the LLC data; the family members have gender difference but they are involved only in mixed-gender conversation but not single-gender conversation as well. Moreover, these Chinese families have been living in New Zealand for at least 6 years and some family members are New Zealand-born Chinese. Their language behaviour may therefore have been influenced by the local culture or have had some deviation from that in China. Yet despite all these limitations, such data is valuable for exploring the nature of Chinese simultaneous speech, ie. to see whether the kinds of features and devices which occur in the LLC data are also present in the Chinese data. The results of such a Chinese case study can thus serve as a complementary analysis to the quantitative study of the LLC subcorpus to assist in the interpretation of the LLC data.

The following aspects of simultaneous speech in Chinese were examined.

(a) general frequency of five categories of simultaneous speech.

(b) environmental features occurring before simultaneous speech.

(c) linguistic devices and strategies used for simultaneous speech.

(d) sociolinguistic features associated with simultaneous speech.

Owing to the small sample of the Chinese data, the analyses in (c) and (d) are concentrated on turn-bidding speech acts. They cover only the instances of STB, UTB, STC and UTC, but not BC.

3.4.2 PROCEDURES USED IN ANALYSING SIMULTANEOUS SPEECH IN CHINESE

Since we had developed a set of criteria for analysing English simultaneous phenomena, the procedures for sorting and analysing Chinese simultaneous speech were generally the same as for the LLC data, only with some modification because of the characteristics of the Chinese language.

(1) Identifying and Classifying Simultaneous Speech in Chinese

The purpose of identifying and sorting simultaneous speech was to obtain a count of the overall frequency of simultaneous speech in the corpus of Chinese and the specific frequency in each of the five categories in particular. It would then be possible to find out if there is a similar trend of simultaneous speech in Chinese conversation and if each of its five categories has similar functions as in English conversation. Using the same formal criteria as were discussed in **2.2.2**, all instances of simultaneous speech were firstly identified. Then with the help of functional criteria (see **2.2.2**), the content of some instances was examined so as to subdivide the simultaneous speech into five categories: STB, UTB, STC, UTC and BC. This provided a foundation for a further analysis of the various features of Chinese simultaneous speech.

(2) Identifying Environmental Features Associated with Simultaneous Speech in Chinese

The purpose of identifying environmental features of Chinese TB and BC was to see if simultaneous speech in Chinese also occurs frequently at a unit boundary in the current speaker's turn. As the Chinese data was not prosodically transcribed, the focus was on syntactic and lexical boundary markers. In Chinese, a grammatical unit is defined as an independently used word, a phrase, a clause or a sentence. Therefore, a unit boundary can be signalled by a complete syntactic

structure such as a sentence, a clause, a phrase and a single-word unit as in English. eg.

[46]   ngoh wah ngoh di pengyauh daaih bouh fahn haih gongyahn

    I   said my   friend   most   be   worker

    (I said most of my friends were workers)

Note:
The Chinese (Cantonese) examples in this chapter are transcribed by the well known Yale system for Cantonese with a gloss for each word or a free translation in the brackets below or after them. As the tones of Cantonese are not pertinent to the present analysis, the nine Cantonese tone markers are omitted here.

Here the whole utterance "ngoh wah ngoh di pengyauh daaih bouh fahn haih gongyahn (I said most of my friends were workers)" can be regarded as a complete sentence. It involves clauses such as "ngoh wah (I said)" and "ngoh di pengyauh daaih bouh fahn haih gongyahn (most of my friends are workers)"; phrases such as "ngoh di pengyauh (my friends)" and "daaih bouh fahn (most)"; and independent words such as "pengyauh (friend)" and "gongyahn (worker)". In other words, a syntactic boundary can be identified at the beginning and end of each of the above units. But if an overlap occurs between the two morphemes such as "pengyauh" (friend) or "gongyahn" (worker), it cannot be counted as a syntactic boundary case.

Lexically, analysis was carried out in terms of the five sets of discourse items appearing just before the current speakers' speech is overlapped. They include address tags, hesitation markers, termination tags, emphasizers, connectors (for detail, see **3.3.2 (2)**). Examples of each of these include:

[47]   an addressing tag can be " ......mh haih meh?" (..., isn't it?);

       a hesitation marker can be: "@:", "mh" or "ni go ni go ni go" (it it it);

       a termination tag can be "......dengdeng" (...and so on) or "......jyu yu

                  chi leuih" (...something like that);

       an emphasizer can be "jauh haih!" (Right!);

       a connector can be "wak jeh" (or), "bìng cheh" (and) and "dàn haih" (but).

In spoken Cantonese, as in other varieties of spoken Chinese, there is a pervasive use of mood auxiliaries.. A Chinese mood auxiliary is a word or a combination of two words which have no equivalent in English. It always occurs at the end of a sentence, a clause, a phrase or a single-word unit. Though it is not a meaningful unit by itself, it helps to convey the speaker's varying mood, depending on the conversational context and the pitch and amplitude attached to it (see Liu Yuehua et al., 1983: 237-249). The major types of Cantonese mood auxiliaries are used as follows:

---- to mark an inquiry or a question such as "meh",

---- to mark an imperative, pleading, commanding, urging or persuading such as "la",

---- to mark an exclamation, surprise, or appreciation such as "wa",

---- to mark confirmation, emphasis, explanation or a reminder such as "ya ma",

---- to mark a pause, hesitation or to follow an example such as "a".

As the usual position of a mood auxiliary word or phrase is at the end of a meaningful unit, its appearance can be regarded as the termination of a Cantonese utterance. Therefore mood auxiliaries are also termed "sentence-final particles" by Li & Thompson (1981) or "bound forms" by Luke (1990). In the present study, such mood auxiliaries are identified as discourse items and categorised into the set of "termination tags", because they can signal a transition-relevance place for the hearer to start an overlap. .

(3) Identifying Linguistic Devices and Strategies Used for Turn-bidding in Chinese

The purpose of sorting linguistic devices and strategies used in Chinese turn-bidding was to see if native Chinese speakers share any similar strategies in making TB as the native English speakers do in the LLC data. Because of the lack of

prosodic transcription of the Chinese data, the focus here was again laid on the discourse items occurring at the beginning of TB and the repetition strategies used at the beginning, middle and end positions of TB. Identification and classification of these items used the same criteria as in the analysis of the LLC subcorpus (see **3.3.3 (2)** and **(3)**). The corresponding Cantonese for the five sets of discourse items can be as follows:

[48]   an initial of agreement can be "deuih" (right) or "haih ya" (yes);

an initial of disagreement can be "mh haih" (no);

an initial of hesitation can be "@::" or "gam::" (well);

an initial of continuity can be "yi cheh" (and) or "soh yi" (so);

an initial of exclamation can be "wa" (wow).

(4)   **Identifying Sociolinguistic Features of Turn-bidding in Chinese**

As the Chinese data does not cover as wide a range as the LLC subcorpus in terms of speech domains, degree of familiarity between interlocutors and speakers' status, the analysis of the social aspects of Chinese simultaneous speech was concentrated on the speakers' age and gender. It was structured with a view to find whether differences in age and gender could influence the Chinese speakers' frequency in making different categories of simultaneous speech.

In the orthographic transcription of the Chinese data, the identification of every speaker in the family conversation was annotated in terms of age and gender as follows:

GF = grandfather

GM = grandmother

F = middle-aged father

M = middle-aged mother

S = grandson of the grandparents

D = granddaughter of the grandparents

Based on the coding, the present researcher classified all the tokens of turn-bidding into three age groups and two gender groups. In terms of age, they have:

(a) same-generation type. That is, turn-bidding occurring between speakers within the same generation such as grandfather to grandmother, middle-aged father to middle-aged mother, and grandson to grand daughter.

(b) old-to-young type. That is, turn-bidding produced by speakers of the older generation to speakers of the younger generation such as the grandparents to the middle-aged parents or to their grandchildren, and the middle-aged parents to their children.

(c) young-to-old type. That is, turn-bidding produced by speakers of the younger generation to the speakers of the older generation such as the grandchildren to their parents or to their grandparents, and middle-aged parents to grandparents.

In terms of gender, as all the data are mixed-gender conversations and every speaker's gender was marked in their identity codes, all the tokens of TB were classified into:

(a) same-gender type --- turn-bidding occurring between interlocutors within the same gender such as female to female and male to male.

(b) female-to-male type --- turn-bidding produced by female to male.

(c) male-to-female type --- turn-bidding produced by male to female.

Since the Chinese data are all cross-generation and mixed-gender conversations, the study could indicate whether or not the difference in gender and in age would affect the frequency of turn-bidding, especially the frequency of successful and unsuccessful turn-bidding in the Chinese conversations.

3.4.3   SUMMARY

In summary, the coding and sorting of Chinese simultaneous speech were mainly a manual process because of the small database and the lack of a software package operating on the Chinese character data. The analysis is thus not comparable to the LLC subcorpus study. However, the manual sorting on a limited amount of data allowed an inspection of the content in more detail so that we may obtain some subtle usages and functions of the Chinese speaking strategies. The analysis of simultaneous speech in Chinese was designed to address the last research question of the present study, namely how the results of the LLC study could compare with simultaneous speech in Chinese conversation and how any difference could help account for difficulties experienced by Chinese learners of English. The results are discussed in Chapter V.

CHAPTER IV

RESULTS AND DISCUSSION OF ENGLISH DATA ANALYSIS

In this chapter the results of the analysis of the LLC spoken English data are presented and discussed in four sections: (a) the frequency of different categories of simultaneous speech; (b) the features of the linguistic environment associated with simultaneous speech; (c) the linguistic devices and strategies used by overlappers in different categories of simultaneous speech; (d) the sociolinguistic characteristics associated with different categories of simultaneous speech.

4.1 FREQUENCY OF DIFFERENT CATEGORIES OF SIMULTANEOUS SPEECH

Table 3 shows the relative frequency of occurrence and the frequency per thousand words of simultaneous speech in each of the five categories. First, the table shows that simultaneous speech is a rather frequent phenomenon in English conversation. It occurs 2011 times in the 91802-word subcorpus, or about 22 times for every thousand words. As Markel (1975:190) points out:

... Overtalk may be the listener's positive reinforcement to the speaker indicating that the listener is in tune with the speaker and is encouraging him to continue. Or, in some cultures switching overtalk may be the accepted means of obtaining a speaking turn.

Table 3   **Occurrences of Simultaneous Speech in the Subcorpus**

	Category	Tokens	%	Frequency/1000 w.
	STB	830	41.0	9.0
	UTB	333	17.0	3.6
Subtotal	TB	1163	58.0	12.7
	STC	71	3.5	0.8
	UTC	71	3.5	0.8
Subtotal	TC	142	7.0	1.6
	BC	706	35.0	7.7
Total		2011	100.0	21.9

Note:
As turn-competing (TC) involves two competitors; one is regarded as a successful turn-competitor and the other is an unsuccessful turn-competitor. Therefore one case of turn-competing is counted as two tokens of TC in this study; the former is STC and the latter is UTC.

Table 3 shows that instances of turn-competing (which include STC and UTC) occur less frequently than the other categories of simultaneous speech. It involves only 7% of the total tokens of the subcorpus, with a frequency of only 1.6 per thousand words. The possible explanation for this is firstly, turn-competing usually occurs in a multi-party conversational situation when the prior speaker finishes the turn without allocating the next turn to a specific speaker. In the subcorpus, there is quite a large amount of two-party conversation and this reduces the occurrence of TC in general. Secondly, in multi-party conversations, the communication is usually one-to-one instead of one-to-all; as Stenström (1994: 190) points out, "a multi-party talk tends to split into two-party talk", hence the lower frequency of TC. Thirdly, even in a one-to-all situation of multi-party

conversation where there is an absence of a prior speaker's appointment of the next speaker, the next self-selected speaker will be aware of the other participants' intention to speak (which may be conveyed by their eye contact, gestures or body movement) and will then decide whether he or she will compete with another self-selected speaker. This is especially the case in a more formal conversational setting which will be further discussed in **4.4**.

Turn-bidding (which includes STB and UTB) is the most frequent kind of simultaneous speech in the corpus, for it makes up 58% of the total tokens and has a frequency of 12.7 per thousand words. This is much more frequent than backchannelling, which makes up 35% of the total tokens and has a frequency of 7.7 per thousand words. This indicates that the participants use simultaneous speech for turn-bidding more often than backchannelling when they are in the position of a listener. Moreover among the 1163 TB tokens, the majority (71.4%) are successful turn-bidding cases. This further suggests that most of the turn-bidding speech acts are smoothly involved in the development of the conversation and cause no serious break down in communication. However, turn-bidding is often described in negative terms as if it is interrupting others' speech. It is regarded as a "trouble" in Sacks et al.'s turn-taking system (see Sacks et. al., 1974: 701) or as "not permissible" by Duncan (1973: 36) or as "disruptive" by West & Zimmerman (1983: 104). But we may ask why it nevertheless occurs so frequently in English conversation and how it can be carried out so smoothly in turn transactions? Such questions indicate a need to examine the linguistic environment in which TB and other categories of simultaneous speech are likely to occur.

4.2 THE LINGUISTIC ENVIRONMENT ASSOCIATED WITH SIMULTANEOUS SPEECH

In answering the question of when and where simultaneous speech is likely to occur, it was necessary to examine closely the features of the current speakers' language which occur before the overlappers' turn-bidding and backchannelling. Here the boundary of Sacks et al.'s "unit type" is a key concept. They point out:

> the allocation of turn-space is organised around the construction of talk in the turn. That organisation appears to key on one main feature of the construction of the talk in a turn --- namely, that whatever the units employed for the construction, and whatever the theoretical language employed to describe them, they still have points of possible unit completion, points which are projectable before their occurrence.
>
> (Sacks et al.: 1974: 720)

If Sacks' statement mainly refers to syntactic boundary signals, McLachlan's summary of those non-syntactic signals is a complement to it. He says:

> when a turn was not grammatically completed but ended with any of four additional cues, namely, a prolonged drawl on the final word or syllable, a sociocentric sequence (Duncan, 1972), a prolonged unfilled pause (typically see Beattie, 1977), and a generally falling pitch contour over the final clause (Cutler & Pearson, 1986), it was deemed to be complete.
>
> (McLachlan, 1991: 211)

The present study explores where such possible completion points (prosodic, lexical and syntactic) occur in the current speaker's turn and how they function as a transition-relevance place for simultaneous speech such as TB and BC.

4.2.1 PROSODIC CONTEXT

Prosodically speaking, a piece of spoken discourse can be divided into one or more tone or intonation units (TU). A full TU structure consists of a pre-head, a head (or an onset), a nucleus and a tail. (see Brazil et al.: 1980) In the present study the most relevant features being analysed are the TU boundaries (which are marked with "#" in the corpus by the LLC compilers) and the nuclear tone patterns which are marked with "\" (falling), "Λ" (rising-falling), "/" (rising), "V" (falling rising) and "=" (level) in the original corpus. A TU boundary occurs between two TUs (which are separated by "/" in the original corpus), each of which has at least one nuclear tone pattern. A nuclear tone pattern is the most essential part of a TU, for it carries the information focus and the most prominent prosodic features such as an obvious pitch change movement, a stressed amplitude, and a longer duration of the vowel in the stressed syllable of the nuclear word. Here, the important features occurring in the prosodic context immediately before STB, UTB and BC include:

(a) the appearance of the TU boundary marker,

(b) the appearance of the nuclear tone pattern,

(c) the appearance of a silent pause,

(d) the appearance of inaudible syllables and paralinguistic signals.

The results are presented in **Table 4**.

Table 4   Last Features of the TU Before STB, UTB & BC in the Corpus

Prosodic Features	STB		UTB		BC		Total	
	tokens	%	tokens	%	tokens	%	tokens	%
1.TU boundary	318	44.9	137	48.2	350	55.2	805	49.5
2.nucleus	414	58.4	160	56.3	418	65.9	992	61.0
3.pauses	193	23.3	83	24.9	134	19.0	410	21.9
4.inaudible syllables	4	0.5	2	0.6	2	0.3	8	0.4
5.paralinguistic sound	5	0.6	4	1.2	4	0.6	13	0.6

In **Table 4**, the percentage of TU boundary markers and nuclear patterns (feature 1 & 2) is based only on the prosodically transcribed cases on the current speakers' side, ie. 709 in STB, 284 in UTB, 634 in BC. However, the percentage of pauses, inaudible syllables and paralinguistic signals (features 3-5) is based on the total tokens of each category of simultaneous speech, whether they are prosodically transcribed or not, ie. 830 in STB, 333 in UTB and 706 in BC (for a detailed explanation, see **3.1.2 (c)**). Some tokens of STB, UTB or BC have more than one prosodic feature and therefore the percentages do not add up to 100%.

Moreover, feature 2 (ie. tone units with a nucleus) includes all those TUs which end with a TU boundary marker (see [1]) and all those TUs which have a nuclear pattern but are not yet completed by a TU boundary (see [2]).

[1]
1 2   3210   A: I think everyone `pr\esent#. / %*((of the ^British deleg\ation .
// &*((^no well)) I th=ink# . ...

[2]
1 9   1310   A: .../and then ^in the **/end** %*of course# / ^people `always d\o#*
// &*you ^kn\ow# .*$ / [@:m} - it ^has ...

Example [1] is an instance of STB occurring at a TU boundary of the current speaker's turn because the word "present" has a falling nuclear tone and is followed by "#" (which is a TU boundary marker) just before the STB starts. Example [2] is another instance of STB occurring at a place where the current speaker has uttered a nuclear patterned word "end" in rising tone, but has not yet finished a TU, because its tone tail "of course" (which is usually in a continuing but diminishing rising pitch) occurs just after the overlap begins. Therefore this STB in fact occurs at a place which is near a TU boundary.

In **Table 4**, feature 3 (ie. tone units with pauses) include all those instances in which the pauses occur immediately after a TU boundary marker (see [3]) and all those instances in which pauses occur at a non-TU boundary place (see [4]).

[3]
1 2   3210 A: I think everyone `pr\esent# . / %*((of the British deleg\ation . #
// &*((^no well)) I th=ink# . ...

[4]
1 2   1850 A: ... that's ^gone =on# / with^in [dhi] . %*((foundation its\elf#))*
// &*^inside ((Chirk its\elf#)) . / ^y\eah ...

Example [3] is an instance of STB which occurs at a TU boundary accompanied by a pause. Example [4] is another instance of STB which occurs after a pause but not at a TU boundary.

Table 4 shows that nearly one half (49.5%) of the simultaneous speech tokens occur at a moment when the current speaker has just finished a TU. In Duncan's terms (1972), a TU is a phonemic clause which is relevant to a grammatical unit. In Halliday's terms (1970), a TU is an information unit which is usually equal to a short clause. Therefore, a TU is a unit type whose boundary can be regarded as a transition-relevance place in one's speaking turn, thus making it possible for a hearer to enter the conversation. But how can a hearer know that the current speaker is coming to the end of a TU? As noted by Altenberg (1987: 47, cited by Svartvik, 1990: 73), a prosodic unit (which is also called a tone unit):

> (is) manifested as a coherent intonation contour optionally bounded by a pause and containing a salient pitch movement with a principal accent ('nucleus', 'tonic', 'main stress' etc.) normally occurring at the end of the unit.

Although in the LLC, the compilers said a tone unit is contour defined but not pause defined (see Svartvik, 1990: 74), the present research still identified the following three features as prosodic unit boundary signals:

(a) the occurrence of a word or words with a nuclear pattern, ie. an obvious change in pitch direction which is in contrast to that of the previous words.

(b) the occurrence of a nuclear tone pattern in a TU, a falling tone in particular.

(c) the occurrence of a silent pause which occurs immediately after a TU boundary or not.

As **Table 4** shows, 49.5% of the last TUs before simultaneous speech have a TU boundary marker which means that these TUs also have a nuclear tone pattern (see [1] above). 61% of the last TUs before simultaneous speech have a nuclear pattern, though some of them do not have a TU boundary marker. However, the appearance of a nuclear patterned word or words, no matter whether it is followed by a TU boundary marker or not, can itself serve as a boundary signal, because as Crystal & Davy (1969: 26) point out: "...the usual position of a nuclear is at or towards the end of a tone unit". For example in [2] above, the nuclear word "end" occurs only two syllables ahead of the TU boundary, indicating that the most important information of this TU has already been delivered. Therefore the appearance of a nucleus in a TU can be considered as being near a prosodic boundary.

A further examination of the nuclear patterns in the last TUs before simultaneous speech may find that a falling tone nucleus is a frequently occurring feature. **Table 5** shows that among the 992 nuclear patterned instances, about two thirds (64.7%) have a falling tone --- an intonation contour which usually conveys a tendency towards completeness and confirmation in the speakers' attitude (see O'Connor: 1967). The rising and level tones, which usually mean unsureness or incompleteness in a speaker's attitude, only occur in one third of the total nuclear

patterned instances. This indicates that simultaneous speech is likely to occur at the current speaker's tone unit boundary especially when its nuclear pattern is a falling tone. **Table 5** also shows that in the instances of turn-bidding such as STB and UTB, this tendency is even stronger, for they have a slightly higher proportion of falling tones than BC ($67.6 > 65.0 > 61.7$).

Table 5 **Intonation Contour of the Last TU before STB, UTB & BC**

Nuclear patterns	STB (414)		UTB (160)		BC (418)		Total (992)	
	tokens	%	tokens	%	tokens	%	tokens	%
Falling tone	280	67.6	104	65.0	258	61.7	642	64.7
Rising tone	125	30.2	54	33.8	154	36.8	333	33.6
Level tone	9	2.2	2	1.3	6	1.4	17	1.7
Total	414	100.0	160	100.0	418	100.0	992	100.0

Note:
The percentage for each type of nuclear pattern is based on the number of tokens (see the figures in the brackets on the top row) of which the TUs where the nucleus has already appeared before simultaneous speech begins.

The third signal of a TU boundary is a pause. **Table 4** shows more than one fifth (21.9%) of the simultaneous speech instances occur at the moment when the current speaker is making a silent pause, eg. [3] above is an instance marked with a brief pause "." after the TU boundary marker. (In other cases, a longer pause is marked with "-" or "--"). Pauses in such a place obviously emphasise the termination of a TU. However, a pause which does not appear at a TU boundary above can also be a possible boundary marker such as in [4], because it can signal the current speaker's hesitation or planning during the speech, thus offering a chance for the hearer to speak. Besides, as is shown in **Table 4**, the moment when the current speaker is making paralinguistic signals such as laughing or sighing or coughing, or when the speaker's voice is diminishing (ie. in the case marked by "((sylls))" which indicate inaudible syllables and can be interpreted as a symbol of lower

degree of clarity) can also provide a chance for a hearer's insertion, though these do not occur as frequently as the other prosodic features discussed above.

One conclusion which can be drawn from the above discussion is that simultaneous speech is most likely to occur at a full-prosodic-boundary (such as a TU with a nucleus and immediately followed by a TU boundary marker and a pause) or at a near-prosodic-boundary (such as a TU with a nucleus or a pause but without a TU boundary marker). The boundaries are usually signalled by a nuclear pattern, a falling nuclear tone in particular, a silent pause, or diminishing amplitude. They can be regarded as a possible completion point of a unit type, thus becoming a transition-relevance place for a new turn.

Generally the prosodic features discussed above occur before all categories of simultaneous speech, yet in specific categories of simultaneous speech some of these features occur more frequently than in others. **Table 4** shows that in terms of the occurrence of TU boundaries and nucleus, BC has the highest frequency. Although it is said that there is no restriction in BC occurrence (Bublitz, 1988: 183), the present study shows 55.2% of the BC in the subcorpus occur at a TU boundary and 65.9% of the BC occur after the appearance of a nuclear pattern. This is much more frequent than those in the categories of STB and UTB. An interpretation of this is that the hearer's brief feedback is usually issued **after** he or she has received new or important information but not **before**. A TU can be regarded as an information unit (see Halliday, 1989: 55) where the rheme (ie. the new information) always follows the theme (ie. the old or given information). This is consistent with the position of the nucleus which is usually at the latter part of a TU and imposed on the information focus --- the new or most important information.

Another feature which distinguishes STB from UTB or BC is that the TU structure at the end of the current speaker's overlapped part contrasts with the TU structure which occurs immediately before the current speaker's overlapped part. **Table 6** shows a comparison between the two:

Table 6   **Features of Last TU at the End of the Overlapped Part in Current Speaker's Turn**

Prosodic features	STB			UTB			BC		
	tokens	%	(%)	tokens	%	(%)	tokens	%	(%)
1.with TU boundary	483	68.1(44.9)		82	28.9(48.2)		146	23.0(55.2)	
2.with nucleus	433	61.1(58.4)		100	35.2(56.3)		165	26.0(65.9)	
3.inaudible syllables	84	10.1(0.5)		17	5.1(0.6)		17	2.4(0.3)	
4.overlap in pause	0	0 (0)		14	4.2(0)		125	17.7(0)	

Note:
(a) The percentage of each feature is counted in the same way as was done in **Table 4**.
(b) All the figures in the brackets are the corresponding figures in the last TU just before the current speakers' overlapped part, as have been shown in **Table 4**.
(c) Some tokens have more than one feature and therefore the percentages do not add up to 100%.

Table 6 shows that in the category of STB, the last TU at the end of the overlapped part in the current speakers' turn is much more likely to be completed than the TU structure occurring immediately before the current speaker's overlapped part, because it has a larger proportion of feature 1 and feature 2 (68.1 > 44.9, 61.1 > 58.4). This means that a substantial number of STB instances, though not occurring exactly at a prosodic boundary of the current speaker's turn, occur at a place which is very close to a TU boundary; eg. in [2] :

1 9   1310   A: .../and then ^in the /end %*of course# / ^people `always d\o#* // &*you ^kn\ow# .*$ / [@:m} - it ^has ...

Here, the TU on the current speaker's side is not completed before the STB starts (ie. before "%") but it ends immediately after the STB starts (ie. after "%"). One

explanation of the phenomenon may be that the turn-bidder in [2] may have got used to the prosodic pattern that a nucleus always occurs at or towards the end of a TU. He therefore anticipates a possible boundary when he notices the occurrence of the falling nuclear pattern on the word "end" and starts to bid for a turn. The current speaker's tendency to end the turn is also presented by a much more frequent occurrence of inaudible syllables at the end of the overlap than before the overlap (11.8% > 0.5%). This indicates that the speakers' voice is diminishing which is another signal of a possible completion point in a speaking turn. All these features show that the current speakers are coming either to the end of an utterance or to the end of a turn, thus signalling to the potential overlappers that it is appropriate to bid for a turn.

However in the category of UTB, there is a sharp decrease of completed TU structures at the end of the current speaker's overlapped part. **Table 6** shows that both feature 1 and 2 have a much smaller proportion (28.9 < 48.2, 35.2 < 56.3), accompanied by a much lower occurrence of inaudible syllables in comparison with the category of STB (5.1 < 10.1). This shows a clear indication of incompleteness of the current speaker's turn, which means that the current speaker still keeps the floor without giving many boundary signals, and hence the outcome is unsuccessful turn-bidding.

In the category of BC, the tendency towards incompleteness is even stronger. **Table 6** shows that both feature 1 and 2 have an even smaller proportion than before (23.0 < 55.2, 26.0 < 65.9). A plausible interpretation of this is that since an overlapper in a BC instance is not aiming at bidding for a turn when he or she is making a backchannel, the current speaker may simply maintain his or her turn which is not at all threatened by the hearer's backchannel. This is further supported by the fact that 17.7% of the instances of the current speaker's overlapped part are in fact silent pauses, which means that the current speaker accepts the hearer's

backchannel without uttering a word. The current speaker may assume, firstly, that the backchannel will not last long and secondly, he or she might be quite confident of continuing his or her turn after the backchannel ends. This is never the case in STB. **Table 6** shows no occurrence of feature 4 (ie. overlap in silent pause) in the situation of STB, but such a feature appears occasionally in the situation of UTB (4.2%). The reason why the current speakers occasionally keep quiet in the situation of UTB could be that they feel the overlappers' words are unlikely to be for turn-bidding, so they just let the overlappers finish speaking first, without competing with them, and then continue their own turns.

In summary, the prosodic context for simultaneous speech is concentrated on TU boundaries. That is, the appearance of a TU boundary marker with a nucleus (falling tone in particular) before it, and a silent pause after it can be regarded as an explicit prosodic boundary. However those TUs without a boundary marker but with a nucleus or a silent pause can be regarded as having an implicit prosodic boundary, because the turn bidder can make use of the contextual cues, such as the appearance of a nucleus, or a pause or diminishing amplitude, to anticipate the coming unit boundary and start turn-bidding or backchannelling. Further discussion on the anticipation mechanism is made in the next section of this chapter.

4.2.2   LEXICAL CONTEXT

One of the most striking results of the analysis of the lexical environment in which simultaneous speech occurs, is shown in **Table 7**. About 24% of the simultaneous speech instances occur at the time when the current speakers have just uttered lexical items or hedges such as "@" (equal to [ ], see **Figure 1** on page xiv), "m", "well", "and" , "but", "sort of" or phrases like "you see", "you know", and other tag questions (see examples from [5] to [35] in this chapter). These items cannot be adequately accounted for at the grammatical level but they become more

meaningful when being analysed at the discourse level in terms of speech-organisation, interactional and communicative devices. They have been called "interactional signals" and "discourse markers" by Stenström (1994: 61-63). The interactional signals are used to start, carry on and terminate the conversation; the discourse markers are used to organise and hold the turn and to make boundaries in the discourse (see Stenström, ibid.). In this part of the chapter, the emphasis is on how such signals and markers terminate utterances or make boundaries in the discourse, how they address hearers' responses and how different categories of simultaneous speech use different sets of these discourse items. In **Table 7**, these discourse items are classified into five sets and later each set is analysed in terms of its position, its prosodic contour and its pragmatic function.

Table 7   **Five Sets of Discourse Items Before STB, UTB & BC**

Function	STB (830)		UTB (333)		BC (706)		Total (1869)	
	tokens	%	tokens	%	tokens	%	tokens	%
Address tags	28	3.4	10	3.0	19	2.7	57	3.1
Hesitation markers	27	3.3	24	7.2	28	4.0	79	4.2
Termination tags	24	2.9	10	3.0	12	1.7	46	2.5
Emphasizers	106	12.8	41	12.3	55	7.8	202	10.8
Connectors	23	2.8	17	5.1	20	2.8	60	3.2
Total	208	25.1	102	30.6	134	19.0	444	23.8

Note:
The percentages of each feature is based on the total tokens of each category of simultaneous speech which are displayed in the brackets in the top row of the table. The same method is used in **Table 8** and **Table 9**.

(1)   Address Tags

Table 7 shows that 3.1% of the simultaneous speech tokens in the corpus occur after the current speakers produce address tags. Such tags include tag questions after statements and other words or phrases with a similar function to a tag question. eg.

[5]   B: ...**^darling** %*have some ch/erries#*
           d: &*have a hand*$ ful

                                      (LLC, S2.10: 1254-1255)

[6]   A: ...a``dvises them / on \options and `things **/are you#**. / |*or ar^ranges*
           B: ||*^w\ell# / ((^very V̇often#))* / .
           A: their . tut\orials# /

                                      (LLC, S1.9: 1005-1006)

[7]   A: ...it's be`cause he's / t\eaching **/isn't it#** - / |*((^up or ^not))* until
               `one
           B: ||*^well*
           A: o'cl\ock# /

                                      (LLC, S8.2f: 1132-1133)

[8]   B: I kn\ow# / ^I've been through this bef/ore# / it's **^all r/ight#** -
               /%*[@:m]*
           C: &*^bio*$ l\ogically you c/an# / ^\any `way# /

                                      (LLC, S8.2a: 244-250)

[9]   A: ...tell these g/uys# / ((that we'll)) ^carry \on# - / **^you s/ee#** / >*-*
               ^and
           B: **>>*^y\ep#***

                                    (LLC, S1.2: 170-177)

[10]   a: ...he's a very heavy man **you know** >*. Gooch I don't* know
              whether...
           A: >>*^y\es# / ^y\es#*

                                    (LLC, S1.9: 308-310)

In these examples, all the bold type words and phrases have the function of addressing the hearers' attention or response and expecting the hearers to share the current speakers' opinions. In [5], speaker B is addressing speaker "d", offering her some cherries. Speaker "d" gives an answer before speaker B makes the offer, suggesting that she may have heard the address and understood speaker B's offer, perhaps by body language, thus producing the overlapped speech. In [6], [7] and

[8], the overlapped speech occurs immediately after the current speakers' tag questions such as "are you", "isn't it" and "all right". They can be questions or requests demanding the hearers' answers, for they all end with a rising nucleus, an intonation pattern which is usually associated with interrogative utterances. As is pointed out by Stenström (1984: 224):

> Q-tags in final position have a stronger elicitative force than softeners (such as "you see", "you know") and are generally followed by a confirming reaction regardless of pitch contour.

Sometimes the current speaker may not really expect the hearer to make a verbal answer at the moment, yet the address-like tags such as "you see" and "you know" still make the hearer's response especially relevant for starting the next turn. In [9] and [10], the bold type phrases are not actually questions nor requests, but they can be used as a "sympathetic circularity sequence, just like a tag question" (Schourup, 1985: 72). Stenström (1990: 145) also points out:

> softeners (such as "you know", "you see") in final position are unique in that they serve as explicit appeals for feedback, especially if they carry a rising tone. They are therefore typically turn-yielding and interpersonally oriented.

Thus, the addressees in the above examples all take these tags as signals of a transition-relevance place and then make a backchannel as in [10], or they start bidding for the next turn, as in [5], [6], [7], [8] and [9].

(2) Hesitation Markers

Table 7 shows that 4.2% of the simultaneous speech tokens in the corpus occur just after the current speakers produce some hesitation markers which are also called "filled pauses" in Stenström's terms (1990: 215). Such markers include sounds such as "@", "@:m", "m:"; and words or phrases such as "well" and "sort

of". If we include those instances of silent pauses occurring before the current speakers' overlapped parts (which can also be regarded as non-lexical markers of hesitation or planning), hesitation markers before simultaneous speech make up more than a quarter (25.3%) of the total tokens of the three categories (ie. STB, UTB and BC, see **Table 8**). This suggests that a pause is a very salient signal in addressing the hearer's response. It is particularly frequent in the case of turn-bidding. The proportion of STB and UTB preceded by hesitation markers is 26.5% in the case of STB and 32.1% in the case of UTB (see **Table 8**).

Table 8   <u>**Hesitation Markers Before STB, UTB & BC**</u>

Type of pause	STB (830)		UTB (333)		BC (706)		Total (1869)	
	tokens	%	tokens	%	tokens	%	tokens	%
Silent pauses	193	23.3	83	24.9	134	19.0	410	21.9
Filled pauses	27	3.3	24	7.2	12	1.7	63	3.4
Total	220	26.5	107	32.1	146	20.7	473	25.3

Note:
The percentages are calculated using the same method used in **Table 7**.

As instances of silent pauses have been illustrated and discussed in **4.2.1**, the focus here is on the lexical hesitation markers. eg.

[11]   a: yes . yes . but **[@:m]** %* . ((5 to 6 sylls)) yes yes yes . yes*
       C: **&***come ^/on# / ^what's the ^what's the `dirt in `Rufford it*$``s/\elf# /.
                              (LLC, S1.9: 329-331)

[12]   A: ...``t\ime and# / **^w/ell#** / l*^I said* `to[@:m] . Steve . `and
          c/ompany# /
       B: ll*and . [ri^s@i]*
                              (LLC, S7.2k: 967-970)

[13]   B: ^not - [@:m] - wanting to `be **sort of** %*[@m] ((`2 to 3 sylls))#* /
       A: **&***^c\ourse not# / ^n\o# / you can be ^absolutely*$ b\usiness - `like# /
                              (LLC, S3.2a: 288-291)

[14] B: ...[dhi] - ^how far were y\ou# / [@:m] >*-* [@:m] ^b\anking on
        this...
      A: >>*((^ye g\ods#))* /

(LLC, S1.2: 851-852)

[15] C: but he ^doesn't . `live **in** `**in** `**in** . l*[@m]* / T\undraland at \all# /
      A: ll*^T\/undraland#*- /

(LLC, S1.9: 399-400)

All the words and phrases in bold type can be described as "fillers" when the
current speaker is planning what to say or looking for the right word, or just
wanting to keep the floor. Most of them are unstressed and accompanied by a silent
pause as in [11], [14] and [15] or followed by another filled pause as in [13] and
[14]. In [11] and [12] the hesitation markers [@:m] and "well" occur after the
connectors "but' and "and", indicating the current speakers have finished the prior
utterance and are planning what to say next. They use the filled pauses to keep the
floor, or maybe use these markers to relinquish the floor, or to make way for the
hearer to add ideas (see Schiffrin, 1987: 148). In [13], [14] and [15], the phrase
"sort of", the sound ''@:m" and the three repetitions of "it" are uttered within
unfinished utterances, suggesting that the current speakers hesitate in finding a
suitable word to continue. These hesitation markers serve the function of
addressing the hearer's response or signalling a termination of the utterance.

In a normal turn-taking system described by Sacks et al. (1974), it is unusual to
find more than one speaker at a time (or of course, no speaker at a time). When
such a situation occurs, a repair is necessary. The best repair for filling an
awkward pause is either for the hearer to give a backchannel to fill the "planning" or
"breathing" pause, or for the hearer to take a turn to speak. In this sense, turn-
bidding and backchannelling made at the moment of the current speaker's hesitation
may sound cooperative and supportive.

(3)   Termination Tags

Table 7 shows that 2.5% of the simultaneous speech instances in the corpus occur just after the current speakers have made a "terminating tag", in Aijmer's terms (1985: 117). They include the words and phrases whose usual position is at the end of an utterance. eg.

[16]   B: ...`possible to `sort of do **th\at sort of** `**thing**# / %*((2 to 3 sylls
               ^w\ould
        A: &*^c/\ourse it `will#*$ . / ^c/\ourse it `will#
                                                   (LLC, S3.2a: 477-480)

[17]   g: ...even - `with my own `books and r\ecords **and** `**so on**# / |*and ^not
                and...
        w: ||*aged ^fifty . aged ^fifty-s/ix# / ((that's))*
                                                   (LLC, S6.5: 69-73)

[18]   B: ...turned us d/\own# . / ^on `**something** `**like th\at**# / |*``^w\e* `like
               `them#...
        C: ||*^who*
                                                   (LLC, S8.2a: 80-82)

[19]   w: ...im`provement of `social con`ditions and `**so on and s\o forth**#
               %*[b@]*
        n: &*but*$ it's ^n/ot h\alting you# . /
                                                   (LLC, S5.3: 630-632)

[20]   A: so we ^don't know if they `taste nice **or n\ot**# - / >*((^d\o we#*)) /
        B: >>*^n\o#* /
                                                   (LLC, S2.10: 1447-1449)

[21]   A: ...rather exp\ensive `**I'm** / **afr/aid**# . / >*I'm ^prac*tically
               prepared to...
        B: >>*uh^/uh#*   (LLC, S7.3f: 704-706)

[22] h: our our our hy^gienic services **as w\ell#** / %*- ^shut the lot from
\every...
c: &*oh ^m\adam# / ^don't mention roads#*$ / ^don't mention r\oads ...

(LLC, S5.4: 990-994)

The phrases in bold type from [16] to [19] are all terminating tags which create a kind of vagueness. This is a characteristic of informal conversation in which social relations, sharing and intimacy are often more important than being explicit or precise. According to Aijmer's view, "one position where the speaker can choose to be vague is before a possible boundary in the discourse" (Aijmer, 1985: 117). These termination tags all appear at the end of the utterances, thus giving the hearer signals of transition-relevance places in the current speaker's turn. From [20] to [22], the bold type phrases may not be vague expressions, but they always occur at the end of an utterance, so their appearance also signals a unit boundary place. Some other phrases such as "'.., please.", "..., thank you." and "..., sorry." also have a similar location and discourse function.

(4) **Emphasizers**

Table 7 shows that 10.8% of the simultaneous speech instances in the corpus occur at the time when the current speaker has just finished a one-word emphatic utterance such as "yes", "right", "ok", "really", "good", "fine", "oh", "no". All those prosodically transcribed one-word utterances are in a separate TU and the majority of them (90%) have a falling tone. Thus a single word with a nuclear pattern and with a TU boundary marker may appear quite prominent in the context. A closer look at their meaning shows that most of these one-word units have a major function which is to display the speaker's positive confirmation of the current statement.

[23]   B: ^VI thought so# . / ^y\es# %*((2 to 3 sylls)*
       A: &*[@:] and *$ ^clearly there have been people in the Chirk Endowment
                                                    (LLC, S1.2: 474-477)

[24]   A ...I'm ^just not ``\interested e`nough# . / **^r\Veally#** / %* . to ^do th/at#*
       B: &*^and you know that*$ `there's a linguistic c\ontent ...
                                                    (LLC, S3.1a: 134-138)

[25]   b: **^g\ood#** / %*well ^n/ow# / ((^let's . go on to the n\ext `point#))* /
       m: &*of ^c\ourse# . / ^may I `just `ask you . *$ ^may I `ask `you \one ...
                                                    (LLC, S5.6: 750-754)

[26]   A: ^o/k `then# /
       B: r\ight# / I'll t\ell him# /
       A: **^r\ight# .** / |*^bye b/ye#* / ^b/ye# / ^he ...
       B: ||*^th\anks ((very)) `much#* /
                                                    (LLC, S7.21: 1225-1232)

[27]   B: ...the de^partment of Yiddish language and medieval l/iterature# . / ^and
            [dhi] div\Vision# / ^of contemporary Y\iddish# . / ^o\k# / **^f\ine#** /
            >*^this is the main ((b\edding#)) / and there's* pro^vison for / ...
       A: >>*^y\es# . / ^y\es# . / ^y\es#* /
                                                    (LLC, S1.2: 46-52)

Each of the words in bold type in the above examples can be called an "emphasizer", using Stenström's terms (1994). That is, they act as a reinforcer to the previous statements and then terminate the topics. For example, the falling tone "yes" in [23] and the falling-rising tone "really" in [24] express the speakers' re-confirmation of their previous ideas and indicate that they will add no more new ideas. The word "good" in [25] is a topic-shift marker. It concludes the previous topic and shifts to a new topic (which is reflected by the following words of "now let's go on to the next topic"). [26] is a closing of a telephone conversation where speaker A's "right" expresses her absolute agreement to speaker B's decision and indicates the point has been settled and it is time to end the talk. In [27], speaker B has gone through a long utterance and finally come to an end. He uses a falling

tone "ok" and "fine" to conclude his idea with such an emphasis that it is echoed by a confirmative response made by speaker A. The other one-word utterances such as "no" and "oh" could have a similar function, but are represented in a different way. eg.

[28] (b: I never heard of the thing
   m: I thought you were)
   b: ^n/\o# / %*^n\o# / ^n\o#* /
   m: &*oh I'm ^so s\orry#*$. / the ^f\ield [spo]# . /
                (LLC, S5.6: 1125-1131)

[29] (A: no it's OK)
   B: ^\oh# / %*^hang /on# . / ^I'll `go and g\et her#* /
   A: &*^[=m]# / ^O/K# - *$ / ^\oh# . / ^r\ight# ---
                (LLC, S9.1d: 203-211)

Here, the word "no" in [28] reiterates the negative statement which speaker "b" had made in the previous turn, indicating that he has no more to say on this point. The word "oh" in [29] is an initiating marker which indicates a new idea is coming. Both "no" and "oh" can be regarded as transitional points in speech. In summary, these emphasizers can serve the function of reinforcing the previous speech and terminating it, hence creating a context for the next speaker to start speaking.

(5) Connectors

Table 7 shows that 3.2% of the simultaneous speech tokens in the corpus occur just after the current speaker utters a connector, including conjunctions such as "and", "or", "but", "so", "however", "because", "if". eg.

[30] A: ...on the ^way to the /airport# - / **and** >*[@:m]* ^you were /out# /...
   B: >>*^that was /it#* /
                (LLC, S1.2: 5-9)

[31]   A: ...^Miss Str\ang's `now in`volved in it /all# / ^so . l*- ((2 to 3
          sylls))*...

    B: ll*(- laughs) ^{wh\o} \isn't#* /

                               (LLC, S7.2l: 1219-1221)

[32]   A: ... were being re``versed in Am\erica# / **or** >*where* A^merican was ...

    B: >>*^[=m]*

                               (LLC, S1.2b: 1236-1239)

[33]   B: ... w\ards {for ^s\o long#}# / ^t\oo# / **how^\ever#** / %*((3 to 4
          sylls))*

    A: **&***((2 to 3 sylls)) ^this is the *$ logical ...

                               (LLC, S1.2: 466-470)

[34]   w: ...`not en`tirely ``tr\ue {Mr Na^b\arro#}# . / **because** l*^what they
          ^\also...

    n: ll*((sylls))*

                               (LLC, S5.3: 298-300)

[35]   B: ...we'll get some `kind of d\ay`light# / (- laughs) **^even if** %*it's
          r\aining...

    C: **&***^y\es# / that*$ will `be . ^that ((would)) be g\ood# / ^y/\es# /

                               (LLC, S8.2g: 1235-1239)

A common function of connectors is to show that the current turn is unfinished and to indicate that another part is coming after the connector. There are two reasons why they become such a salient environmental feature associated with simultaneous speech. Firstly, the appearance of a connector signals the termination of a unit which could have a phrasal, clausal or sentential boundary, as shown in the above examples. Therefore, they can be regarded as the end of the prior utterance and the start of the next utterance as well. Secondly, overlapped speech cannot be expected to start exactly at a boundary, because the speech sounds are coming in such a continuous flow.

In summary, the five sets of discourse items discussed above can be further generalised into two categories. The first category functions as an appeal to the hearer's feedback: for example, address tags and hesitation markers. These provide an opportunity for the hearer to give a prompt response, whether by way of a brief backchannel, or by turn-bidding. The second category functions as a boundary signal: for example, termination tags, emphasizers and connectors. They mark out a transition-relevance place in the current speaker's speech which may encourage the hearer to attempt to take a turn.

Comparing the five sets of discourse items in each category of simultaneous speech, **Table 7** shows that a higher proportion of STB and UTB than BC are preceded by discourse items (25.2% > 19.0%, 30.6% > 19.0%). Even with regard to the occurrences of each set of discourse items, STB and UTB still have a higher proportion than BC in almost every set of discourse items (except that STB has a slightly lower proportion of hesitation markers than BC (3.3% < 4.0%)). This may indicates that the overlapper in BC, who is not aiming at bidding for a turn, may pay less attention to the five sets of discourse items. Looking at the frequency order of the five sets of discourse items, it is obvious that emphasizers are the most frequent, accounting for 202 (45.5%) of the 444 tokens of discourse items. This is particularly the case with STB, which has the highest proportion of this set. It suggests that taking full advantage of this set of discourse items may be particularly effective in turn-bidding.

As pointed out by Stenström (1994: 61), discourse items and interactional signals can do more than one thing in the discourse. They can have different functions in different places, or play more than one role in one place. Some of their other functions, which will be discussed in **4.3,** are related to the linguistic devices for producing simultaneous speech.

4.2.3   GRAMMATICAL CONTEXT

In the turn-taking system described by Sacks et al. (1974), a syntactic boundary is likely to be associated with overlapped speech. In the present study, the syntactic unit boundary is identified as a sentence boundary, a clause boundary, a phrase boundary or a "one-word" boundary such as the single-word utterances "yes" and "no" (see examples [29] to [32] in **3.3.2**). The word boundary is identified as a boundary between words within a phrase and between syllables within a word (see examples [33] to [36] in **3.3.2**). Their distribution in the corpus is shown in **Table 9**.

**Table 9   The Last Grammatical Unit Structure Before
STB, UTB & BC**

Boundary type	STB (830)		UTB (333)		BC (706)		Total (1869)	
	tokens	%	tokens	%	tokens	%	tokens	%
Syntactic boundary	624	75.2	259	77.7	564	79.9	1447	77.4
Word boundary	197	23.7	68	20.4	136	19.3	401	21.5
Laughs & inaudible syllables	9	1.1	6	1.8	6	0.9	21	1.1
Total	830	100.0	333	100.0	706	100.0	1869	100.0

Note:
The percentages are calculated in the same way as in **Table 7**.

Table 9 tells us that 77.4% of simultaneous speech occurs at a syntactic boundary of the current speaker's turn which can be identified by an explicit syntactic boundary marker. This, to a large extent, confirms Sacks et al.'s statement (1974: 721) which regards these syntactic structure boundaries as possible transition-relevance places for turn-taking. Though they are not necessarily possible turn completion places, they recur discretely in the course of a turn, thus offering a possibility for the next speaker to start a new turn or for the hearer to give a backchannel. Backchannels make up the highest percentage of such syntactic boundaries (79.9% > 77.7% > 75.2%), though it has been said that backchannels

"are subject only to very few, if any, restrictions as to their placement." and "there do not seem to be any restrictions on the hearer signal occurring simultaneously with any of the speaker's utterances" (Bublitz, 1988: 183-184). The present study shows that backchannels do not occur randomly, but mostly occur at a syntactic boundary which is usually coincident with an information chunk. Therefore it is a reasonable place for the hearer to give a comment or response. But Bublitz correctly points out that "Hearer signals are less frequently given simultaneously with the initial words of a speaker contribution, since they would then be redundant." (ibid: 184). This is supported by the present study, which shows fewer than one fifth (19.3%) of backchannels occurring at a word boundary.

However, this is not the whole problem, because **Table 9** also shows there are still many instances of STB and UTB (23.7% and 20.4% respectively) occurring at a word boundary. In other words more than one fifth of simultaneous speech instances occur at places which have no explicit syntactic markers. They appear at the "non-possible-completion point of a unit type" in Sacks et al.'s terms. According to their description, those overlaps which do not occur at "possible completion points" are what they term "turn-taking errors, violations and troubles" which should be avoided or need repair (see Sacks et al., 1974: 723). This raises the question of whether this number of turn-bidding instances can really be described as turn-taking errors which should be avoided. It also needs to be explained why so many hearers do not make turn-bids until the syntactic unit boundary actually occurs. One possible answer could be that the hearer's need to turn-bid is so urgent that he or she has to start speaking before it is too late. The other reason may relate to another problem --- whether a word boundary can really be a syntactic boundary which is predictable before its actual occurrence. Some data from the corpus suggests an affirmative answer to this question. eg.

[36]   B: ...you were ^going to be able to come and **pick it %*\up#*** /
        A: **&*^y\eah#*$** / ^where ^where \is your `office# / in the Str\and# /
                                    (LLC, S7.3l: 1347-1350)

[37]   A: ... could `almost k\ick your `way into `sandstone **c/ouldn't %*you#***
        B: **&*^you*$** c\ould# / if ^you were in b\oots# /
                                    (LLC, S7.3f: 612-614)

[38](After speaker A tells speaker "a" that she is actually recovering from several
        days of strange meals and drinks and things which were given by some
        terrible Greeks and she then finished up with a session with Mary Patrick all
        last night which is just about to finish her off, speaker "a" takes the turn as
        follows:)

        a: ((so it's)) for whom the bell . **for I*whom*** the bell tolls
        A: ||*((^d\esperate# .))*
                                    (LLC, S1.9: 53-54)

[39]   C: ...how I `go for . bits and `pieces that are going \on `then **/all I*right#**
                /and   ^I will -* go into your financial ...
        B: ||*^r\ight# . / ^f=ine# / ^thanks very `much ind\eed#* /
                                    (LLC, S8.2a: 905-909)

[40]   B: ...^Mrs Kay Gr/een#- / ^and the phone **n=um %*ber#*** /
        A: **&*^four*$** `six n/ine# /
                                    (LLC, S9.1n: 1270-1272)

[41]   f: ...))^I wouldn't kn\ow# . / it's ^only my id\/ea# / **^Victor
            %*M/ishcon#***
        c: **&*((well))*** . you may be r\ight# /
                                    (LLC, S5.4: 1087-1089)

Here, all the turn-bidding instances occur at a word boundary place. Yet a closer
examination of the current speakers' speech before the turn-bidding finds that their
structures and content are usually predictable. In other words the turn-bidders can
anticipate a possible completion of the current speakers' utterance by noting the
existing syntactic patterns or collocative phrase structures, or the other semantic

context of the topic being discussed. For example, in [36] and [37], if the hearer has the knowledge that "pick + pronoun + up" is a fixed verbal collocation, and that the statement "you could" can sometimes be followed by a tag question "couldn't you", he or she will not find it difficult to predict the coming words from the current speaker, which by then may be a complete syntactic unit, because all these word boundary structures have involved some predictable information in the communicative context.

[38] and [39] are UTB instances in which the turn-bidders make use of the prior semantic context to bid for a turn. In [38], speaker A could be familiar with the connotation of the phrase "for whom the bell tolls" --- which is quite relevant to their present topic and therefore she does not wait until speaker "a" finishes the whole sentence before starting to speak. In [39], speaker B may take speaker C's rising tone "all right" as a usual feedback-seeking phrase at the end of a statement, and so quickly makes a response before speaker C has actually finished the phrase. Yet she does not know speaker C still has something to say after that, and hence there is an unsuccessful turn-taking.

The occurrence of turn-bidding may be associated with other contextual cues. In [40], both speakers may be quite familiar with the telephone conversation routine which requires that to contact a person, the speaker should not only know that person's name but also the telephone number as well. Therefore speaker A immediately says the number even before speaker B has completed the inquiry. In [41], speaker "c" happened to be the participant whose name was being addressed, and therefore he had no hesitation in taking up the turn before speaker "f" had finished addressing his name.

In summary, any one of the syntactic boundaries such as the end of a sentence, a clause, a phrase or a one-word utterance is likely to be a proper linguistic

environment for simultaneous speech. Yet quite a number of word boundaries which have provided a structural or semantic context for prediction can also signal a transition-relevance place for turn-bidding or backchannelling. Such word boundary structures have embedded redundant information in the communicative context and it is such redundancy that makes such a prediction possible. In this sense such a word boundary can be called an implicit syntactic boundary because it may signal a possible completion point for a sentence or a clause or a phrase. Therefore it can also be a relevant place for turn-bidding or backchannelling. **Table 9** shows that STB has a slightly higher frequency of word-boundary occurrence than UTB and BC (23.7% > 20.4% > 19.3%). Although it is difficult to count how many word-boundary instances are being predicted exactly by the overlappers as an oncoming syntactic boundary, one could assume that the overlappers in STB are better in anticipating the boundaries by using predicting skills, resulting in a much higher frequency of STB than UTB in this corpus. As a spoken language is structured with redundancy, and as a human speaker has the capacity to predict the forthcoming structures and meaning, STB is, in fact, a process of interaction between the hearer's anticipation and the predictable structures and information in the current speaker's turn. In other words, a current listener can make full use of the language features and other contextual cues to anticipate a possible completion point in the current speaker's speech, taking it as a relevant transitional time and starting a new turn, thus achieving a STB speech act. This may support, from another aspect, the existence of what Sacks et al. called: "the projectability of possible completion or transition-relevance places" in the turn-taking system of conversation (Sacks et al., 1974: 707).

4.2.4 ASSOCIATION BETWEEN THREE LEVELS OF BOUNDARY MARKERS

As was mentioned in the literature review, a computerised and prosodically transcribed spoken corpus can provide an opportunity to study comprehensively a language phenomenon from its prosodic, lexical and grammatical aspects. Here an analysis of boundary markers on these three levels reveals some interesting findings. For the validity of comparison, all the tokens in **Table 10** are those which have been prosodically transcribed. A chi-square test with two degrees of freedom was undertaken to find if there is any association between any two of the three levels of boundary markers.

Table 10 Association between Boundary Markers at Three Levels

(1) Syntactic boundary and Prosodic boundary:

	with syntactic boundary	also with tone unit boundary	%	with tone unit boundary	also with syntactic boundary	%	p-value
STB	536	310	57.8	318	310	97.5	0.00*
UTB	219	134	61.2	137	134	97.8	0.00*
BC	508	342	67.3	350	342	97.7	0.00*
Total	1263	786	62.2	805	786	97.6	0.00*

(2) Prosodic boundary and Lexical boundary:

	with tone unit boundary	also with lexical boundary	%	with lexical boundary	also with TU boundary	%	p-value
STB	318	129	40.6	179	129	72.1	0.00*
UTB	137	50	36.5	88	50	56.8	0.00*
BC	350	73	20.9	120	73	60.8	0.00*
Total	805	252	31.3	387	252	65.1	0.00*

(3) Lexical Boundary and Syntactic Boundary:

	with lexical boundary	also with syntactic boundary	%	with syntactic boundary	also with lexical boundary	%	p-value
STB	179	171	95.5	536	171	31.9	0.00*
UTB	88	81	92.1	219	81	37.0	0.00*
BC	120	100	83.3	508	100	19.7	0.00*
Total	387	352	91.0	1263	352	27.9	0.00*

Note:
* = significant at p =< 0.004

Table 10 shows the p-values are all less than 0.004 which are highly significant and provide strong evidence that the boundary markers at three levels are associated. In section (1) of the table, almost two thirds (ie. 62.2%) of the syntactic boundary instances also end with a TU boundary marker, while nearly all of the prosodic boundary instances (ie. 97.6%) also end at a syntactic boundary. This is consistent with Halliday's point (1989: 36) that a clause (a grammatical unit) is always associated with a tone group (a phonological unit characterised by pitch movement). It is also consistent with Duncan's point (1973: 37-38) that a phonemic clause is often additionally marked by the completion of a grammatical clause, involving a subject-predicate combination. In section (2) of the table, almost two thirds (ie. 65.1%) of the instances with a lexical boundary (which is displayed by the five sets of discourse items) are also immediately followed by a TU boundary marker, and about one third of instances with TU boundary markers (31.3%) are associated with discourse items. With the salient prosodic features such as an obvious pitch change or a pause, these discourse items become more prominent signals to the hearers. In section (3) of the table, 91% of the instances with discourse items can be accounted for as a syntactic unit. About 30% of instances with syntactic boundary markers are also associated with discourse items. The coincidence of boundary markers at these three levels makes it easy for hearers to recognise the appearance of the unit boundary in the current speakers' speech. Moreover the interrelation of the three-level boundary markers indicates that most of simultaneous speech instances do not occur randomly. They are closely associated with boundaries which are signalled by a combination of prosodic, lexical and syntactic boundary markers.

4.2.5   SUMMARY

The findings of the study in terms of the environmental features associated with the occurrence of simultaneous speech are as follows:

(a) The occurrence of simultaneous speech is based on a context (linguistic and non-linguistic) which is rule-governed.

(b) The most important "rule" is that simultaneous speech is most likely to occur at a unit boundary.

(c) The unit boundary is associated with three levels of the speaker's language: prosodic, lexical (discoursal) and grammatical.

(d) The unit boundary on each linguistic level is linguistically marked by salient features; and the features at different levels are closely associated.

(e) A unit boundary can be signalled both by explicit markers or by implicit cues. The implicit cues usually occur before or near the actual boundaries. Hearers mainly seem to note the explicit boundary markers as a signal to start simultaneous speech, but they also make use of the implicit cues to anticipate the coming boundary.

(f) The hearer's anticipation relies on the redundant information involved in language. Such redundancy is presented by linguistic forms, content, and other contextual cues.

(g) Different categories of simultaneous speech have a similar tendency to use the above environmental features, but can have a stronger or weaker tendency to use particular features. In comparison with STB and UTB, BC occurs more frequently at a prosodic and syntactic boundary but less frequently at a lexical boundary displayed by a discourse item. In comparison with UTB, STB occurs more frequently at the three levels of boundaries, especially at those implicit ones which can positively signal an oncoming unit boundary.

4.3 LINGUISTIC DEVICES AND STRATEGIES USED BY OVERLAPPERS IN SIMULTANEOUS SPEECH

After analysing the linguistic environmental features which occur immediately before simultaneous speech, ie. features on the current speaker's side, we now turn to explore the linguistic devices and strategies which are used by the overlapper in carrying out simultaneous speech. The analyses were carried out from three aspects: prosodic devices, discourse items and repetition strategies. The purpose is to see how these devices and strategies function in realising simultaneous speech and how they are associated with each category of simultaneous speech in particular.

4.3.1   PROSODIC DEVICES

Speech sounds and their intonation contours are essential in transferring spoken information. Therefore their representation in simultaneous speech can be very significant. In this section, all five categories of simultaneous speech are involved, ie. besides STB, UTB and BC (as were analysed in **4.2**), STC and UTC are also taken into account, because they both have overlappers' linguistic features. The data in **Tables 11-14** show that the most prominent prosodic features are concentrated at the beginning and ending of the overlap and are also associated with speech rate and amplitude change through the overlap. These features include:

devices initiating the overlap:	a) stressed initials
	b) initials with nucleus
	c) initials as single-word tone units
devices used during the overlap	a) faster speech rate
	b) stronger amplitude
devices ending the overlap	a) incomplete tone unit structures

(1)   Devices Initiating the Overlap

Table 11 shows that 67.5% of the simultaneous speech tokens in the corpus are associated with a stressed word at onset --- that is, the majority of the initial words in the first TU of the overlappers' speech are stressed onsets. Second, 46.6% of the initial words in the overlaps are highlighted by an obvious change of pitch direction, ie. they are not only stressed but also adopt a nuclear pattern. Third, 41.7% of the initials are followed by a TU boundary marker, which means that they are not only stressed and nuclear patterned, but also construct a complete TU. Therefore, more than two thirds of the instances of simultaneous speech have a stressed initial, and more than 40% of the initials are further strengthened by having a nucleus or by forming a single tone unit. All this contributes to a prominent start to the simultaneous speech on the overlapper's side.

Table 11   Prosodic features of Overlappers' Initial Word

categories	Total tokens	With stress		With nucleus		As a TU	
		tokens	%	tokens	%	tokens	%
STB	830	502	60.5	254	30.6	198	23.9
UTB	333	174	52.3	68	20.4	51	15.3
STC	71	37	52.1	18	25.4	16	22.5
UTC	71	32	45.1	22	30.9	22	30.9
BC	706	613	86.8	575	81.4	551	78.0
Total	2011	1358	67.5	937	46.6	838	41.7

Note:
The percentages of each feature are based on the total tokens of each category of simultaneous speech.

Examining each category of simultaneous speech, BC has the highest frequency of such prominent starting. This is because most of the backchannels consist of only one word or one sound such as "yes", "ok", "right", "m". The briefness of backchannels causes a high frequency of stressed single words, which are at the same time nuclear patterned and constructed of a single-word TU. Previous study

of backchannels (Bublitz, 1988) concluded that the nuclear tone in backchannels is generally lower than that of the preceding or simultaneous uttered syllables. Yet the present study does not show such a feature, possibly because this feature is not marked up in the original corpus and because the tape-recorded version of the corpus was not accessible. However, comparing STB and UTB, the former has a much higher frequency of prominent beginnings than the latter in terms of the three beginning features (60.5% > 52.3%, 30.6% > 20.4%, 23.9% > 15.3%). Comparing STC and UTC, successful turn-competitors also have a higher frequency of stressed initial sounds than unsuccessful turn-competitors (52.1% > 45.1%). This indicates that a prominent start favours successful turn-bidding, because, as noted by Levinson (1983: 301):

> The speaker who "upgrades" most wins the floor, upgrading consisting
> of increased amplitude, slowing tempo, lengthened vowels and other
> features.

Therefore, a prominent prosodic start can serve as a turn-bidding signal. It serves the function of addressing the current speaker with "I want to speak now!" In achieving such prominence or attraction, quite a number of usually unstressed words become onsets of the first TU in STB and STC. This involves personal pronouns such as "you", "he", "we", "they", "it", "his"; prepositions such as "for", "in"; conjunctions such as "and", "or", "but", "if" and auxiliaries such as "be", "can", "may". eg.

[42]   c: ...the old Godfather . first God %*father*
        A: &*^**be** a*$:pretty good :double :bVill `that `actually# / we'd be ^out of
                                                                (LLC, S2.10: 647-649)

[43]   (B: ...I think it will just be be a few days in the Lakes
        A: uhuh (. laughs))
        B: ^really %*`just*
        A: &*^**for** a*$ ch/ange# - /
        (B: for a change)   (LLC, S7.3f: 844-849)

[44] (C: ...you stand to make a capital gain on that of at least two and a half
         thousand three thousand pounds within the first year)
      B: ^y\eah# - / ^yeah I s\ee th/at# / %*^y/eah#* /
      C: &*^and*$ it's a \very :s\Vensible `house \for you to ``\b\uy# /
         be^cause...
                                          (LLC, S8.2a: 445-455)

[45]   (c: I can recommend you a book)
        B: >*^it's a ``form of an``:\alysis my ``dear#*
        A: >>*^if you were . ^if you were* if ^you were if ^you were sort of -
           un\edu`cated# ^and :ph\ili`stine#
                                          (LLC, S210: 351-354)

Here the overlappers in [42] and [43] are achieving STB by helping the current
speakers to finish their incomplete utterances. The highlighted "be" in [42] and
"for" in [43] indicate that the coming words are not a feedback, but a continuation
of the current statements. In [44], speaker C's stressed conjunction "and" stops
speaker B's agreement with the previous statement, instead, he lays emphasis on a
further development of the topic. In [45], both turn-competitors are starting the
overlap with an unusual stressed onset such as "it" by speaker B and "if" by
speaker A. But the successful competitor A persists in stressing the "if" and "you"
and finally gains the floor. Moreover, in the four examples above, the words
following the onsets all have a tendency to increase in pitch range, as reflected by
the booster markers in front of the words such as "pretty", "double", "bill" in [42],
"very", "for", "buy" in [44] and "if", "you" in [45]. The abnormally stressed
words in the initial position of STB and STC and the subsequent increasing pitch
can have the function of foregrounding the new turn or strengthening the coherence
between the current speaker's and the overlapper's speech. In UTB and UTC there
are also some similar stressed cases, but they are not as frequent as those in STB
and STC.

(2) Devices Used during the Overlap

The present study generally supports the previous research findings that a more rapid speech rate and higher vocal amplitude than normal speech is associated with successful turn-competition (see Duncan & Fiske: 1977, Roger: 1989, Ng et al.: 1993). These two devices are discussed respectively as follows:

(2a) Speech Rate

Here the speech rate within the simultaneous speech is evaluated by comparing the average number of words spoken by the current speaker and by the overlapper during the time of simultaneous speech. It is based on the assumption that, within a given time, a larger average number of words and more unit pauses in the overlap indicate a faster speech rate (for a detailed rationale see **3.3.3 (1)**).

Table 12   Comparison of Speech Rates between Current Speakers and Overlappers during the Overlaps

	(a) total number. of overlapping words	(b) total tokens of overlaps	(c) average number of words in overlaps	(d) total number of unit pause in overlaps
STB				
current speaker	2127	830	2.6	124
overlapper	2363	830	2.9	144
UTB				
current speaker	1022	333	3.1	89
overlapper	975	333	2.9	46
BC				
current speaker	1091	706	1.5	115
overlapper	1014	706	1.4	63
STC	203	71	2.9	10
UTC	179	71	2.5	8

Note:
The average number of words is obtained by dividing the total number of overlapping words by the total tokens of overlaps, ie. (c) = (a) ÷ (b).

Table 12 shows that in different categories of simultaneous speech, the current speaker and overlapper have a different speech rate. In the case of STB, the overlapper appears to have a slightly faster speech rate than the current speaker,

because the former has a larger average number of words and more pauses (2.9 w. > 2.6 w., and 144 > 124 in pauses) than the latter. Yet this is not the case in UTB. **Table 12** shows that it is the current speaker who has a slightly faster speech rate than the overlapper (3.1 w. > 2.9 w. and 89 > 46 in pauses). The contrast of speech rates in these two categories suggests that the turn-bidder who has a faster speech rate than the current speaker during the simultaneous speech is more likely to gain the floor. This is further supported by the results from the comparison between STC and UTC in this study. **Table 12** shows that the successful turn-competitor also has a slightly faster speech rate than the unsuccessful turn-competitors during the simultaneous speech (2.9 w. > 2.5 w. and 10 > 8 in pauses), suggesting that even in a turn-competing situation, the one who speaks faster is more likely to gain the floor. Apparently, a faster speech rate can be a positive factor leading to a successful turn-bidding and competing. However, the situation in BC is different. **Table 12** shows that the overlapper has a slightly slower speech rate than the current speaker (1.4 w. < 1.5 w. and 63 < 115 in pauses). This could be due to the non-turn-bidding nature of the backchannel speech act. That is, the backchannel makers need not increase their speech rate during the overlaps, for they are not aiming at taking over a turn.

(2b) Amplified Volume

In simultaneous speaking, the overlapped speech can influence the sound quality and amplitude of both speakers. Those who have a higher volume or a louder voice may suppress those who have a weaker or lower voice. Thus the former can be heard clearly while the latter cannot. If we accept that degree of clarity is closely associated with vocal amplitude, the degree of clarity in the overlap may be very significant for the realisation of STB and STC. In the present study, degree of clarity is evaluated by collecting the occurrence of unclearly-heard cases in the overlaps on both sides. A comparison was made of the clarity between the overlap of the current speakers and that of the overlappers.

Table 13 shows that different categories of simultaneous speech in the corpus have different degrees of clarity in the overlap and that even within the same category of simultaneous speech, the degree of clarity varies between the start and the conclusion of the overlap.

Table 13   Cases of Unclearly-Heard Speech in the Overlaps

Category	Position of overlap	Current speaker		Overlapper	
		tokens	%	tokens	%
STB	initial	161	19.4	119	14.3
	end	217	26.1	72	8.7
UTB	initial	44	13.2	114	34.2
	end	43	12.6	134	40.2
BC	initial	65	9.2	64	9.1
	end	65	9.2	70	9.9
STC	initial			10	14.1
	end			10	14.1
UTC	initial			22	31.0
	end			20	28.2

Note:
The percentage is obtained by dividing the tokens of unclearly-heard cases by the total tokens of simultaneous speech in each category.

Comparing the two turn-bidding categories, ie. STB and UTB, the overlappers in STB have a much lower frequency of unclearly-heard cases than those of the current speakers, whether it is at the initial part or at the ending of the overlap (14.3% < 19.4% and 8.7% < 26.1%). It is assumed that this means the turn-bidders in STB generally have a louder amplitude than the current speakers. However, the overlappers in UTB have a much lower degree of clarity than the current speakers, because the former have a much higher frequency of unclearly-heard cases in the two positions (34.2% > 13.2%, 40.2 > 12.6). The lower clarity may be due either to the diminishing volume of one speaker or the increasing volume of the other. Moreover, the speech in STB has an increasing clarity during the overlap, because unclearly-heard cases occur much less at the

end of the overlap than at the initial part (8.7% < 14.3%). However, the current speaker's speech, ie. of the speaker being interrupted, has a decreasing clarity during the overlap: there is a much greater occurrence of unclearly-heard cases at the end of the overlap than at the initial part (26.1% > 19.4). On the contrary, the overlappers' amplitude in UTB is decreasing during the overlap (40.2% > 34.2%) while the current speakers' voices do not change much (12.6% : 13.2%). A similar contrast also occurs in turn-competing speech acts, ie. between STC and UTC. The degree of clarity in STC is much higher than that in UTC, because in STC the occurrence of unclearly-heard cases is much less than UTC, whether it is at the beginning or end of the overlap. (14.1% < 31%, 14.1 < 28.2) As Meltzer et al. (1971: 392) note:

> the change in interruptees' amplitude from before to during the interruption was particularly effective in determining the outcome of very short interruptions.

The amplitude difference between STB and UTB and between STC and UTC again suggests that overlapper's "higher volume provides a powerful resource for the achievement of turn-competition interruption" (French & Local, 1986: 178).

(3)   Devices Ending the Overlap

Prosodic features at the end of the overlap are important for the outcome of simultaneous speech, ie. who can continue to speak after the overlap. **Table 14** presents the prosodic features of the last TU in the overlaps of the turn-bidders, turn-competitors and producers of backchannels. It reveals a sharp contrast between the successful turn-bidding/competing cases and the corresponding unsuccessful ones in terms of the prosodic features at the end of overlaps. The TU structure at the end of STB and STC has a strong tendency towards incompleteness in comparison with that of UTB and UTC. As mentioned in **4.2.1**, a complete TU is marked by a TU boundary or a nucleus. In **Table 14**

both STB and STC have a much lower percentage than UTB and UTC in terms of TU boundary markers (29.2% < 47.7, 32.4% < 46.5%) and nuclear patterned words (42.9% < 48.6%, 42.3% < 47.9%). The fewer occurrences of complete TU structure at the end of STB and STC indicates that the overlappers still have something more to say, and thus do not provide an opportunity for others to insert speech.

Table 14   Features of the Last TU at the End of the Overlap

Category	Total tokens	(1)With TU boundary tokens(%)	(2) With nucleus tokens(%)	(3) Falling tone tokens(%)	(4) Rising tone tokens(%)	(5) Level tone tokens(%)
STB	830	242(29.2)	356(42.9)	249(69.9)	87(24.4)	20(5.6)
UTB	333	159(47.7)	162(48.6)	103(63.6)	55(34.0)	4(2.5)
STC	71	23(32.4)	30(42.3)	24(80.0)	6(20.0)	0(0)
UTC	71	33(46.5)	34(47.9)	20(58.8)	13(38.2)	1(2.9)
BC	706	648(91.8)	648(91.8)	454(70.1)	106(16.4)	88(13.6)
Total	2011	1105(54.9)	1230(61.2)	850(69.1)	267(21.7)	113(9.2)

Note:
The percentage of feature 1 and 2 is based on the total token of each category of simultaneous speech, while the percentage of feature 3 to 5 is based on the total tokens of feature 2 in each category of simultaneous speech.

However, instances of UTB and UTC present a much higher percentage of prosodic boundary features. First, 47.7% of UTB and 46.5% of UTC have a TU boundary marker; second, 48.6% of UTB and 47.9% of UTC have a nuclear pattern in the last TU of the overlap. These unit boundaries may indicate a turn transition-relevance place or signal that the overlapper might relinquish the floor. **Table 14** also reveals a strong tendency towards complete TU structure in BC, for it has the highest percentage of TU boundary markers and nuclear patterned words (both are 91.8%). This means that all the nuclear-patterned words are in fact followed by a TU boundary marker, thus forming a large number of single-word TUs in backchannels. The strong tendency towards prosodic completeness in BC

can be interpreted as evidence that the overlappers have no intentions of continuing to speak after the overlaps, but rather, they are just supporting or encouraging the current speakers to go on with their speech.

In terms of the nuclear patterns of the last TU in the overlap, Table **14** shows that the majority of the nuclear patterned instances in each category end in a falling tone (STB: 69.9%, UTB: 63.6%, STC: 80%, UTC: 58.8%, BC: 70.1%) . Yet STB has a higher percentage of level tones than UTB (5.6% > 2.5%). As level tone is a signal of an incomplete phonemic clause in Duncan's terms (1972), and as it is a tone pattern mostly used to indicate unfinished speech (see O'Connor: 1967; Halliday: 1985), this feature is consistent with the general tendency towards incomplete prosodic endings in STB. However, BC has an even higher frequency of level tones than any other categories (13.6%). This can be explained by the fact that many single-word backchannels with a level tone may not indicate the overlappers' intention to continue the speech, but rather, their short and non-prominent pitch contour may sound less intrusive to the current speakers, ie. allowing them to continue their speech without being interrupted.

In summary, the prosodic strategies for each category of simultaneous speech are different. STB and STC are characterised by:

(a) more frequent prominent beginnings (ie. an initial stressed onset often with a nuclear pattern and followed by a TU boundary marker);

(b) a slightly faster speech rate than the current speaker or the other competitor (ie. a greater number of words and more pauses in the overlap);

(c) a presumed higher amplitude than the current speaker or the other competitor (ie. less occurrence of unclearly-heard cases and an increasing degree of clarity during the overlap)

(d) a stronger tendency towards prosodic incompleteness at the end of the overlap.

For UTB and UTC, all the above features occur with a comparatively lower percentage.

BC is characterised by:

(a) the highest percentage of prominent prosodic features at the beginning of the overlap;

(b) no obvious difference in amplitude between the current speaker and the overlapper during the overlap;

(c) a slightly faster speech rate of the current speaker during the overlap;

(d) a very strong tendency towards complete prosodic units at the end of the overlap.

4.3.2   DISCOURSE ITEMS

This analysis of discourse items used by the overlappers focuses on the initial words of each overlap. A wordlist was made by computer of all the words which occurred in the corpus at the beginning of the overlappers' speech in each category of simultaneous speech. The most frequently occurring words are presented in descending order of frequency in **Table 15**. They include only those types whose tokens make up 2% or more of the total tokens in each category of simultaneous speech. Therefore they can be regarded as more frequent initial types than the rest.

Accounting for the proportion of the most frequent **types** of initials in relation to the proportion of their **tokens**, **Table 16** shows that in each category of simultaneous speech there is a limited number of initial **types** which make up quite a large proportion of total initial **tokens**. For example, in STB the eleven most frequent initial types make up only 7.4% of the total of 149 initial types, but the number of initial tokens covered by these ten types makes up 61.8% of the total 830 initial tokens. In BC, the five most frequent initial types (ie. "yes", "m", "no", "I",

"right") make up 81.3% of the 706 tokens of this category. This indicates that there is a small set of initial items frequently used by the overlappers in a large number of instances of simultaneous speech. It is thus worthwhile to study what this set of items is and why it is so frequently used in simultaneous speech.

Table 15   The Most Frequent Initial Words in Simultaneous Speech

STB		UTB		BC		STC		UTC	
type	tokens(%)	type	tokens(%)	type	tokens(%)	type	tokens(%)	type	tokens(%)
yes	102(12.3)	I	28(8.4)	yes	281(39.8)	I	6(8.5)	m	11(15.5)
I	76(9.2)	no	22(6.6)	m	237(33.7)	bye	5(7.0)	no	6(8.5)
no	61(7.4)	yes	13(3.9)	no	26(3.7)	oh	5(7.0)	I	4(5.6)
m	56(6.7)	oh	12(3.6)	I	15(2.1)	well	4(5.6)	bye	3(4.2)
well	49(5.9)	and	12(3.6)	right	15(2.1)	and	4(5.6)	so	3(4.2)
oh	45(5.4)	i t	11(3.3)		574(81.3)	m	3(4.2)	hello	2(2.8)
and	34(4.1)	@	10(3.0)			how	2(2.8)	ok	2(2.8)
i t	27(3.3)	we	7(2.1)			i t	2(2.8)		31(43.7)
you	26(3.1)	what	7(2.1)			but	2(2.8)		
but	20(2.4)		122(36.6)			no	2(2.8)		
the	17(2.0)					what	2(2.8)		
	513(61.8)					when	2(2.8)		
							39(54.9)		

Note:
(a) For the convenience of analysis in the wordlist, words like "yes", "yeah", "yup" are classified as being the type "yes", because of their similarity in form and in meaning. The same is done for the type "m" which includes "m", "@", "@m", "mhm", "uhuh", "aha" and the type "oh" which includes "oh", "ooh", "ah".
(b) The percentage in the table is based on the total tokens of each category of simultaneous speech.

Table 16 **The Most Frequent Initial Types and Their Proportion of Total Initial Tokens**

category	(a) total types	(b) total tokens	(c) number of types which have 2% and more of tokens	% (c)÷(a)	(d) number of initial tokens covered by (c)	% (d)÷(b)
STB	149	830	11	6.7	452	54.5
UTB	98	333	9	9.2	122	36.6
BC	47	706	5	10.6	574	81.3
STC	37	71	12	32.4	39	54.9
UTC	37	71	7	18.9	31	43.7

Based on the results of **Table 15**, a further generalisation was made from those types so as to obtain the most frequently occurring types which are shared by at least two or more categories of simultaneous speech. The result is that eleven types of initials are shared by at least two or more categories of simultaneous speech. They are "I", "m", "no", "yes", "oh", "it", "well", "and", "but", "what" and "bye". These eleven types were further reduced to seven types, because the initial words like "I", "it", "what" and "bye" can only be analysed when accounting for their association with the other words or structures, while the other seven types, ie. "yes", "no", "well", "oh", "and", "m" and "but" are typical discourse items (see Stenström, 1990: 174). Considering their meaning and pragmatic effect, the seven types of initials words were further categorised into five sets as follows:

(a) agreement set: including "yes", "yeah", "yup", affirmative "no" and falling toned "m", "mhm", "uhuh", "aha"

(b) disagreement set: including negative "no'

(c) hesitation set: including "well", "@", "@m" and level toned "m"

(d) exclamation set: including "oh", "ooh", "ah"

(e) continuity set: including "and", "but"

As the function of a discourse item is closely related to its position in the discourse and its prosodic presentation, items such as "m" and "no" are put into more than one set of initial categories according to their different prosodic presentations or prior context. For example, a falling tone "m" usually indicates the hearer's agreement with the current speaker's opinion (see example [47] and [49]) while an unstressed or level tone "m" usually indicates the hearer's hesitation or planning in speech (see example [54]). Hence the former was put into the agreement set while the latter was in the hesitation set. Similarly the word "no" can be put into the agreement set when it follows a negative statement (see [50]) or it can be in a disagreement set when it expresses a negative answer or shows a different opinion to the prior statement (see [52] and [53]). **Table 17** outlines the distribution of the five sets of discourse items in the five categories of simultaneous speech.

Table 17   <u>Five Sets of Discourse Items at the Beginning</u>
<u>of Simultaneous Speech</u>

Function	STB (830)		UTB (333)		BC (706)		STC (71)		UTC (71)		Total(2011)	
	tokens	%	tokens	%	tokens	%	tokens	%	tokens	%	tokens	%
agreement	156	18.8	22	6.6	501	71.0	8	11.2	9	12.7	696	34.6
disagreement	29	3.5	14	4.2	0	0.0	1	1.4	3	4.2	47	2.3
hesitation	84	10.1	31	9.3	44	6.2	4	5.6	2	2.8	165	8.2
exclamation	45	5.4	12	3.6	48	6.8	5	7.0	1	1.4	111	5.5
continuity	54	6.5	16	4.8	0	0.0	6	8.5	0	0.0	76	3.8
Total	368	44.3	95	28.5	593	84.0	24	33.8	15	21.1	1095	54.5

Note:
Figures in the brackets of the top line of the table are the total tokens for each category of simultaneous speech based on which the percentage of each feature is counted.

Table 17 shows that more than half (54.5%) of the 2011 instances of simultaneous speech are initiated by any of the five sets of discourse items. This suggests that these items are salient discourse devices associated with simultaneous

speech. Because of their importance, a further analysis needs to be made of the pragmatic functions of each set.

(1)   Agreement Set

Preference for agreement and continuity is a general principle in human communication. **Table 17** shows that more than one third (34.6%) of the overlappers' speech is introduced by an agreement. If other initial words or phrases displaying the same function, including "ok", "right", "good", or "I agree", "that's right", "that's ok" are included, this portion will be even larger. In BC, agreement initials account for 71% of the tokens. This is consistent with the major functions of backchannels, ie. to confirm that the hearer is taking note of or agreeing with what the current speaker is (or has been) saying or meaning (see Bublitz, 1988: 189). eg.

[46]   A: ...^from S\eahaven /onwards# >*.* / . ^what has now hVappened# - /
              can
         B: >>*^y/es#* /

(LLC, S1.2: 269 -271)

[47]   A:...^this is something that could be running through your m/\ind# /
              >*^or* something on `inton/ation# /
          B: >>*^[\m]#*. /

(LLC, S3.2a: 492-494)

Here in [47], hearer B gives a backchannel "yes" which may tell speaker A that he is listening to him and encourages him to continue. In [48], the falling toned "m" indicates that hearer B is agreeing with the prior statement made by speaker A.

Since agreement initials have both the function of taking note and of stating position, they can be a hearer's signal and a speaker's contribution to the current conversation. In previous research, words like "yes", "m", "aha" in the agreement

set were usually put into the backchannel category and seldom associated with interruption. The present result shows that these words are frequently used as a starting point for turn-bidding speech acts. In other words, turn-bidders are likely to begin their speech with a backchannel word, ie. agreeing with what is being said first, and then using it as a precursor to take over the floor. eg.

[48]　a: ...w\ell# . / sym^b\olic of # . / the ^future of w\/omen#- / %*
　　　　　((^m\/ilitant#))*
　　　b: &*^y\es#* / of ^course I never [@m]. `felt in the least ---^sort of .
　　　　　^w\ell.

　　　　　　　　　　　　　　　　　　　　　　　　(LLC, S6.4b: 714-719)

[49]　B: ...you ^get to `know the st\udents ((much % *`better*))# /
　　　A: &*^[\m]#*$ - / sup^pose r/eally ...

　　　　　　　　　　　　　　　　　　　　　　　　(LLC, S1.9: 978-980)

[50]　A: ...I ^don't want to b/uy it# / it's %*^all r\ight#* /
　　　B: &*^n\o# / because*$ it'll be ^all over the `bedroom fl\oor# / ...

　　　　　　　　　　　　　　　　　　　　　　　　(LLC, S2.10: 293-297)

[51]　A: ...have you d\one with {^that `French `packet#}# . / %*^Deb `bought...
　　　B: &*^y\es# / but ^these*$ are s\uper# -/ ^Arthur `doesn't sm\oke
　　　　　them...

　　　　　　　　　　　　　　　　　　　　　　　　(LLC, S2.10: 532-536)

In these four examples, the turn-bidders all use agreements to start the turn-bidding. The initial word "yes" in [48] and "m" in [49] are confirmative in character. Similarly, in [50], the word "no", though it is negative in meaning, is in fact confirming the previous negative statement. [51] is an instance of using an agreement as a precursor to express a side step. The word "yes" preceding "but" may soften the effect of the counter. Therefore it is a polite way of showing disagreement. Since turn-bidding can be face-threatening behaviour, both to the turn-bidder and current speaker, starting to bid for a turn with an agreement can be "a cooperative 'sharing' device" (Altenberg 1986: 28). It tends to make the

insertion less provocative and thus smooths the way to gain the floor. Comparing STB and UTB in **Table 17,** the former has a much higher frequency of agreement initials than the latter (18.8% > 6.6%), suggesting that using agreeing initials may be an effective strategy for successful turn-bidding.

(2) Disagreement Set

A disagreement initial which consists of the word "no" has the lowest percentage among the five sets (see 2.3% > 3.8% > 5.5% > 8.2% > 34.6% in the right hand column of **Table 17**), perhaps because saying "no" as an initial in simultaneous speech is likely to cause a disagreeable effect on the current speaker. It is noteworthy that it does not occur in BC at all, except in a few cases of using "no" as an agreement to confirm the previous negative statement like that in [50]. Comparing STB, STC with UTB, UTC in **Table 17**, the first two have less use of such disagreement initials than the last two (3.5% < 4.2%, 1.4% < 4.2%). Successful turn-bidders or turn-competitors tend to avoid using it but instead, use a more polite way to show disagreement, as was discussed in [49]. Even when "no" actually appears at the beginning of the overlap, it is displayed with a mild or soft intonation contour. For example:

[52] m: ... it with an`other qu\estion# . / ^do you %*appr/ove# / your^s/elf#* /
     d: &*((**no . no no**))*$ ^may I `ask this one qu\estion# / and be^fore
      we...

                            (LLC, S5.6: 220-224)

[53] a: but that site must have been cleared a long %*time now*
     A: &*^nVo#*$. / it ^hVasn't# - /

                            (LLC, S1.9: 833-835)

In these two examples, the turn-bidders seem to weaken the disagreement "no" to some extent, such as in [52], speaker "d" uses an unstressed "no" instead of a stressed onset. In [53], speaker A uses a falling-rising tone "no" in stead of a

falling tone "no". Such weakened prosodic features may soften the sound effect and thus convey a kind of mild disagreement. The rising tone "no" may show a kind of uncertainty, because in spoken English, utterances ending with a falling tail are usually associated with confirmation, resoluteness and completeness while those ending with a rising tail are usually associated with a shared feeling, a more polite manner, or a casual and interested attitude (see O'Connor: 1967, Brazil et al.: 1980). A prosodic comparison between the agreement "no" and the disagreement "no" in STB has made this point clearer as shown in **Table 18**.

Table 18 **Prosodic Features of Initial _"No"_ in STB**
(61 cases in total)

Functions	Stressed		Unstressed		Falling		Rising		Level		Total
	tokens	%	tokens	%	tokens	%	tokens	%	tokens	%	tokens
agreement "no"	29	90.6	3	9.4	18	56.3	3	9.4	1	3.1	32
disagreement"no"	25	86.2	4	13.8	9	31.0	6	20.7	1	3.5	29

Note:
The percentage is based on the total tokens of each type of "no", ie. 32 tokens for agreement "no" and 29 tokens for disagreement "no". Some tokens have more than one prosodic feature and therefore the percentages do not add up to 100%.

Table 18 shows that most of the instances of disagreement "no" in the initial position of STB have a less prominent prosodic contour than that of the agreement "no". The former are less stressed (86.2% < 90.6%), with less falling tone (31% < 56.3) but more rising tone (20.7% > 9.4%). Although there are only a small number of tokens, it still indicates that firstly, STB initiated by "no" does not necessarily mean a disagreement; and secondly, even those STB instances which are initiated by a disagreement "no" are likely to diminish the disagreeable effect by some prosodic devices such as an unstressed start or a less prominent intonation contour, so that the disagreement can be dealt with in such a way that an opposite opinion is conveyed on the one hand, yet on the other hand, the interpersonal bond or the solidarity between the interlocutors is still maintained.

(3) Hesitation Set

Hesitation initials have the second highest percentage of occurrence among the five sets. They are composed of words like "well" or sounds like "@", "@:m", "m" (usually unstressed or level toned).

[54] (C: ...when they're buying their first flat or a house)
        B: ^that's `right I can see %*th\at#*
        C: &*[@:m]*$. / ^if `you could `do th\/at# . /
                                                   (LLC, S8.2a: 779-781)

[55] A: it ^is it ^is a``possible %*st\arter /is it#* /
        B: &*^w\ell# - *$ / yeah I'd ^like to th\/ink about it# / ^y\ou know# /
                                                   (LLC, S3.2a: 347-350)

[56] d: ...it's the oldest . oldest cinema in London ---
        B: <*((a^mazing `new l\ights#))* /
        A: <<*well ^what were the `nice c\uddly* `seats `like# / ^sort of two two
                                                   (LLC, S2.10: 684-687)

The overlappers in the above examples all initiate their turn-bidding with a hesitation marker. On the one hand, this displays the overlapper's awareness of the current speakers' message and on the other hand, it provides the overlapper with increased flexibility and a certain measure of looseness in complying with their new turns (see Schiffrin, 1987: 111). Therefore hesitators soften or cushion the abruptness caused by the overlapping and create an informal atmosphere as well.

However, the hesitation initials such as the unstressed "@m" and "m" in BC do not display the overlapper's hesitancy in speaking, but instead they serve the function of taking note. They are used by the hearer to confirm to the current speaker

that he is listening and thus paying attention, that he is understanding what he has heard, that he is accepting it as relevant (at the present stage of the conversation) and adding it to his knowledge (so that henceforth the speaker may refer to it as known.

(Bublitz, 1988: 189)

[57] A: ... the h/ierarchy# / and ^this this in >*^now declared p\olicy#* / and ...

B: >>*(([m . m . m]))*

(LLC, S1.2: 480-483)

Here the unclearly-heard and unstressed "m" repeatedly made by hearer B is just confirming his status as a hearer who, for the moment, has no intention of interrupting the current speaker. In addition, it has the function of encouraging the current speaker to go on with the speech.

(4)   **Exclamation Set**

Words like "oh", "ah", "ooh" are traditionally defined as interjections to express strong emotional state. The pragmatic function of these exclamation markers is noted by Schiffrin (1987: 99) as follows:

"oh" has a role in information state transition because it marks a focus of (the) speaker's attention which then also becomes a candidate for (the) hearer's attention. This creation of a joint focus of attention not only allows transitions in information state, but it marks information as more salient with a possible increase in speaker/hearer certainty as to shared knowledge and metaknowledge. ... "Oh" as backchannel not only marks information receipt, and marks an individual as an occupant of a specific participation status (active recipient), but it also ratifies the current division of turn-taking responsibilities in the exchange structure.

The present research finds that among the five categories of simultaneous speech, such exclamation markers occur more frequently in STB and STC than in UTB and UTC (see 5.4% > 3.6%, 7% > 1.4% in **Table 17**) and they have a proportion of 6.8% in BC.

[58]   A: I'm ^quite `` s/\ure {it's ^unsh/\akeable#}#- / >*^unshakeable* in the ...
        B: >>*((^\ah#))*

                                   (LLC, S1.2: 482-485)

[59]   c: ((good idea)) yes she's somewhere in . Alberta . %* - ((several sylls))*
        B: **&*^oh** `` g/\olly# - / it's ^so `` c\old*$ in `Alb\erta {in the ^w/inter#}# /
                                   (LLC, S210: 314-316)

[60]   A: ... `meant we(('d))# / `` ^come out of /Europe#. / ^pretty qu\ick#--- /

        c: <*[m]*

        B: <<*^oh* but this is very nice n\ews `Gordon#. / ^very `nice ind\eed#

                                   (LLC, S2.10: 188-193)

Here, the exclamation markers in the three examples all signal the current listeners' recognition of the current speakers' messages, ie. serving as focus signals. And then in [59] and [60], with such highlighted initials which indicate that new turns are about to start, the overlappers continue to add their own messages.

(5)   **Continuity Set**

Connectors like "and" or "but" are regarded as discourse coordinators because they play a role in continuing, developing and even completing the current speakers' speech. As backchannels are not aimed at developing nor continuing the current topic, hearers are not likely to use such linking words as backchannels. Yet turn-bidders or turn competitors may use them to strengthen the coherence between the current speakers' turns and their new turns. **Table 17** shows that STB and STC have a higher frequency of this set of items than UTB and UTC (6.5% > 4.8%,

8.5% > 0), indicating this is also an effective strategy for successful turn-bidding or turn-competing. eg.

[61]   A: yes# / com^pared with [@:] Macb\eth %*and C/aesar#* /
        B: &*^**and** you*$ character`ised it `as a `tragedy of c\ircumstance# - /
                                        (LLC, S3.5: 917 - 919)

[62]   l: ^/\evidence#* / - ^/\evidence# . / ^that's \/all# - /
        n: <*I ^don't think*
        w: <<*((**and** we ^still))* are waiting for y\ours# /
                                        (LLC, S5.3: 654-658)

[63]   B: and ^then you could `ask for the most monu%*mental pr\esents#* /
        A: &***but** ^this is `why I asked that*$ qu\estion# / ((cos)) ^I was quite ...
                                        (LLC, S210: 949-951)

[64]   (a: I don't think it's a very comfortable position to be constantly
            hyperventilating and rushing off to Casualty ---)
        B: <*[m]*
        c: <<((***but*** ^[z=is]#)). / his `sister's an\/alysis# / ^and [`@:m]-- .
            there's...
                                        (LLC, S6.8: 604-607)

Here, the word "and" in [61] and [62] shows that the overlappers are continuing or giving a conclusion to the current topic, while the word "but" in [63] and [64] indicates that the overlappers are developing the topic, but in a contradictory way. These discourse connectors have made coherent and cohesive links in transferring turns. Furthermore they have shortened the gap between turn-yielding and turn-taking, thus realising a smooth transition.

Having discussed the five sets of discourse items in terms of their discourse and pragmatic functions, additional comments should be made about their prosodic prominence through the use of stress or by forming separate tone units.

Table 19 **Prosodic Features of the Five Sets
of Discourse Items**

Functions	Total tokens	With stress		With a nucleus	
		tokens	%	tokens	%
Agreement set	696	635	91.2	613	88.1
Disagreement set	47	40	85.1	28	59.6
Hesitation set	165	135	81.8	70	42.4
Exclamation set	111	79	71.2	39	35.1
Continuity set	76	26	34.2	1	1.3
Total	1095	915	83.6	751	68.6

Note:
As some tokens have more than one feature, the percentages do not add up to
100%.

Table 19 shows that 83.6% of the five sets of initials are stressed and 68.6% of

them have nuclear patterns, thus displaying the strategic importance of these

discourse items in simultaneous speech. Although the five sets of discourse items

constitute only a small number of initials in terms of word types, they make up

67.4% of the 1358 stressed initial words and 80.1% of the 937 nuclear-patterned

initial words in the simultaneous speech (see **Table 11 in 4.3.1 (1)**). If we

assume that stress and nucleus give prominence, the prosodic prominence ranking

order of the five sets is:

    (a) agreement set
    (b) disagreement set
    (c) hesitation set
    (d) exclamation set
    (e) continuity set

Prominent prosodic features are likely to enhance the discourse functions of these

sets and once again display the combination of prosodic devices and discourse

devices when language is being used.

The five sets of initials are salient discourse items in simultaneous speech in general. Yet some categories of simultaneous speech tend to use some specific discourse items more than the other categories do. The present study finds that STB uses more initials of agreement, hesitation, exclamation and continuity than UTB, and STC also uses more initials of hesitation, exclamation and continuity than UTC. This suggests that these initials are potential sources for successful turn-bidding and turn-competing. In other words, in UTB and UTC, the less frequent use of these four sets of initials and the greater use of disagreement initials could be one of the possible conditions associated with unsuccessful turn-bidding and turn-competing.

With regard to the five sets of discourse items used in BC, **Table 17** shows that BC uses agreement initials more frequently (71%) than any other categories of simultaneous speech though its total number of initial-word types is more limited than that of the others. This may be due to the positive nature of agreement initials in supporting the current speaker, ie. to display the hearer's awareness or acceptance of the current speech. However, BC does not use initials of disagreement and continuity, because it does not aim at arguing with the current speaker, nor does it aim at getting a new turn by continuing or developing the current topic.

4.3.3 REPETITION STRATEGIES

Repetition is a frequent phenomenon in English conversation (see Persson: 1974, Brown: 1977). The present research reveals that repetition also frequently occurs in simultaneous speech, especially used by the overlappers in turn-bidding and turn-competing situations. Its distributional frequency is displayed in **Table 20.**

Table 20   Occurrence of Three Types of Repetition
in Overlaps

Category	Total tokens	Initial		Mid		Post		Total	
		tokens	%	tokens	%	tokens	%	token	%
STB	830	61	7.4	79	9.5	114	13.7	254	30.6
UTB	333	31	9.3	23	6.9	0	0	54	16.2
STC	71	0	0	4	5.6	7	9.9	11	15.5
UTC	71	0	0	3	4.2	0	0	3	4.2
BC	706	0	0	26	3.7	0	0	26	3.7
Total	2011	92	4.6	135	6.7	121	6.0	348	17.3

Note:
The percentage of each type of repetition in each category of simultaneous speech
is based on the total tokens of simultaneous speech in that category. ·

Table 20 shows that repetition items at the beginning, middle and end of the
overlappers' speech make up 17.3% of the total 2011 simultaneous speech
tokens, yet the distribution is quite different in each category of simultaneous
speech. As **Table 20** shows, the five categories of simultaneous speech all
have tokens of mid-repetition, because all overlappers can repeat their own
utterances during the overlaps. Yet only STB, UTB and BC have tokens of
initial-repetition, because only in these three categories of simultaneous speech
can the overlapper have an opportunity to repeat the current speaker's speech,
while speakers in STC and UTC are only overlapping each other. Moreover,
only STB and STC have tokens of post-repetitions, because only in these two
situations do overlappers continue to speak after the overlap. As a result, STB
has repetition in three positions, showing that 30.6% of the total tokens of STB
include repetition. UTB has the first two types of repetition, showing that 16.2%
of the total tokens of UTB include repetition. STC has the later two types of
repetition, showing 15.5% of the total tokens of STB include repetition. UTC
has only the second type of repetition, showing that 4.2% of the total tokens of
UTC include repetition. BC can have the first two types of repetition but it turns
out to have only the second type, showing that only 3.7% of its total tokens
include repetition. From this analysis we can conclude that repetition has several

functions in simultaneous speech according to where it occurs, whose speech it repeats and how it is produced. The following is a discussion of these functions in each type of repetition.

(1) Initial Repetition

Initial repetition occurs when the overlapper starts his or her speech by repeating the current speaker's last word or syntactic pattern occurring immediately before the simultaneous speech. eg.

[65] B: ...[@] in [dhi:] **^publi`cations** %*dep\artment#* /
   A: **&*^publi`cations*$** dep\artment# / ^[/mhm]# /
   (LLC, S3.1b: 977-979)

[66] A: and **^that's [dhi] ^that's [dhi]** %*((. ^wh\at do you [m] `call
   it#))* /
   B: **&*^that's the** ``g\auleiters#* / ^y\es# /
   (LLC, S1.2: 1391-1393)

[67] B: ^no no n/o# / but I ^mean . do ^you `think **it's `worth** %*s\/eeing#*
   A: **&*di*$ verting perhaps**...
   (LLC, S2.10: 622-624)

[68] A: ... @] contr/ols {^on the b/ack#}# **^is for** [@] |*setting it's a
   ``^h\and*
   B: ||**^looking at the h/ands#*** /
   A: `setting# / [@:m] . ^is I don't know whether it's a scr/ew ...
   (LLC, S7.2d: 309-316)

Here in [65] the turn-bidder is repeating the current speaker's last word "publication"; in [66] the turn-bidder is repeating the current speaker's incomplete utterance "that's the ...". As pointed out by Brown (1977: 113-114), by repeating such lexical items and sentence patterns, the turn-bidder shows that he

accepts the formulation of the previous speaker and is prepared to agree with the first speaker's identification of the topic under discussion, even though he may disagree with his views on it. It has then a *solidary* social function.

In [67] and [68], the turn-bidders are not simply repeating the current speaker's original words or pattern: they are actually repeating the pre-structured patterns (ie. "to be worth + doing something" in [67] and "is for + doing something" in [68]). By utilising what appears to be a prefabricated utterance, the turn-bidder does not have to construct the whole of his new utterance from the beginning, but just establishes what is being discussed as a common basis and helps the current speaker to finish the utterance. He then takes the floor to continue the discourse. As is summarised by Bublitz (1988: 228) from Brown's (1977) analysis of repetition, such "repetitions facilitate formulation and help the speaker to focus on the repeated item". This type of repetition relies on the turn-bidder's anticipation of the coming utterance in the current speaker's turn, as has been discussed in **4.2.2**. Although only 7.9% of the 1163 turn-bidding instances were identified as initial repetition , they reinforce the view that a hearer's catching up with the current speaker's ongoing structure and helping to complete the utterance is a positive strategy for turn-bidding.

(2)   **Mid Repetition**

Mid-repetition occurs during the overlapper's simultaneous speech. Its major function appears to be to reiterate the overlapper's words, because the communicative channel can be affected by the current speakers' voice during the overlap. This type of repetition mostly involves a brief utterance, especially a single-word utterance. eg.

[69] n: ... `` ^b\ut# / %*^these ^these `^w/ait# . / ^let me just f/inish#* /
    l: &*ex^c\use me# / `^y\our point# . / ^your ``ly\our case#*$ / is
                                    (LLC, S5.3: 338-344)

[70] A: ...^/is it# / l*^y\es# / ^c\ourse it `is# / ^we're the* twenty -\eighth# /...
    B: ||*^y\es# . so ^he . ^so*
                                    (LLC, S7.2l: 1205-1208)

[71] B: ^o\k# / ^f\ine# / >*^this is the main ((b\edding)) and there's*
            pro^vision
    A: >>*^y\es# . / ^y\es# . / ^y\es#* /
                                    (LLC, S1.2: 51-55)

[72] (C: (-laughs) terribly yes (- - laughs))
    d: <*((boringly settled))*
    B: <<*I ^mean **is it ^is it** ``n\ice `being ((well)) ^is it n\ice#* /
            ^getting it ...
                                    (LLC, S2.10: 921-924)

[73] (A: when you're back in London I'll be in touch - .
    B: eh)
    A: <<*((^when)) - ^\oh# / you ^\are# - - * / ^w\ell# / ^would you `like .
    B: <*I ^am I ^\am back in _London# / I'm ^h\ere {``y\es#} [@:m]# /
                                    (LLC, S9.1i: 637-647)

Here the overlappers repeat their own utterances as shown in the bold type words
in each example. Such repetitions may reveal the speaker's need for time to plan
his or her utterance, or it may be because within the overlaps, neither speakers'
sounds are clearly heard. As the simultaneous speech in all these examples
(except [71]) involves a situation of bidding or competing for a turn, such
repetitions can enhance the turn-bidding or turn-competing behaviour, both in
terms of meaning and in terms of clarity of the sound. In [71], the overlapper's
repetition is not for turn-bidding but is a backchannel. The repetition of "yes"
shows that the hearer is quite aware of, or acknowledges, what the current

speaker is saying, and it also strengthens the hearer's positive response to the current speaker.

(3)   **Post Repetition**

Post-repetition occurs immediately after the overlap. It is produced by the overlapper immediately after he or she gains the floor. If we consider only STB and STC, that is, the only two categories that can have the opportunity to make post repetition, this type of repetition has a much higher frequency than the other two types of repetition. (STB: 13.7% (type 3) > 9.5% (type 2) > 7.4% (type 1); STC: 9.9% (type 3) > 5.6% (type 2), see **Table 20**), because at this moment, the overlapper has already gained the floor and he may need to recycle the start of the new turn.

[74]   a: ...talk [t@:m] about this sort of thing %*[@ @:] before the candidate*
       B: &*^n\o# - / **^I think you would***$. **I ^think you |w\ould**
          `find it# / ex...

<div align="right">(LLC, S3.1a; 485-488)</div>

[75]   l: ...nothing to d\o with it# / %*I ^beg your p\ardon#* /
       n: &*I ^**didn't***$ [@:] . **I ^didn't** say that they were irr\elevant# . /

<div align="right">(LLC, S5.3: 873-875)</div>

[76]   (v: ...so don't take him seriously
       aud: (laugher))
       f: <*^Colin H/odgkinson#* /
       c: <<***there ^is one*** - **there ^/\is one** thing# / [@:] . if ^one may

<div align="right">(LLC, S5.4: 1257-1261)</div>

Here the overlappers' repetitions all occur immediately after the overlaps. When the turn-bidders in [74], [75] and the turn-competitor in [76] are sure that they have gained the floor, they are likely to restate their new turn-beginning by repetition. In Sacks et al.'s turn-taking system, this is a recycled turn-opening which is regarded as a repair technique for a mistimed turn-taking error. As a

turn-initial place is very important in shaping and planning a new turn, identical repeats of turn-beginnings occur regularly when there has been an overlap of the turn-beginning with the prior turn (see Schegloff, 1987: 74). Levinson (1983: 300) also describes such repetition in that: "as soon as one speaker thus emerges into 'the clear', he typically recycles precisely the part of the turn obscured by the overlap." French and Local (1986: 173) further note that:

> by producing either simple or expanded repeats of incomplete syntactic units, interrupters can signal both that there is more to follow and that it is being withheld until they have gained sole occupation of the floor.

The present study supports these views and further recognises repetition as a positive strategy for successful turn-bidding and turn-competing.

(4) Prosodic Features of Self-repetition

Self-repetition refers to the type 2 and type 3 repetition, that is, the overlapper repeats his or her own words or utterance within or immediately after the overlap. As discussed in (2) and (3), self-repetitions are associated with the outcome of turn-bidding and turn-competing. It is thus worthwhile to go further into its prosodic presentation in terms of successful and unsuccessful turn-bidding and turn-competing.

Here a comparison of the intonation contour was made between the original version and the repeated version by the same speaker. Three types of intonation contour were identified:

(a) the unchanged type, ie. the intonation contour of the repetition is the same as the original version;

(b) the more prominent type, ie. the intonation contour of the repetition becomes more prominent than the original version. It is enhanced by imposing a stress

on the originally unstressed word or imposing a nucleus on a word which is originally only stressed ;

(c) the less prominent type, ie. the intonation contour of the repetition becomes less prominent than the original version. It is downgraded by reducing the stress on the originally stressed word or losing the nucleus on the originally nuclear-patterned word.

The results are displayed in **Table 21**:

Table 21   Intonation Contour of Self-repetition
in TB and TC

Category	Total tokens	Unchanged	More prominent	Less prominent
STB	193	60.1%	24.4%	15.5%
UTB	23	56.5%	21.7%	21.7%
STC	11	36.4%	45.5%	18.2%
UTC	3	66.7%	33.3%	0
Total	230	58.7%	25.2%	16.1%

Table 21 indicates that in general, more than half (58.7%) of the self-repetition instances in turn-bidding and turn-competing retain their original prosodic features as shown in [71] and [75]. However, with regard to the changed types of intonation contour, the more prominent type has a higher proportion than the less prominent type (25.2% > 16.1%). Therefore the conclusion may be that the intonation contour of self-repetition in turn-bidding and turn-competing tends to retain the prior pattern or become more prominent. For example in [69], the onset "^your" is repeated as double-stressed and falling tone "`y\our"; in [70], the

Comparing the changed types of intonation contour of repetitions in terms of successful and unsuccessful turn-bidding and turn-competing, it is quite obvious that repetitions in STB and STC have a higher percentage of more prominent intonation contour than those in UTB and UTC (24.4% > 21.7%, 45.5% > 33.3%). However, with regard to the percentage of less prominent intonation contours, UTB is higher than STB (21.7% > 15.5%). This suggests that repetitions which maintain the original intonation contour, or upgrade to a stronger version, are likely to make turn-bidding or turn-competing successful. The results here are consistent with the results of increasing amplitude during the overlap in the case of STB and STC (see **4.3.1. (3)**).

4.3.4   SUMMARY

The results of this section are summarised as follows:

(a) In general, linguistic devices and strategies associated with simultaneous speech are multi-layered. They involve:

  i) salient prosodic features such as prominent initial sounds, various speech rate and amplitude, and incomplete TU structures at the end of the overlap;

  ii) initial discourse items which are associated with agreement, disagreement, hesitation, exclamation and continuity;

  iii) repetitions which are associated with the beginning, middle and post overlap positions.

(b) Different linguistic devices and strategies are interrelated and combined, ie. the frequently used initial discourse items are usually accompanied by some distinct prosodic features; and the repetitions are often highlighted by a more prominent intonation contour. In other words, it is the combination of these linguistic devices and strategies that contributes to the same discourse function or to the same pragmatic effect. To take STB as an example, the most frequently

occurring initial word "yes" (which makes up 18.8% of 830 STB tokens) is associated with:

    (i) 19.7% of 502 tokens of stressed initials,

    (ii) 35.8% of 254 tokens of nuclear patterned initials,

    (iii) 17.4% of 189 tokens of single-word TU initials,

    (iv) 43.0% of 363 frequently occurring discourse items,

    (v) 15.5% of 193 tokens of self-repetitions within or after the overlap.

To analyse the above findings in another way, among the 156 tokens of "yes" which occur at the beginning of STB,

    (i) 63.5% are stressed,

    (ii) 58.3% are nuclear patterned,

    (ii) 21.2% are a single-word TU,

    (iv) 19.2% are in a repeated form.

All these features concentrated on "yes" enhance its pragmatic function and make it an effective device with which to start turn-bidding, turn-competing and backchannelling. This is consistent with Grice's (1975) cooperative principle and Levinson's (1983) politeness principle in conversation --- two essential principles for successful human communication.

(c) Different categories of simultaneous speech are likely to emphasise particular linguistic devices and strategies. This is clearly displayed by the difference in the relative frequencies of many devices when comparing successful and unsuccessful turn-bidding and turn-competing. The results of the present study show that the overlapped speech of STB and STC is more likely than UTB and UTC to have the following features:

    i) more frequent use of prominent prosodic features at the beginning of the overlaps,

    ii) a faster speech rate during the overlaps,

iii) a higher amplitude during the overlaps

iv) a stronger tendency towards incomplete prosodic units at the end of the overlaps

v) more frequent use of discourse items of agreement, hesitation, exclamation and continuity at the beginning of the overlaps,

vi) more frequent use of repetition within and immediately after the overlaps, and the repetitions are more often accompanied by more prominent intonation contours.

However, BC has by nature the highest frequency of agreement initials such as "yes" and "m", mostly accompanied by prominent prosodic features and ending as a single-word tone unit. Yet it has the lowest frequency of repetition and no occurrence of initial discourse items of disagreement and continuity at all.

4.4 SOCIOLINGUISTIC FEATURES ASSOCIATED WITH SIMULTANEOUS SPEECH

Language behaviour is one aspect of social behaviour and therefore a study of any language phenomenon cannot neglect its social dimensions. In the last part of this chapter, the analysis shifts to the association between simultaneous speech and various social factors. Here the five categories of simultaneous speech are generalised into two:

(a) TB, which includes STB, UTB, STC and UTC, ie. simultaneous speech aiming at biddding for a turn (see rational in **3.3.4.** p.81);

(b) BC, which includes only backchannels, ie. simultaneous speech not aiming at bidding for a turn.

The social factors being analysed in relation to TB and BC are:

(a) Three speech domains, ie. casual conversation, public discussion and telephone conversation.

(b) Two degrees of familiarity between interlocutors, ie. equal conversation and disparate conversation. The corpus compilers defined the relationship between interlocutors as being "equal" or "disparate" depending on the degree of familiarity or intimacy between them. Therefore equal conversation means that interlocutors are familiar with each other while disparate conversation means that interlocutors are strangers to each other (for detail, see **3.3.4**).

(c) Two degrees of relative status of interlocutors, ie. within the disparate conversation, speakers with higher social status and speakers with lower social status.

(d) Speakers' gender, ie. single-gender conversation and mixed-gender conversation; female speakers and male speakers.

The analysis was carried out in three steps with each pair of social variables so as to find out:

(a) whether different social variables are associated with the frequency distribution of TB and BC;

(b) whether different social variables are associated with the overlappers' choice of a unit boundary to start TB and BC;

(c) whether different social variables are associated with the overlappers' use of particular linguistic devices or strategies to produce TB and BC.

Chi-square tests were carried out on the distributional frequency of TB and BC in terms of different social variables. The statistical results are presented in **Tables 22, 26, 29, 32**. A log likelihood ratio statistical test (asympototically the same as chi-square test) was made on the percentage of boundary markers and linguistic devices and strategies of TB and BC in terms of different social variables. The

results are presented in **Tables 24, 25, 27, 28, 30, 31, 33, 34**. In these tables, the p-values are located in the line or under the column entitled "p-value". The degree of freedom (df) and the significant p-value are noted in the line just below each table. Under the column for p-value, those figures which are marked with an asterisk "*" indicate that the frequencies being compared in the table are significantly different, while those p-values which are unmarked indicate that the observed differences between the frequencies being compared in the table are not convincing evidence against chance.

4.4.1 SIMULTANEOUS SPEECH AND SPEECH DOMAIN

(1) Relation between Speech Domain and the Occurrence of TB and BC

Comparing the frequency of TB and BC per thousand words, **Table 22** shows that TB clearly has a higher frequency than BC in all of the three speech domains (17.9 > 11.5, 9.6 > 3.8, 15.3 > 7.9). This is consistent with the finding in **4.1,** which shows that in the simultaneous speech of these three speech domains, participants are making many more turn-bids than backchannels when they are in the position of a listener.

Table 22 **Frequency of TB & BC in Three Speech Domains**

Domain	Total words	TB tokens	Frequency /1000 w.	BC tokens	Frequency /1000 w.
Casual conversation	31134	557	17.9	357	11.5
Public discussion	31827	307	9.6	120	3.8
Telephone conversation	28841	441	15.3	229	7.9
Total	91802	1305	14.2	706	7.7
p-value		0.00*		0.00*	

df = 2,   * = significant at p =< 0.004.

In this comparison of the three speech domains, casual conversation has the highest frequency of both TB and BC (TB: 17.9, BC: 11.5), telephone conversation ranks second (TB: 15.3 , BC:7.9) and public discussion ranks last (TB: 9.6, BC: 3.8). The chi-square test shows that the different frequencies among the three domains are significant. This indicates that the occurrence of simultaneous speech is not only associated with different speakers' speech habits or personal characters, but also with the characteristics of communicative settings. This can be further explained by the different degrees of formality related to each domain. We might expect texts from the speech domain which has a higher degree of formality to have similar characteristics, such as:

(a) a longer speaking turn by every speaker on average, ie. each turn contains more tone units;

(b) a longer tone unit on average, ie. each tone unit contains more words;

(c) a lower frequency of paralinguistic sounds such as laughs, giggles or sighs.

In order to test this assumption, a comparison was made of three texts selected from three speech domains in the subcorpus. They are S2.10, S5.4 and S7.2 (see **Table 2** in p. 52).

Table 23 shows that the degree of formality in the text of casual conversation is lower than that in public discussion. As casual conversation is usually carried out in a less formal situation and is usually unrehearsed, it is characterised by lack of explicitness, which is displayed by the frequent use of anaphoric pronouns, phonologically obscure utterances and incomplete sentence structures (see Crystal & Davy, 1969: 103). This inexplicitness is due to the participants' shared background and knowledge about each other. Therefore, the hearer is able to take a great deal of what the speaker is trying to say for granted. This capacity for predicting is important in simultaneous speech. The text in this domain (S2.10) consists of a conversation between four young people aged between 25 to 30: a merchant banker and his wife, a male computer specialist and a female research

worker (see Greenbaum & Svartvik, 1990: 23). The topics all refer to daily life such as family affairs, living conditions and travelling. The atmosphere is light and happy and the frequency of "laughs" is 25.7 per thousand words. The speaker's average turn length is only 2.7-TU long and the average TU length is only 3.8 words, which means the speakers' turns are short and transferred frequently. All this displays an informal setting which enables the hearers to bid for a turn more easily, thus accounting perhaps for the highest frequency of TB (28.1/1000 words) in this text.

Table 23 __Comparison of Three Texts from Different Domains__

Domain	Casual conversation	Public discussion	Telephone conversation
text code	2.10	5.4	7.2
total turns	549	132	619
total tone units	1462	1390	1414
total tokens of "laughs"	141	10	66
total words	5489	6545	4801
total tokens of TB	154	24	109
total tokens of BC	11	10	34
average turn length	2.7(TU)	10.5(TU)	2.3(TU)
average TU length	3.8(words)	4.7(words)	3.4(words)
frequency of "laughs"	25.7/1000w.	1.5/1000w.	13.7/1000w.
frequency of TB	28.1/1000w.	3.7/1000w.	22.7/1000w.
frequency of BC	2.0/1000w.	1.5/1000w.	7.1/1000w.

Note:
The average turn length = the total number of tone units ÷ the total number of turns.
The average turn unit length = the total word count ÷ the total number of tone units.

However, in a public domain such as the radio discussion in this corpus, the conversation is carried out in an "open to all" situation, leading to a higher degree of formality than in the other two domains. As all public discussions were non-surreptitiously recorded, it meant the participants were aware that their talk was to be heard by a public audience, a setting which could move them in a direction of

formality, such as being more careful in planning what to say and how to say it, caring more about other's face and their own self-image. All this could reduce the occurrence of simultaneous speech. **Table 23** shows that the text in this domain (S5.4) has the fewest occurrences of TB, even though it has the largest word count. It is a radio discussion among three journalists and a lawyer talking about an accident. The discussion was chaired by a radio journalist who for most of the time allocated each turn by addressing the particular speaker by name. Therefore, there were far fewer opportunities for the next speaker to self-select a turn, not to mention fewer opportunities for turn-bidding or turn-competing. Besides, the atmosphere seems much more formal than that in Text S2.10, because the frequency of "laughs" is much lower than in Text 2.10 (1.5/1000 w. < 25.7/1000 w.). Moreover the average speaker's turn is 10.5-TU-long, which is much longer than that in Text S2.10 (10.5 > 2.7), and the average TU length is 4.7 words, which is the longest among the three texts (4.7 > 3.8 > 3.4). This indicates that the speech in this domain is more planned and more formal, resulting in the lowest frequency of TB (3.7/1000 w.) among all three texts.

The situation with telephone conversations is quite different because, although a telephone conversation is carried out between individuals and is not open to the public (all the telephone talks were surreptitiously recorded in this corpus), the occurrence of simultaneous speech is limited by the lack of visibility. The interlocutors could not rely on extralinguistic context such as eye contact, facial expression and body movement in order to understand the referencial meaning from each other. They had to depend solely on audible sounds, which are usually diminished in quality over the telephone. As a result, telephone conversation is likely to have no long pauses but more voiced hesitations, more short utterances and more questioning, responding and interacting patterns (see Crystal & Davy, 1969: 119-120). As any long pause may give a false impression of a mechanical break in communication, speakers and hearers have to present more signals such as

"@m", "uh", "yes" to keep the channel open. Besides, the speakers' turns involve more regular patterns of opening, addressing and closing in telephone conversation. In this sense, this domain may still in some ways create particular opportunities for turn-bidding, for the hearer can make use of the filled-pause marker, the short utterance boundary or the predictable routine patterns as a transition-relevance place to start a new turn. Hence this domain has more occurrences of TB than public discussion, though not as many as face-to-face casual conversation. **Table 23** shows that the text in this domain (S7.2) has the shortest average turn-length among the three texts (2.3 TUs) and the shortest average TU length as well (3.4 words). Furthermore, the speech turns involve many pre-fabricated patterns of opening, addressing and closing routines, thus providing conditions for the hearers' anticipation so that they could find a place to bid for a turn before the current speakers end their turns.

The occurrence of backchannelling in the three texts contrasts with that of turn-bidding: telephone conversation has the most brief backchannels and public discussion the fewest. This is consistent with the characteristics of each setting discussed above. In a telephone conversation, the hearer has to make backchannels frequently so as to signal that the communicative channel is kept open. In a casual face-to-face conversation between close friends or family members, overlapped backchannels could decrease either because each speaker's turn or utterance is comparatively short so that the backchannels may not overlap the current speaker's turn, or because the participants are so familiar with each other that the hearer's attention, support or agreement with the current speaker can be expressed by other means such as facial expressions or physical motion in the communicative setting. In a public discussion, giving backchannels to the current speaker can be a polite way of showing the hearer's attention to the topic being discussed, but this could be limited by the nature of public discussion in which the hearer has to be more active in preparing to air different opinions in his or her own turn.

In summary, the frequency distribution of TB and BC varies according to the degree of formality of the communicative setting. Formality can be reflected by the participant's relationship, the occurrence of paralinguistic features and the speaker's average turn length and tone unit length.

(2)   Relation between Speech Domain and the Overlappers' Choice of Unit Boundary

In order to test whether speakers in the speech domains with different degrees of formality would use unit boundaries with different frequencies for making TB or BC, a comparison was made between three speech domains in terms of their percentages of three levels of unit boundary markers (prosodic, lexical and syntactic). The results are presented in **Table 24** below.

Table 24   Percentage of TB & BC which are Immediately Preceded by Boundary Markers in Three Domains

Domain	TB			p-value	BC			p-value
	casual	public	tele.		casual	public	tele.	
Total tokens	479	281	403		357	120	229	
Prosodic tokens	333	258	401		299	106	228	
Prosodic boundary	38.4	45.0	52.4	0.00*	54.9	50.0	58.3	0.35
Lexical boundary	23.8	15.3	38.2	0.00*	18.5	13.3	22.7	0.09
Syntactic boundary	72.4	74.7	81.4	0.01*	75.9	77.5	79.5	0.60

df = 2,   * = significant at p =< 0.05
Note:
In counting the percentage of prosodic boundary markers, only those prosodically transcribed instances were taken into account. Their total number of tokens in each domain is displayed in the fourth row of the table named "prosodic tokens". The same pattern is in **Tables 27, 30** and **33**.

Table 24 shows that in a more formal conversational setting like public discussion, the hearers have a significantly higher percentage of choosing a tone unit boundary and a syntactic boundary for turn-bidding than those in a less formal conversational setting such as casual conversation (45%> 38.4%; 74.7% >

72.4%). This indicates that in a formal conversation, the turn-bidders may adopt a more polite way of turn-bidding, because interrupting someone's speech at a unit boundary is less disruptive than an interruption at a non-unit boundary. However, in telephone conversation, the three levels of boundary markers have the highest percentage (52.4%, 38.2% and 81.4%), though this may not be the domain with the highest degree of formality among the three. A possible explanation for such a result is that in a non-face-to-face conversation which lacks visual contact with the current speaker, the hearer has to rely more on the explicit boundary markers in order to identify a transition-relevance place for making TB. Though the chi-square test did not show a significant effect for the three level boundary markers in the cases of BC, the observed figures shows these boundary markers nevertheless have the highest percentage in the BC tokens of telephone conversation than the other two domains (58.3%, 22.7% ,79.5%).

(3) Relation between Speech Domain and the Overlappers' Linguistic Devices and Strategies

Here the analysis of linguistic devices and strategies discussed in **4.3** is again concentrated on the most salient features of TB and BC according to three aspects (prosodic devices, discourse items and repetition strategies) to find out if the speakers in different speech domains prefer to use particular devices or strategies. The results are presented in **Table 25**.

Table 25 shows that in the domain of casual conversation, there are significant differences in some linguistic devices and strategies of turn-bidding.

(a) TB has the highest percentage of disagreement initials (5.2%), but the lowest percentage of hesitation initials (8.8%) and continuity initials (4.0%). The first two features may be due to the nature of casual conversation, in which participants may engage in a less formal way of talking, ie. they may introduce TB with a direct disagreement or in a less hesitant way. The last feature may be

due to the short utterance structure and the lack of preparedness in this kind of

conversation with the resulting lower occurrence of continuity initials.

Table 25 **Percentages of Linguistic Devices & Strategies**

Associated with TB & BC in Three Domains

Domain	TB			BC			p-value	
	casual conv.	public disc.	tele. conv.	casual conv.	public disc.	tele. conv.	TB	BC
Total tokens	557	307	441	357	120	229		
Prosodic device								
stressed initial	59.7	45.5	65.5	91.0	72.5	88.2	0.00*	0.00*
unclear case	26.8	26.1	22.7	11.5	10.8	7.9	0.31	0.34
incomplete TU	37.2	41.7	33.8	5.0	26.7	10.5	0.09	0.00*
Discourse item								
agreement	15.4	10.1	17.2	78.2	64.2	62.5	0.02*	0.00*
disagreement	5.2	4.6	0.9	0	0	0	0.00*	
hesitation	8.8	11.1	10.2	3.1	16.7	6.1	0.00*	0.00*
exclamation	5.2	5.9	3.4	5.6	5.0	8.7	0.22	0.26
continuity	4.0	8.5	5.4	0	0	0	0.03*	
Total	38.6	40.1	37.1	86.9	85.9	77.3	0.72	0.01*
Repetition								
type 1 repetition	6.8	5.2	8.6	0	0	0	0.19	
type 2 repetition	19.2	18.6	12.0	10.6	6.7	4.4	0.01*	0.02*
type 3 repetition	21.0	18.9	17.9	0	0	0	0.45	
Total	47.0	42.7	38.5	10.6	6.7	4.4	0.03*	0.02*

df=2, * = significant at p =< 0.05

Note:
Type 1 repetition: the overlapper repeats the last words or sentence pattern of the current speaker at the beginning of the overlap;
Type 2 repetition: the overlapper repeats his or her own words or sentence pattern during the overlap;
Type 3 repetition: the overlapper repeats his or her own words or sentence pattern immediately after the overlap. These are the same in **Tables 28, 31** and **34**.

(b) TB also has the highest percentage of repetition (47.0%) and the type 2

repetition in particular (19.2%). The more frequent use of type 2 repetition may

display the turn-bidders' informal speech style or their hesitant way in casual conversation.

The linguistic devices and strategies associated with backchannelling in casual conversation are characterised by

(a) the highest percentage of stressed initials (91.0%), yet the lowest percentage of incomplete TU endings (5%);

(b) the highest percentage of agreement initials (78.2%) but the lowest percentage of hesitation initials (3.1%);

(c) the highest percentage of the type 2 repetition (10.6%).

This means the frequent occurrence of agreement initial in BC such as "yes" and "m" are mostly accompanied by prominent prosodic features and repetitions.

In the domain of public discussion, the TB devices and strategies which are significantly different from the other two domains are as follows:

(a) TB has the lowest percentage of stressed initials (45.5%);

(b) TB also has the lowest percentage of agreement initials (10.1%) but the highest percentage of hesitation initials (11.1%) and continuity initials (8.5%).

The lesser occurrence of agreement initials in this domain may be consistent with the nature of argument usually occurring in public discussion. That is, the interlocutors are encouraged to raise different opinions instead of agreeing with each other during the discussion. Yet the turn-bidders in this domain may still try to make their TB less abrupt or sound more cooperative, and hence they use more hesitation initials. This is also displayed in the BC devices used in this domain. They have a lower occurrence of agreement initials (64.2% < 78.2%) than in the domain of casual conversation and the highest percentage of hesitation initials (16.7%) amongst the three domains. Moreover, these discourse items are accompanied by the fewest stressed initial sounds (72.5%).

In the domain of telephone conversation, the TB devices and strategies which are significantly different from the other two domains are as follows:

(a) TB has the highest percentage of stressed initial sounds (65.5%) which indicates that in a non-face-to-face conversation, turn-bidders tend to start the bidding with more prominent sounds so as to go through the mechanical channel of telephone conversation.

(b) TB has the highest percentage of agreement initials (17.2%), but the lowest percentage of disagreement (0.9%). Yet the frequent occurrence of initial items like "yes", "yeah", "@m" does not necessarily mean the turn-bidders are always involved in agreements with the current speakers. It just shows that these words are especially needed in telephone conversation, because hearers have to use such backchannels to signal that the channel is open and that the hearers are still on the line, listening to the speakers.

(c) TB also has the lowest percentage of repetition as a whole (38.5%) with a particularly low level of type 2 repetition (12%).

The BC devices and strategies in this domain have the lowest use of five sets of discourse items (77.3%), particularly in agreement initials (62.5%). They also have the lowest percentage of the type 2 repetition (4.4%). Telephone conversation is likely to reveal frequent use of words like "yes" or "m" in a repeated way. Yet the result here indicates only that such frequently used backchannel sounds may not occur in simultaneous speech as often as they do in the other categories.

In summary this analysis suggests that there are some major differences between the three speech domains because of the different degrees of formality. First, the most formal speech domain, such as public discussion, has the lowest frequency of TB and BC, while the least formal speech domain, ie. casual conversation, has the highest frequency of TB and BC. Second, the overlappers in public discussion

have a higher percentage of making TB at prosodic and syntactic boundaries than in casual conversation. But the overlappers in telephone conversation have the highest percentage in this aspect. Third, the overlappers in public discussion , though they do not use as many prominent prosodic features and agreement initials to introduce TB as in the other two domains, have the highest percentage of hesitation initials and continuity initials for TB. However, the overlappers in casual conversation tend to use the repetition strategies more frequently in both TB and BC, whereas the overlappers in telephone conversation tend to use the stressed agreement initials more frequently to introduce TB. However they use fewer repetitions in TB and BC than the other two domains.

4.4.2 SIMULTANEOUS SPEECH AND DEGREE OF FAMILIARITY BETWEEN INTERLOCUTORS

(1) Relation between the Familiarity of Interlocutors and the Occurrence of TB and BC

In **4.4.1** the association between simultaneous speech and the degree of formality in communicative settings was discussed. In the present section, the formality is further associated with the degree of familiarity between speakers and hearers.

Table 26   Frequency of TB & BC in Equal & Disparate Conversations

Relationship	Total tokens	TB tokens	Frequency/ 1000 words	BC tokens	Frequency/ 1000 words
Equal	47223	861	18.2	406	8.6
Disparate	44579	444	10.0	300	6.7
Total	91802	1305	14.2	706	7.7
p-value		0.00*		0.00*	

df=1,   * = significant at p =< 0.004

In **Table 26**, the observed differences in the frequency of TB and BC between equal conversation and disparate conversation are quite obvious. The differences are further confirmed to be significant by the chi-square test of which the p-value is below .004. First, the frequency of TB per thousand words in equal conversation is much higher than in disparate conversation (18.2 > 10). Such a difference may be due to the different degrees of familiarity between the interlocutors in these two kinds of conversation. For example, in a conversation where the two interlocutors are not familiar with each other, they would typically share much less common background than would familiar interlocutors. This could hamper the hearer's anticipation during the communication, a mechanism which was discussed in **4.2** and **4.3**, and which is crucial for turn-bidding. In other words, the hearer who is a stranger to the current speaker may find it difficult to anticipate the possible completion of the current speaker's utterance. He or she may try to keep a profile of being polite or respecting the unfamiliar interlocutor by making less turn-bidding, and hence a much lower frequency of TB in disparate conversation. This result indicates that degree of familiarity between interlocutors greatly affects the occurrence of TB in conversation. People who are familiar with each other may feel more comfortable in bidding for a turn with each other, while people who are strangers to each other may be more hesitant in doing so. Dunne & Ng (1994: 47) point out:

> ... listening to the speech of unfamiliar people (strangers) is cognitively more demanding than listening to the speech of familiar people (friends). This is because with strangers one has little or no idea of what to expect and as a result has to pay more attention to what they are saying, thus reducing available cognitive resources to plan and execute one's own utterance.

The present results support this point of view, because interlocutors in disparate conversations seem to anchor their behaviour more in a one-speaker-at-a-time mode by making much less TB. Yet this is not necessarily so in making BC, for **Table**

26 shows that the difference in BC frequency between equal and disparate conversation (8.6 > 6.7) is not as big as the difference in TB frequency.

(2)   Relation between the Familiarity of Interlocutors and the Overlappers' Choice of Unit Boundary

Table 27 shows that there is no significant difference between equal and disparate conversation in terms of the percentages of unit boundary markers occurring before TB and BC. For example, 44.8% of the TB in equal conversation occurs at prosodic boundaries and 47.7% of the TB in disparate conversation also has this feature. The chi-square test shows that these observed percentage differences are not statistically significant, because the p-value is far beyond 0.05 (p = 0.38). **Table 27** shows this is also the case with other boundary features and in both TB and BC cases. These results may be interpreted as evidence that whether the interlocutors are familiar with each other or not, they seem to choose a unit boundary to start TB and BC in a similar frequency.

Table 27   **Percentage of TB & BC Immediately Preceded by Boundary Markers in Equal & Disparate Conversations**

Relationship	TB		p-value	BC		p-value
	equal	disparate		equal	disparate	
Total tokens	749	414		406	300	
Prosodic tokens	663	329		387	247	
Prosodic boundary	44.8	47.7	0.38	54.5	56..3	0.66
Lexical boundary	27.4	25.6	0.51	20.4	17.0	0.24
Syntactic boundary	75.0	78.0	0.25	77.1	77.7	0.81

df=1, * = significant at p =< 0.05

(3) <u>Relation between the Familiarity of Interlocutors and the</u>
<u>Linguistic Devices & Strategies Used by Overlappers</u>

As **Table 28** shows, the TB devices and strategies used by turn-bidders in both equal conversation and disparate conversation do not differ greatly. The chi-square test shows that most of the p-values are above 0.05, except for the features marked with "*". This means that interlocutors in both equal conversation and disparate conversation are likely to use the same linguistic devices and strategies in producing TB.

However, a closer look at **Table 28** shows that interlocutors in disparate conversation have:

(a) a significantly lower percentage of unclearly-heard cases (19.8% < 28%);

(b) a significantly lower percentage of hesitation initials (3.2% < 8.9%);

(c) a significantly higher percentage of repetitions overall (48% > 40.6%), though the chi-square test shows no significant difference in each respective type of repetition.

All these features indicate that turn-bidders in disparate conversation tend to have a higher amplitude during the overlap (with evidence of less occurrence of unclearly-heard cases) and a firmer attitude in starting turn-bidding (with evidence of less occurrence of hesitation initials). The hearers actively anticipate the current speakers' oncoming words or patterns and reiterate their own words or patterns during and after the overlap. This is consistent with the atmosphere in disparate conversation in which the speaker is trying to take an active part in communication, but cooperative communication is still maintained between the unacquainted partners.

**Table 28   Percentages of Linguistic Devices & Strategies
Associated with TB & BC in Equal & Disparate Conversations**

Relationship	TB		BC		p-value	
	equal	disparate	equal	disparate	TB	BC
Total occurrence	861	444	406	300		
Prosodic						
stressed initial	58.7	57.3	87.7	86.0	0.62	0.51
unclear case	28.0	19.8	12.1	7.7	0.00*	0.05*
incomplete TU	36.7	37.8	8.9	12.7	0.69	0.11
Discourse item						
agreement	13.6	17.3	65.3	78.0	0.46	0.00*
disagreement	3.4	4.1	0	0	0.56	
hesitation	8.9	3.2	5.2	8.0	0.00*	0.01*
exclamation	5.5	3.4	8.6	3.7	0.15	0.01*
continuity	4.8	7.0	0	0	0.10	
Total	36.2	35.0	79.1	89.7	0.67	0.00*
Repetition						
type 1 repetition	6.4	8.3	0	0	0.20	
type 2 repetition	16.1	17.6	7.9	8.0	0.51	0.95
type 3 repetition	18.1	22.1	0	0	0.09	
Total	40.6	48.0	7.9	8.0	0.01*	0.95

df=1,   * = significant at $p =< 0.05$

Note:
Type 1 repetition: the overlapper repeats the last words or sentence pattern of the current speaker at the beginning of the overlap;
Type 2 repetition: the overlapper repeats his or her own words or sentence pattern during the overlap;
Type 3 repetition: the overlapper repeats his or her own words or sentence pattern immediately after the overlap.

Comparing the BC devices and strategies in these two kinds of conversation, the BC tokens in disparate conversation have a higher amplitude during the overlap, because the percentage of unclearly-heard cases is lower than that in the equal conversation (7.7% < 12.1%). They have a significantly greater use of the five sets of discourse items as initials (89.7% > 79.1%) and in particular a higher percentage of agreement initials (78% > 65.3%) and hesitation initials (8% > 5.2%). This can be interpreted to mean that in a conversation among unfamiliar people, the hearers

may pay more attention to the speech of others and give positive backchannels more frequently so as to display a more formal and polite manner in communication. However, the BC in equal conversation has a significantly higher percentage of exclamation (8.6% > 3.7%). It may indicate an informal atmosphere among familiar interlocutors in conversation, because when exclamation markers like "ah" and "oh" are used alone, they may indicate strong emotional states (see Schiffrin, 1987: 73), while strongly expressed emotional states are less likely to occur in conversation between strangers. This may possibly explain why in disparate conversation, such exclamation markers are fewer than those in equal conversation. From another angle it also suggests that interlocutors may be aware of the degree of familiarity during the conversation.

4.4.3 SIMULTANEOUS SPEECH AND THE STATUS OF INTERLOCUTORS

The identification of "higher status" (H) speakers and "lower status" (L) speakers was based on the LLC compilers' brief description of the participants' professional ranking and age. It may be a very rough classification, but the data from this corpus still provides some interesting information.

(1) Relation between Interlocutors' Status and the Occurrence of TB and BC

On examining the total tokens, the TB and BC of H speakers and L speakers are both significantly different, ie. H speakers have more TB (205 > 186) but less BC (95 < 168). Yet we cannot simply conclude that H speakers have a higher TB frequency, or that L speakers have a higher BC frequency until further analysis is done in terms of frequency per thousand word.

Table 29 <u>**Frequency of TB and BC Made by**</u>
<u>**High Speakers & Low Speakers**</u>

	Total words	TB tokens	Frequency/ 1000 words	BC token	Frequency/ 1000 words
	44579	391	8.8	263	5.9
Method 1					
High	28347	205	12.6	95	5.9
Low	16232	186	6.6	168	5.9
Method 2					
High	28347	205	7.2	95	3.4
Low	16232	186	11.5	168	10.3
p-value		0.00*		0.00*	

df=1, * = significant at $p = < 0.004$
Note:
The tokens of TB and BC in this table only take account of those instances which involve the overlap between H speakers and L speakers, ie. excluding those H speaker-to-H speaker or L speaker-to-L speaker overlapping instances which may occur in a multi-party disparate conversation.

By using two methods in counting the frequency per thousand words, two contrasting results can be obtained . **Table 29** shows that the total word count of H speakers is larger than that of L speakers (28347 > 16232). By method 1, ie. to divide respectively the total TB tokens and BC tokens of H speakers by the total word count of L speakers, the TB frequency of H speakers is much higher than that of L speakers (12.6 > 6.6) and the BC frequency is the same for both H and L speakers (5.9=5.9). However by method 2, ie. to divide respectively the total TB tokens and BC tokens of H speakers by the total word count of H speakers themselves, the TB and BC frequencies of H speakers are both much lower than those of L speakers (TB: 7.2 < 11.5, BC: 3.4 < 10.3). The results of method 1 may indicate that H speakers are likely to speak more in disparate conversation and interrupt more often the speech of L speakers. But the results of method 2 also need to be considered. It is not unusual to see L speakers making more backchannels when they are listening to H speakers, because giving more attention signals or agreement feedbacks such as "yes" and "uhm" may indicate that L

speakers adopt a more polite or cooperative attitude in disparate conversation. But the higher frequency of TB by L speakers in comparison to H speakers was beyond the present researcher's expectation, because both turn-bidding and turn-competing in respect of H speakers' speech are likely to be seen by native speakers of Chinese as face-threatening acts, and L speakers may prefer to avoid them. Yet in the subcorpus, the TB frequency of L speakers is even higher than that of H speakers. One explanation could be that in native English speakers' conversation, participants may not be aware of social status as much as they are aware of the personal relationship between the speaker and the hearer. In other words, the different degrees of familiarity or intimacy between interlocutors tends to influence the occurrence of TB more than the status difference between interlocutors. A higher TB frequency of L speakers may indicate their active participation in the conversation despite the possible status difference.

Text S3.1 is an extract of casual conversation. This is a 5080-word text with 148 tokens of simultaneous speech, of which 55% of the TB and 68% of the BC are made by an L speaker --- a 20 year-old prospective female undergraduate --- and 45% of the TB and 32% of the BC are made by a H speaker --- a 40 year old academic staff member in the university. This is an academic interview, but the atmosphere seems quite casual, for there are 37 instances of "laughs, giggles, sniffs, coughs" and "sighs" transcribed all through the interview. The young girl student seems to be actively involved in the conversation, for the transcribed extract shows that she seldom hesitates in making turn-bidding and only six out of her 51 instances of TB are initiated by hesitation markers like "well" or "@:". She seems always ready to answer the teacher's question, not even waiting for it to be completely finished. eg.

[77]   (a = teacher, A = student)

> a: you're not very attracted to the modern English novel. %*perhaps you
>   have been reading . novels of earlier in this century*
> A: **&*I'm ^n/\ot# / I'm ^so .`very much beh\ind#*$** / ^in . ^in
>   novels `"\Vearlier#. / ^in you ^kn/ow# / in the ^nineteen h\Vundreds# . /
> a: yes have you been reading a lot of them recen %*tly*
> A: **&*I*$** ^h/ave# / I've been ^reading a `lot of L\Vawrence# . / I've been
>   ^trying to read m\ost of the`works of L/awrence# . /
> a: in the early nineteen hundreds yes %*I see*
> A: **&*^y\es#*$** / ^y\es# /

(LLC, S3.1: 1211-1222)

Here, in speaker A's first STB, she seems so eager to express her idea that she has a long overlap with the teacher's speech. In her second STB, she starts answering the question even before the teacher has actually finished it. In her third STB, she loses no time in showing her agreement with the teacher's words, which in fact is a completion by the teacher of the student's prior statement. The coherent development of this discourse does not seem to be hampered by the frequent occurrence of STB, but on the contrary, it smooths the turn-shifting during the conversation. In a disparate conversation like this, the student may assume that an active participation and a full display of her knowledge and talent will be more important than taking a conservative attitude and keeping at a distance from the teacher because of the status difference.

Another reason for the higher frequency of TB in L speakers' speech may be that in the subcorpus, the total word count of L speakers is much less than that of H speakers ($16232 < 28347$). This could be interpreted as showing that L speakers have fewer opportunities to speak, while H speakers seem to talk much more in a given time, and therefore L speakers have to make more TB to gain the floor. Some previous research findings (see Ng. et al. 1993) show that the power or influential ranking of a speaker in conversation is represented by the total number of turns taken in a given time which is reliably predicted by their total number of

interruptions. The present study finds that a speaker's power or influence in conversation may be represented by his or her total word count in the conversation but not necessarily by his or her interruption frequency. As the present study is interested in exploring hearers' response strategies, the results obtained by method 2 may be more relevant and significant. As noted in **3.3.4. (2),** method 2 for calculating frequency of occurrence involves dividing the total number of TB tokens of one type of speakers by the total word count of these speakers themselves.

(2) Relation between Interlocutors' Status and the Overlappers'
Choice of Unit Boundary

Table 30   Percentage of TB & BC of High & Low Speakers
Immediately Preceded by Boundary Markers

Status	TB			BC		
	high	low	p-value	high	low	p-value
Total tokens	237	177		132	168	
Prosodic tokens	214	115		57	82	
Prosodic boundary	46.3	50.4	0.47	50.4	61.2	0.09
Lexical boundary	23.6	28.2	0.29	21.2	13.7	0.09
Syntactic boundary	82.3	75.1	0.22	72.7	81.0	0.07

df=1,   * = significant at p =< 0.05

Table 30 shows there is no significant difference between H and L speakers in terms of the percentage of TB immediately proceeded by unit boundary markers, for all the p-values in this column are well beyond 0.05. Yet the observed difference shows that L speakers' TB has a higher percentage of beginnings at a prosodic boundary and a lexical boundary (50.4% > 46.3%; 28.2% > 23.6%), but it has a lower percentage of beginnings at a syntactic boundary (75.1% < 82.3%).

Although the chi-square test fails to show a significant difference between H and L speakers in terms of the boundary features associated with BC, the p-values (p=0.09, p=0.09, p=0.07) are much less than those in the TB column. This may

indicate that in making BC, L speakers tend to start at a prosodic boundary and a syntactic boundary more often than do H speakers (61.2% > 50.4%, 81% > 72.7%). As a whole, L speakers still seem to choose a unit boundary to start TB and BC more frequently than H speakers do. It might be interpreted that L speakers seem to be more cautious in simultaneous speech, because beginning to speak at a unit boundary may sound less disruptive to the current speaker whose status is higher than the overlapper's.

(3)   Relation between Interlocutors' Status and the Linguistic Devices & Strategies Used by Overlappers

In **Table 31,** the chi-square test does not show a significant difference between H and L speakers in terms of the linguistic devices and strategies associated with TB, because most of the p-values are more than .05, except two cases marked with an asterisk. This indicates that in general there is no close association between speakers' social status and their linguistic strategies for producing TB.

Nevertheless, a few differences do exist in some linguistic features, such as the following:

(a) Prosodically, L speakers' TB has a significantly lower degree of clarity in speaking than H speakers' TB, for it has more occurrences of unclearly-heard cases within the overlap (24.2% > 16.1%). Such a difference in degree of clarity may indicate that the higher amplitude of H speakers in their speech (especially during the overlapped speaking time) can be evidence of H speakers' dominance in disparate conversation, because the same tendency also occurs in H speakers' BC overlaps (10.7% > 3.8%).

Table 31 <u>Percentages of Linguistic Devices & Strategies
Used by High Speakers & Low Speakers</u>

Status	TB		BC		p-value	
	high	low	high	low		
Total tokens	205	186	95	168	TB	BC
Prosodic device						
stressed initial	56.0	57.8	79.5	91.1	0.78	0.00*
unclear case	16.1	24.2	3.8	10.7	0.05*	0.02*
incomplete TU	36.1	37.1	18.2	7.7	0.83	0.01*
Discourse item						
agreement	17.3	17.4	71.2	83.3	0.36	0.01*
disagreement	1.6	7.4	0	0	0.01*	
hesitation	3.9	4.2	7.6	8.3	0.79	0.81
exclamation	2.8	4.2	2.3	4.8	0.65	0.24
continuity	7.5	5.9	0	0	0.21	
Total	33.1	39.1	81.1	96.4	0.81	0.00*
Repetition						
type 1 repetition	8.9	8.9	0	0	0.95	
type 2 repetition	20.0	19.4	11.4	5.4	0.87	0.06
type 3 repetition	24.5	25.9	0	0	0.74	
Total	53.4	54.2	11.4	5.4	0.91	0.06

df=1,　　　* = significant at p= < 0.05

Note:
Type 1 repetition: the overlapper repeats the last words or sentence pattern of the current speaker at the beginning of the overlap;
Type 2 repetition: the overlapper repeats his or her own words or sentence pattern during the overlap;
Type 3 repetition: the overlapper repeats his or her own words or sentence pattern immediately after the overlap.

(b) In terms of discourse items, the observed figures show that L speakers have a more frequent use of the five sets of discourse initial items as a whole (39.1% > 33.1%) and also have a higher frequency in each set of initials except the continuity set, even though the chi-square test does not show a significant difference for each set separately, nor does it show it as a whole. As initials of agreement, hesitation and exclamation are all positive devices for turn-bidding, this may suggest the active participation of L speakers in disparate conversation.

However the higher percentage of disagreement initials in L speaker's TB (7.4% > 1.6%) does not necessarily mean that L speakers often start turn-bidding by disagreeing with H speakers' statements. It could be due to the fact that H speakers ask more yes/no questions. As a consequence, L speakers are more often required to answer H speakers by saying either "yes" or "no". Such cases make up 43% of the disagreement instances in L speakers' TB.

(c) In terms of repetition, no significant difference is reported by the chi-square test for each type of repetition. This indicates that both H speakers and L speakers used these strategies with a similar frequency.

In terms of the BC devices and strategies used by H speakers and L speakers, some significant differences are shown by the results of the chi-square test in **Table 31** as follows:

(a) L speakers' BC shows a significantly higher percentage of stressed initials (91.1% > 79.5%) and unclearly-heard cases (10.7% > 3.8%) but a significantly lower percentage of incomplete TU at the end of the overlap (7.7% < 18.2%);

(b) L speakers' BC is more frequently introduced by the five sets of discourse items (96.4% > 81.1%), particularly by agreement initials (83.3% > 71.2%).

These features can be interpreted as evidence that L speakers are more likely to initiate backchannels to H speakers' speech by words with prominent prosodic features, but their amplitude during the overlaps is not as high as that of H speakers. Moreover, L speakers tend to make better use of the five sets of discourse items to introduce BC, especially the items which express agreement.

4.4.4   SIMULTANEOUS SPEECH AND GENDER

(1)   Relation between Interlocutors' Gender and the Occurrence of TB and BC

Table 32   Frequency of TB & BC Made by Females and Males

(1) Single-gender Conversation:

Gender	Total words	TB tokens	Frequency/ 1000 words	BC tokens	Frequency/ 1000 words
F-F	8058	93	11.5	59	7.3
M-M	25317	365	14.4	268	10.6
p-value		0.06		0.01*	

df = 1,   * = significant at p =<0.05

(2) Mixed-gender Conversation: (total words = 58427) (Method 1)

Gender	Total words	TB tokens	Frequency/ 1000 words	BC tokens	Frequency/ 1000 words
Female	25411	340	10.3	202	6.1
Male	33016	343	13.5	108	4.3
p-value		0.00*		0.00*	

df = 1,   * = significant at p =< 0.004

(3) Mixed-gender conversation: (total words = 58427) (Method 2)

Gender	Total words	TB tokens	Frequency/ 1000 words	BC tokens	Frequency/ 1000 words
Female	25411	340	13.4	202	7.9
Male	33016	343	10.4	108	3.3
p-value		0.00*		0.00*	

df=1,   * = significant at p =< 0.004

Note:

The tokens of TB and BC in mixed-gender conversations in this table account only for those instances which are involved in the overlap taking place between a female speaker and a male speaker, ie. excluding those female-to-female or male-to-male overlap instances which may occur in a multi-party mixed-gender conversation.

Section (1) of **Table 32** shows that in single-gender conversations, the male-to-male conversations have a higher frequency per thousand words of both TB and BC than the female-to-female conversations (TB: 14.4>11.5, p=0.06, BC:

10.6>7.4, p=0.01) though the difference of TB frequency is not highly significant by the chi-square test. These results are inconsistent with the previous research, which found significantly more female interruptions in single-gender conversation (see Crosby: 1976, Street & Murphy: 1987 and Bilous & Krauss: 1988, cited by James & Clarke, 1993:236). Yet in mixed-gender conversations, **Table 32** (2) and (3) show that both TB and BC occurrence has significant differences in terms of gender, ie. females have fewer TB tokens (340 < 343) but more BC tokens (202 > 108) than males. A further comparison can be made in terms of frequency per thousand words.

In counting the frequency per thousand words, another pair of contrastive results is revealed by using two methods. As **Table 32** shows, in mixed-gender conversations, the males' total word count is larger than that of females (33016 > 25411). By method 1, (ie. to divide respectively the total TB tokens and BC tokens of males by the total word count of females) **Table 32 (2)** shows the TB frequency of males is higher than that of females (13.5 > 10.3) but the BC frequency of males is lower than that of females (4.3 < 6.1). However by method 2, (ie. to divide respectively the total TB tokens and BC tokens of females by the total word count of the females themselves) **Table 32 (3)** shows that it is the females that have a higher frequency of both TB and BC (TB: 13.4 > 10.4 , BC: 7.9 > 3.3). Moreover, females' TB and BC frequencies in mixed-gender conversations are even higher than those in all-female conversations (TB: 13.4 > 11.5, BC: 7.9 > 7.4, see **(3)** and **(1)** in **Table 32**). We may describe the picture in this way: in mixed-gender conversations, females tend to make more BC than males, whether it is counted by method 1 or by method 2. This supports the previous research which shows that females are more supportive hearers than males (see Coates: 1989, Holmes: 1991). Their studies show that females tend to express agreement and awareness of the current speakers more frequently than males do and females tend to support each other's comments, confirming, elaborating and

developing points made by others. The present study again supports their findings and shows that such a tendency is even stronger in mixed-gender conversations than in all-female conversations.

In terms of TB frequency in mixed-gender conversation, the results of method 1 support the previous studies of males' conversational dominance, which say that males tend to speak more than females and interrupt females' speech more often (see Zimmerman & West: 1975, West: 1979, West & Zimmerman: 1983). But this is only one side of the picture. As shown in section (2) and section(3) of **Table 32,** although females' total word count has 7605 words less than males' (33016 - 25411), females' TB is only three instances less than that of males (343-340). Therefore when counting the frequency by method 2, females' TB frequency per thousand words is even higher than that of the males' (13.4 > 10.4). This may be interpreted to mean that since females have fewer opportunities to speak than males, they are likely to engage in more turn-bidding or turn-competing in order to gain the floor, thus resulting in a higher frequency of female TB than male TB. In this sense, male's dominance in mixed-gender conversations could be represented by a larger total word count, but not necessarily by a higher frequency of turn-bidding.

(2)   Relation between Interlocutors' Gender and the Overlappers' Choice of Unit Boundary

Table 33 shows that females' TB has a significantly lower percentage of starting at a prosodic boundary, but a higher percentage of starting at a lexical boundary than males' TB (41% < 48.5%; 31.5% > 24%). No significant difference is found in the case of syntactic boundary features (75.6% < 76.1). This might be interpreted to mean that female speakers may be keener on choosing discourse signals as a chance to bid for a turn. Yet both females and males choose syntactic boundaries in turn-bidding with a similar frequency, for such boundary features appear in the majority of TB instances for both females and males. **Table 33** also

shows that females tend to insert backchannels more often at the prosodic and syntactic boundaries than do males (61.3% > 52%, 84.8% > 76.8%).

Table 33   **Percentage of TB and BC of Females and Males Immediately Preceded by Boundary Markers**

Gender	TB			BC		
	female	male	p-value	female	male	p-value
Total tokens	413	750		270	436	
Prosodic tokens	354	639		217	417	
Prosodic boundary	41.0	48.5	0.02*	61.3	52.0	0.03*
Lexical boundary	31.5	24.0	0.01*	18.9	19.0	0.96
Syntactic boundary	75.6	76.1	0.82	84.8	76.8	0.01*

df=1, * = significant at p =< 0.05

(3)   Relation between Interlocutors' Gender and the Linguistic Devices and Strategies Used by Overlappers

In analysing the devices and strategies for TB and BC particularly preferred by each gender, all tokens of TB and BC made by females and by males were taken into account respectively, whether they occur in a mixed-gender conversation or in a single-gender conversation. The results are presented in **Table 34**.

The statistical analysis in **Table 34** shows some significant preferences appearing in females' and males' use of TB devices. These include:

(a) Prosodically, females' TB has a lower percentage of unclearly-heard cases during the overlap (16.4% < 30.3%) and a lower percentage of incomplete TU structure at the end of the overlap (31.5% <40.3%). This may mean that females are likely to use a higher amplitude to attract the current speakers' attention (or to show their intentions to turn-bid). A higher percentage of complete TU structure at the end of the overlap might, however, indicate that females may not have a strong inclination to speak continuously after their turn-bidding is overlapped with the males' speech.

**Table 34 Percentages of Linguistic Devices & Strategies
Used by Females & Males**

Gender	TB female	male	BC female	male	p-value TB	BC
Total tokens	476	829	270	436		
Prosodic device						
stressed initial	60.6	57.1	90.0	85.1	0.22	0.06
unclear case	16.4	30.3	8.2	10.5	0.00*	0.15
incomplete TU	31.5	40.3	8.1	9.6	0.00*	0.50
Discourse item						
agreement	14.7	15.1	68.2	72.5	0.86	0.22
disagreement	4.2	3.3	0	0	0.38	
hesitation	2.9	4.1	5.6	6.9	0.14	0.48
exclamation	5.5	4.3	8.5	5.3	0.36	0.09
continuity	3.8	6.5	0	0	0.03*	
Total	31.1	33.3	82.3	84.6	0.32	0.40
Repetition						
type 1 repetition	8.4	6.3	0	0	0.15	
type 2 repetition	11.8	19.4	4.1	10.3	0.00*	0.00*
type 3 repetition	15.8	21.6	0	0	0.01*	
Total	36.0	47.3	4.1	10.3	0.00*	0.00*

df=1, * = significant at p =< 0.05

Note:
Type 1 repetition: the overlapper repeats the last words or sentence pattern of the current speaker at the beginning of the overlap;
Type 2 repetition: the overlapper repeats his or her own words or sentence pattern during the overlap;
Type 3 repetition: the overlapper repeats his or her own words or sentence pattern immediately after the overlap.

b) In terms of discourse items, though not supported by a significant p-value, the observed figures still show that females' TB tends to use slightly more exclamation initials (5.5% > 4.3%) while males' TB uses significantly more continuity initials (6.5% > 3.8%) and a few more hesitation initials (4.1% > 2.9%). The more frequent use of "oh" and "ah" in females' TB and BC may indicate that females tend to make good use of the pragmatic function of exclamation markers. As noted by Schiffrin (1987: 99), the exclamation

markers like "oh" and "ah" can profile a hearer as an active recipient of information who acknowledges and integrates information as it is provided and who tries to maintain the alignment between speakers and listeners.

However, the fact that males' TB has used more continuity initials does not necessarily mean that males have a special preference for using connectors like "and" to start TB. A further examination of 51 instances of males' TB initiated by connectors reveals that 41.2% of continuity initial instances are presented by the word "but", whereas among 24 continuity initials by females, only 20.8% use the word "but". As "but" is a connector usually used to express a contradictory idea to the prior statement, this result could indicate that males are more likely to bid for a turn by starting with a challenge, while females are more likely to use the developmental connector "and". This, from another angle, suggests again females' stronger preference for an agreement strategy in turn-bidding.

c) **Table 34** shows males' TB has a significantly higher percentage of repetition as a whole (47.3% > 36%) and particularly in type 2 repetition (19.4% > 11.8%) and type 3 repetition (21.6% > 15.8%). This can be interpreted as evidence that males may prefer to repeat their own words or their turn-openings. One function of repetition is to raise the degree of voice clarity of the turn-bidder. As males' amplitude during the overlapping time is found generally to be not as high or as clear as the females' (see the feature of unclearly-heard cases in the prosodic section in **Table 34)**, males might have to do more self-repetition to make their words clearly heard. The other function of repetition is to express more strongly the turn-bidder's attempt to gain the turn by overlapping the other's speech. The more frequent repetition by males during the overlap may indicate that males are more inclined than females to gain the floor after the overlap.

4.4.5   SUMMARY

This study has suggested that a corpus-based approach can provide a rich source for analysing the association between simultaneous speech and social factors, though the quantitative analysis of the present data is limited by the lack of more detailed description about the speech domain, the participants' relationships and social identity. The analysis has covered three aspects: the distributional frequency of TB and BC, the frequency of choosing a unit boundary to begin TB or BC, and the most salient linguistic devices and strategies used by the overlappers in TB and BC. The focus is on how these three aspects are associated with the sociolinguistic dimensions of formality, familiarity, status and gender.

(1)   Distributional Frequency

(a) There is a great difference between the occurrence of simultaneous speech in the three speech domains. This may be due to the different degree of formality in each domain. The rank order in terms of formality is that public discussion is the most formal, then telephone conversation and finally, casual conversation (see the analysis in **4.4.1 (1)**). However the rank order in terms of frequency of simultaneous speech is the reverse: casual conversation has the most simultaneous speech, then telephone conversation and finally, public discussion. The TB frequency in casual conversation is nearly twice that of public discussion (17.9/1000 w. > 9.6/1000 w.) and the BC frequency is about three times (11.5/1000 w. > 3.8/1000 w.). The TB and BC frequencies of telephone conversation are ranked between these two domains.

b) A great difference in TB frequency is also found between equal conversations and disparate conversations, the former being much higher than the latter (18.2/1000 words >10/1000 words). However, the frequency of BC in the two kinds of conversation is not such a contrast (8.6/1000 w. > 6.7/1000 w.). This may mean that while the hearers are quite aware of the degree of familiarity with

the current speakers when they make TB, they do not pay as much attention to this factor when they make BC.

(c) When we consider hearers' simultaneous speech acts in relation to their total word count in "disparate conversation", L speakers' TB and BC frequencies are both higher than those of H speakers (11.5/1000 w. > 7.2/1000 w. in TB and 10.3/1000 w. > 3.4/1000 w. in BC), though the word count of the former is much less than the latter. This indicates that when making TB, interlocutors (the L speakers in particular) may not be aware of each other's status difference as much as they are aware of the different degrees of familiarity. However they may pay more attention to the status difference when they are making BC, because the L speakers obviously make more backchannels to the H speakers than the H speakers make to the L speakers.

(d) Assuming that the frequency of hearers' simultaneous speech is related to their total word count, females have a lower frequency of TB and BC in single-gender conversations than males (11.5/1000 w. < 14.4/1000 w. in TB and 7.3/1000 w. < 10.6/1000 w. in BC). Yet in mixed-gender conversations, females have a higher frequency of TB and BC than males (13.4/1000 w. > 10.4/1000 w. in TB and 7.9/1000 w. > 3.3/1000 w. in BC). This means that in mixed-gender conversations, females seem to be more active in bidding for a turn and more involved in giving supportive backchannels by means of simultaneous speech. One of the reasons for females' higher frequency of TB may be that in these mixed-gender conversations, females do not gain enough of the floor to speak, for their total word count is about one third less than that of males.

(2) **Environmental Features Associated with TB and BC**

(a) Analysis of the occurrence of unit boundary markers immediately before TB and BC in different social contexts shows that in public discussion, the prosodic and syntactic boundary markers all have a higher percentage of occurrence than in

casual conversation and telephone conversation. This indicates that in the more formal conversational settings, the overlappers may choose more frequently a unit boundary in the current speaker's turn when they attempt to make TB and BC.

(b) With regard to the three-level unit boundary markers as a whole, there are no significant differences in relation to the other social variables such as the different degrees of familiarity between interlocutors, the status difference and gender difference. Yet with regard to a specific level of boundary marker, it is found that in disparate conversation, L speakers' BC tends to start more often at a prosodic and a syntactic boundary than H speakers' BC. This is the same in females' BC in mixed-gender conversation, which indicates that females are more likely to give backchannels to males' speech at prosodic and syntactic boundaries.

(3) <u>Linguistic Devices & Strategies Associated with TB and BC</u>

As a whole, the linguistic devices and strategies which are used in turn-bidding and backchannelling do not differ greatly in terms of the above social variables, but there are some specific devices and strategies preferred by speakers in individual domains and conversational settings, and by interlocutors with characteristic degree of familiarity or status, or difference of gender.

(a) Prosodically, stressed initial sounds have a higher frequency in the TB and BC of telephone conversations. BC also make more frequent use of stressed initial sounds in disparate conversation and they are used more by L speakers in turn-bidding. Second, a higher amplitude or degree of clarity is found in the TB and BC instances of disparate conversation. The same is found in the H speakers' and females' TB and BC. Third, incomplete TU structure at the end of the overlap is more strongly represented in the TB and BC of males and of instances in public discussion.

b) In terms of discourse items used to initiate TB or BC, agreement initials have a higher percentage in the TB of telephone conversation, the TB and BC of disparate conversation, and the BC of casual conversation and of L speakers. Second, hesitation initials and continuity initials all have a higher percentage in the TB of public discussion. Third, disagreement initials have a higher frequency in the TB of casual conversation.

c) In terms of the repetition strategies used during and after the overlappers' simultaneous speech, the three types of repetition have a higher percentage in the TB and BC of casual conversation , in the TB of disparate conversation and in males' TB and BC. Yet no significant difference is evident for the other social variables.

In summary, the four social factors (the formality of speech domain, the familiarity between interlocutors and the participants' status and gender) all put some degree of constraint on different aspects of simultaneous speech. The constraints seem to be strongly related to the frequency of occurrence of simultaneous speech, but not so strongly related to their environmental features, linguistic devices and strategies. The statistical results show that there are some significant differences in the distributional frequencies of TB and BC, but there are many fewer significant differences in terms of the percentages of unit boundary markers before TB and BC. The same is true for the percentages of overlappers' prosodic devices, discourse items and repetition strategies in producing TB and BC. Some of the results in **4.4.1 - 4.4.3** are not statistically significant because of the limited number of tokens from the data. Yet they still reveal certain general sociolinguistic features associated with different categories of simultaneous speech in a British English conversational context. However, the results may vary in other linguistic contexts where the cultural norms and values are different. This possibility is further discussed in the next chapter.

CHAPTER V

RESULTS AND DISCUSSION OF CHINESE CASE STUDY DATA ANALYSIS

In the last chapter different linguistic features and devices associated with simultaneous speech in English conversation were described. This leads us to address the fifth research question posed in Chapter I, which is, how these English features and devices correspond to their equivalents in Chinese conversation. In this way we can address one of the initial questions which motivated the present research, namely, why Chinese learners of English find it difficult to take a turn or to bid for a turn in English conversation. It could be argued that the difficulty may be due to cultural differences, such as the high value Chinese culture places on politeness in speaking etiquette which will discourage interrupting other people. Nevertheless, there are other possible reasons for the learning difficulty. One might be the learners' lack of proficiency in the English language. Chinese learners of English may simply not know how to recognise opportunities for turn-bidding in English. As noted in Chapter III, there is no corpus of prosodically marked-up spoken Chinese to make a comparison with the data from the LLC used for our analysis of English simultaneous speech. Neither is there any published literature on Chinese simultaneous speech. Thus a body of simultaneous speech collected from Chinese family conversations and used as a case study to investigate whether there are possible parallels in simultaneous speech in English and Chinese has been employed to address this research question.

The Chinese data here are all in Cantonese which is spoken in South China. Although the data is in many ways not comparable with the subcorpus of the LLC in this study, nevertheless, as long as the existence of simultaneous speech and associated features can be demonstrated, then, we can eliminate the possibility that the reason Chinese learners of English find it difficult to make simultaneous speech in English is because of their unfamiliarity with simultaneous speech in their own first language. That is, we may explore a broader area to find out the answer to the research question. The case study covered four aspects of Chinese simultaneous speech, including:

(a) the frequency distribution of the same five categories of simultaneous speech used for the analysis of the English data described on p20 of Chapter II;

(b) the environmental features occurring immediately before simultaneous speech;

(c) the linguistic devices and strategies used by overlappers in simultaneous speech;

(d) the social aspects of age and gender associated with simultaneous speech.

Owing to the limited tokens of BC in the Chinese data, the analyses in (c) and (d) were concentrated only on turn-bidding speech acts. That is, the analyses of linguistic strategies and social aspects associated with Chinese simultaneous speech were only based on the instances of STB, UTB, STC and UTC, but did not include BC.

In order to compare the Chinese case study data with the findings from the English data, the researcher selected a sample of English data from the subcorpus of LLC for the comparative analysis. The English sample consists of three texts (S1.2, S1.9 and S2.10) in the domain of casual conversation. They were held among "equals" (ie. among familiar colleagues, friends and family members). This English sample is not comparable with the Chinese family conversation in some

ways. The English informal conversations were all surreptitiously recorded while the Chinese ones were not; and the Chinese speakers in each family involved three generations (grandparents, middle-aged parents and grandchildren) while the English speakers were all middle-aged people. However, it is still possible to find evidence that both English and Chinese speakers used simultaneous speech for turn-bidding, turn-competition and backchannelling and it is even possible to identify similar kinds of devices or strategies used in a similar way and in similar places in the two languages. All this may give the researcher food for thought: why is it difficult for Chinese learners of English to take a turn or bid for a turn in a second language context?

5.1 FREQUENCY OF CHINESE SIMULTANEOUS SPEECH

Using the same formal criteria as in the LLC data to identify simultaneous speech, 288 instances of simultaneous speech were obtained from four Chinese family conversations as shown in **Table 35**.

Table 35   Frequency of Simultaneous Speech in Chinese Data

Family code	Total words	Tokens of STB	Tokens of UTB	Tokens of STC	Tokens of UTC	Tokens of BC	Total tokens	Frequency/ 1000w.
1	2913	42	40	3	3	12	100	34.3
2	1999	12	14	1	1	1	29	14.5
3	2286	31	16	6	6	2	61	26.7
4	3187	19	56	8	8	7	98	30.8
Total	10385	104	126	18	18	22	288	27.7
%		36.1	43.8	6.3	6.3	7.6		
Total TB tokens: 266		percentage: 92.4%			frequency: 25.6/1000w.			
Total BC tokens: 22		percentage: 7.6%			frequency: 2.1/1000w.			

Note:
The percentage is based on the total tokens of simultaneous speech in the Chinese sample.

Although it is not entirely valid to make a detailed comparison between the English data from the LLC and this Chinese data because of the difference in the context of use, it is nevertheless noteworthy that there is a difference in the frequency of simultaneous speech between the English and Chinese samples (see **Table 36**).

Table 36   Comparison of Frequency of Simultaneous Speech in Chinese & English Samples

(1)   Frequency of Simultaneous Speech:

	Chinese	English
Total words	10385	16098
Total speakers	23	7
Total tokens of simultaneous speech	288	589
Frequency /1000 words	27.7	36.6

(2)   Proportion of five categories:

STB	36.1%	32.6%
UTB	43.8%	15.8%
STC	6.3%	6.0%
UTC	6.3%	6.0%
BC	7.6%	39.5%
	100%	100%

Note:
The percentages in Section (2) of this table are based on the total tokens of simultaneous speech in the Chinese and English samples respectively.

Table 36 shows that simultaneous speech is quite a frequent phenomenon in both Chinese and English casual conversation, but the English sample has a higher frequency per thousand words than that of the Chinese one (36.6 > 27.7). One possible interpretation of this difference can be that the Chinese conversations were held among family members with great age and status difference, which may be a factor limiting the occurrence of simultaneous speech in Chinese conversation. In terms of the proportions of the five categories of simultaneous speech between the two samples, two striking differences in **Table 36** are noteworthy: the proportions

in the category of UTB and in the category of BC. UTB in the Chinese data has a much higher proportion than that in the English data (43.8% > 15.8%). This indicates that when the Chinese hearers attempt to bid for a turn and have an overlap with the current speaker, they are more likely to withdraw or give up their attempt at turn-bidding than the English speakers do. This may be a phenomenon related to the Chinese speech etiquette which will be further discussed in this chapter. The other striking difference is that BC in the English data has a much higher proportion than in the Chinese data (39.5% > 7.6%) which suggests that English speakers are more active users of backchannels. The result is consistent with previous research by Tao & Thomson (1991) which claimed that backchannels in Chinese conversation tend to be much fewer than in English conversation and that even fewer Chinese backchannels may occur during simultaneous speaking. The results here show that such a tendency is also strongly evident in the informal family conversations, as the frequency per thousand words of TB is obviously much higher than that of BC (25.6 > 2.1, see **Table 35**). Strangely enough, although backchannel responses may be among the last of the conversational skills acquired by children (as was pointed by Hess & Johnston: 1988), they could be amongst the first communicative strategies adopted by non-native speakers because of their salience and difference from their mother tongue. It is not unusual for native English speakers to find that non-native English learners tend to use the English backchannels such as "yes", "uhm" and "aha" more frequently in English conversation than they use other kinds of hearer response such as questioning, clarifying and challenging. One reason is that such backchannels can serve as a hearer's signal, which indicates that the hearer is listening to and understanding the speaker's words and they are therefore a pragmatically safe and polite strategy in cross-cultural communication. Yet the problem is that sometimes non-native speakers overuse "yes", "uhm" or "aha" to such an extent that it makes the native speakers doubt that the hearers have really understood what is actually being said, because sometimes their backchannels are given in a wrong place or transfer wrong

signals. As described by Hatch (1992: 22), when language learners have difficulty interpreting messages not negotiated to their level of competence, they may "fake it", pretending to understand and continue to interact in the hope that they will catch the theme or focus of the conversation by using backchannel signals. However, such faking strategy may result in a barrage of talk because communication may also break down completely since information that allows the participants to build a common theme or focus is missing. In this sense, making appropriate backchannels can be included as part of the teaching of English to non-native speakers.

5.2 ENVIRONMENTAL FEATURES BEFORE SIMULTANEOUS SPEECH

Just as a unit boundary in the current speaker's turn was found to be the most frequently used transition-relevance place for starting a simultaneous speech act in the LLC data, the Chinese data here also provide evidence that a large proportion of simultaneous speech occurs at a unit boundary which is signalled by syntactic and lexical markers. Owing to a lack of prosodic transcription comparable to the LLC, the prosodic information in the Chinese data cannot be compared with the LLC data in this study.

Table 37 Simultaneous Speech Immediately Preceded by Boundary Markers in Chinese & English Samples

Boundary type	Chinese (288 tokens)		English (589 tokens)	
	tokens	%	tokens	%
Syntactic boundary	194	67.4	388	65.9
Lexical boundary	79	27.4	132	22.4

Note:
The percentage of each boundary type is based on the total tokens of simultaneous speech in the Chinese and English samples respectively.

Table 37 shows that about two thirds of the tokens of simultaneous speech in both English and Chinese data occur at a syntactic boundary of the current speakers' speech, ie. the overlapped speech starts when the current speaker finishes a complete sentence, or a clause or a phrase or a one-word utterance such as "yes" or "no". This further confirms the fact that simultaneous speech does not occur randomly. No matter whether they are Chinese or English, hearers seek an appropriate place to bid for a turn or to give backchannels.

5.2.1   SYNTACTIC BOUNDARY MARKERS

Syntactically, 67.4% of Chinese simultaneous speech tokens occur at a point at the end of a grammatical unit. They include a complete Chinese sentence, clause, or phrase. For the examples in this chapter, a gloss is provided for every Chinese character and a free translation is given below each example in [ ]. The words in { } are the mood auxiliaries or utterance particles in Chinese which have no equivalent English translation. The words in ( ) are added to help the reader to understand the English translation, and the Chinese auxiliaries and particles are represented by "PT" in the free translation.

[1]
F: ...鍾意乜嘢 就 講 我 就 聽 我 無意見　　 G|* 女女　　　 呀 你
　　 like what just talk I just listen I no objection G|*little daughter {ya} you
　　 呃::*- - 禮拜一 ...
　　 @::*- - Monday ...

GD: ||*我 就　 講吓 *
　　 ||* I just talk*

[F: ...(you talk whatever you like I just listen I have no objection*little daughter PT
　　 you @::* --(on) Monday ...
GD: *I just talk about*]　　　　　　　　　　　　　 (F1: 0111-0113)
Note:
{ya} is a Cantonese auxiliary indicating that someone is being addressed.

[2]
GS: 你 可以 星期四　 G%* 或者 星期五 去*
　　 you may Thursday G%* or Friday go*

GD: &*呃:即 喺　　　 呃:即 喺　 * 幾多號　　 啊 依個星期 幾多號　　 啊
　　 &*@: that means @:that means* which date {ah} this week which date {ah}

[GS: you may go on Thursday *or on Friday*
　GD: *@: that means @: that means* on which date PT this week on which date]
　　　　　　　　　　　　　　　　　　　　　　 (F4: 0075-0076)
Note:
{ah} is a mood auxiliary indicating inquiry.

[3]
GF: 就喺　　　 你 要 來 G>* 三次　　 嘅樣 喳* (笑 聲---)
　　 that means you need come G>*three times like this {je}* (laughs---)

GM: >>*喺 呀*
　　 >>*yes {ya}*

[GF: that means you need to come *three times like this PT* (laughs---)
　GM: *yes PT*]　　　　　　　　　　　　　　　　 (F1: 0150-0151)
Note:
{je}and {ya} are both Cantonese particles indicating confirmation with certainty.

In [1], the grand daughter's UTB occurs when the father has just finished a
sentence "我無意見 (I have no objection)". In [2] the grand daughter's STB
occurs when the grandson has just finished an independent word 　"星期四

(Thursday)". In [3] the grandmother's BC occurs when the grandfather has just finished a phrase "你要來 (you need to come)", demonstrating again that Chinese speakers also choose a syntactic boundary place for turn-bidding and backchannelling.

5.2.2 LEXICAL BOUNDARY MARKERS

Following the criteria for lexical boundary markers in analysing the subcorpus of LLC, the lexical boundary markers here are also considered at the discoursal level. They include five sets of discourse items namely address tags, hesitation markers, termination tags, emphasizers and connectors (see **4.2.2**). As **Table 37** shows, 27.4% of the Chinese simultaneous speech tokens begin at the time when the current speakers just utter one of the five sets of discourse items. eg.:

[4] (N asks F if the recorder is still working)
N: 着緊　呀 喺咪 呀 A|*　着緊*　得得得　王太　　呢
　 still on {ya} is it {ya} A|*　still on* ok ok ok Mrs Wong {ne}

F: ‖*着緊*
　 ‖*still on*
[N: (it is) still on PT isn't it *(it is) still on * ok ok ok (how about) Mrs Wong's
　F: *(it is) still on*]　　　　　　　　　　　(F4: 0018-0019)

[5]
GF: 我 講　我個老豆 就 我 講　我個老豆 就　　總之　E|* 大家都
　　 I talk my father just I talk my father right **anyway** E|* everyone all
　　 講吓*自己 個的 趣事
　　 talk* self those interesting things

F: ‖* 傾偈　　呀嘛*
　 ‖* casual talk {ya ma}*

[GF: I talk about my father I just talk about my father anyway *everyone talk
　　　　about* your own interesting things
　F: *just casual talk PT*]　　　　　　　(F1: 0282-0283)
Note: {ya ma} is a Cantonese auxiliary indicating emphasis.

[6]
M: 我 都 唔知　　　你 你　星期五 . 開始　扎 定　C|*　喺　星期六 .*
　　I all don't know you you Friday . begin play or C|* on Saturday.*
　　我 知　人哋　去　但 喺 我...
　　I know others go but I ...

GS: ||* 奵　奵 多　人　　星期四　去　囉嘛*
　　||*very very many people Thursday go {lo ma}*

[M: I just don't know if you start to play on Friday. or *on Saturday. I know others
　　　will go but I ...
　GS: *many many people will go on Thursday PT]

(F4: 0111-0113)

Note:
{lo ma} is a Cantonese auxiliary indicating explanation or emphasis.

[7]
F: ...女女　　　　呀　你　呃:: H|*--禮拜一* 禮拜三　你　都 要
　...little daughter {ya} you @:: H|*--Monday* Wednesday you all need
　喺埋度　個啵　知道　嘛
　be in here {ge bo} know {ma}

M: ||* 女女　　　要　唔 要　聽　爺爺嘅　爺爺　　啊*
　||* little daughter want not want listen grandpa's grandpa {ah}*

[F: hi little daughter PT you @:: *--you need to be in here on Monday and
　　　Wednesday you know PT
　M: *little daughter do you want to listen to (the story of) grandpa's grandpa PT*]

(F1: 0112-0113)

Note:
{ya} is an address tag and {ge bo} indicates confirmation.

In the above four examples the simultaneous speech occurs just after the current

speaker has uttered a discourse item such as the tag question " 喺咪呀 (... is it)" in

[4]; a concluding and emphasising word "總之 (anyway)" in [5]; an alternative

connecting word "定 (or)" in [6]; a prolonged hesitating sound "呃:: (@::)" in [7].

They all function as discourse boundary markers in the current speakers' turn, thus

providing transition-relevance places for the hearers to bid for a turn.

Table 37 shows that the Chinese data has a slightly larger proportion of lexical boundary markers than the English data (27.4% > 22.4%). This may be due to the frequent occurrence of those mood auxiliaries in spoken Chinese. Mood auxiliaries may be regarded as termination tags or "utterance particles" (PT) identified by Luke (1990). Although different dialects of Chinese can have different mood auxiliaries, their position and function are more or less the same. In the examples below, the mood auxiliaries are mainly from Cantonese. The present Chinese data show nearly one fifth of the TB and BC instances occur immediately after these mood auxiliaries. eg.

[8]
M: 呢 我 哋 講 呢個 啦 T|*我 哋 講 呢個呢::　 * 唔系 食
　　this we　talk　this one {la} T|* we　talk this one this::* is not eating
　　如果我 哋...
　　if　we ...
F: ||*乜 嘢 邊個　 食　 食　 喺嘛 最好 系 食*
　　||*what　which one　eating eating　is it　best　is　eating*

[M: let's talk about this topic PT *we talk about this one this::* it is not (about)
　　　eating if we ...
　F: *what which one is it about eating eating eating is the best topic*]
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　F1: 0126-0127)

Note:
{la} is an ending auxiliary indicating emphasis.

[9]
GM:聽 吓 聽 吓 你爺爺 嘅　 爺爺啲　 歷史 呀嘛　T> *.* 你爺爺
　　listen listen　your grandpa's grandpa's story {ya ma}T> *.* your grandpa
　　好叻　　　 㗎
　　very smart {ga}

M: >>* 喺 呀*
　　>>*yes {ye}*

[GM: listen to listen to the story of your grandpa's grandpa PT *.* your grandpa
　　　(used to be) very smart PT
　　M: *yes PT*]　　　　　　　　　　　　　　　　　　　　(F1: 0105-0106)

In [8] the father's UTB occurs just after the mother utters the mood auxiliary "啦 (la)". In [9] the mother's BC occurs just after the grandmother utters the mood auxiliary "呀嘛 (ya ma)". Both are very frequently used Cantonese utterance-final particles. As a mood auxiliary has no independent semantic content and has to be attached to the end of an utterance as a bound form (see Luke, 1990:3), it can also function as a termination tag, signalling the coming end of an information chunk. Or in other words, it can indicate that the previous utterance is already a completed, meaningful unit. In this sense, the hearer can anticipate the appearance of a mood auxiliary and start a turn-bidding even before it actually occurs or before it is fully finished, as many Cantonese mood auxiliaries are prolonged. eg.

[10]
GF: <u>Friday</u> 啲人 就 去 **W%*** 之嘛*
     Friday those people just go **W%*** {ji ma}*

M: **&***星期*六 同 星期一 之嘛
   &*Satur*day and Monday {ji ma}

[F: the people will go just on Friday *PT*
 M: *(just) Satur*day and Monday PT]   (F4: 0091-0092)
Note:
{ji ma} is a Cantonese ending auxiliary indicating epistemic modality.

[11]]
GM: <u>你 唔好 開 房個機 啵 唔開得 錄音機</u>
     you not good turn on room's machine{bo}not allow turn on recorder
     W%*個啵*
     W%*ge bo*

M: **&***唔開得* 電視機 呀 細佬
   &*not allow turn on* TV {ya} little boy

[GM: don't turn on the machine in the room PT you're not allowed to turn on the
     tape-recorder *PT*
  M: *(you're) not allowed to turn on* the TV PT little boy]
                         (F3: 0053-0054)

Note: {ge bo} is a Cantonese ending particle indicating emphasis.

[12]
GM: 喺 Auckland 邊度 得 W|* 㗎::* 喺 華聯會 呀
    in Auckland which part {de W|* ga::}* in Chinese Association {ya}

M: ||*唔知 呀 而家 佢*
    ||*don't know {ya} now he*

[GM: in which part of Auckland P*T::* in the Chinese Association PT
 M: *I don't know PT now he*]   (F4: 0125-0126)
Note:
{de ga} is a Cantonese particle indicating impatience.

The above are three examples showing that the turn-bidding just precedes the
mood auxiliary or occurs before the auxiliary actually ends. In [10] the mother's
STB starts just before the grandfather utters the epistemic auxiliary "之嘛 (ji ma)";
in [11], the mother's STB starts just before the grandmother utters the emphasising
particle "個啵 (ge bo)"; and in [12] the mother's UTB starts even before the
grandmother's ending auxiliary "得㗎 (de ga)" has actually been completed. This
indicates that unit boundaries ending with mood auxiliaries are predictable. The
mechanism of prediction can be applied to other discourse items as well. The
following is an example of a predictable address tag.

[13]
M: 三點半 阿 Aunty Betty 話 三點半 啊 唔系%*咩*
    three o'clock half {ah} Aunty Betty said three o'clock half {ah}isn't %*it*

GD: &* 四* 點 半 啊
    &* four* o'clock half {ah}

[M: at half past three Aunty Betty said at half past three PT isn't *it*
GD: *half* past four TP]   (F3: 0082-0083)

In summary, simultaneous speech is not a randomly occurring phenomenon in
Chinese conversation. Both TB and BC occur in similar places in both English and

Chinese. As with the native English speakers in the LLC data, the Chinese here tend to choose a unit boundary to start turn-bidding or to give a backchannel. The boundary is signalled by a complete syntactic structure or by a lexical item which may function as an appeal to the hearer or as a termination of an utterance. **Table 37** reveals that more than two thirds of TB and BC occur at a syntactic boundary in both samples and about one fourth of TB and BC occur at a lexical boundary. Although some unit boundaries are not explicitly marked, they can be predicted by the hearers through use of contextual cues such as utterance-final auxiliaries. Further, some of the boundary markers in Chinese have the same function as those in English, but some are specific to Chinese. This suggests that the difficulty for Chinese learners of English to make TB in English conversation may be due to their unfamiliarity with the English boundary signals. On the one hand they may not be aware that some of the English boundary markers have an equivalent in Chinese, and therefore they may not recognise the proper place to start turn-bidding in English. On the other hand, even if they are aware of the Chinese boundary signals for turn-bidding such as the mood auxiliary markers, they could still find it difficult to make TB in English should they not find similar particles appearing in English. This has implications for teachers of English in China.

5.3 LINGUISTIC STRATEGIES USED BY OVERLAPPERS

It may be expected that any language can have effective strategies to carry out different types of speech acts. However it could not necessarily be assumed that Chinese would use similar linguistic strategies to those in the LLC data in order to

make successful turn-bidding. Because the Chinese data was not prosodically analysed, the analysis here concentrates on the five sets of discourse items which are associated with the beginning of simultaneous speech and the three types of repetition used during and immediately after simultaneous speech. A comparison was made between the Chinese and English samples to see if the strategies identified in the LLC data also occur in the Chinese turn-bidding instances (turn-bidding here includes STB, UTB, STC and UTC).

Table 38   <u>**Occurrence of Discourse Items and Repetition**</u>

<u>**in Chinese and English Samples**</u>

Strategy	Chinese(288 tokens)		English (589 tokens)	
	tokens	%	tokens	%
Discourse items	52	18.1	188	31.9
Repetition	66	22.9	157	26.7

5.3.1 DISCOURSE ITEMS

Table 38 shows that the English sample has nearly one third (31.9%) of the TB tokens initiated by one of the five sets of discourse items. Although the Chinese sample does not have such a high proportion as the English sample (18.1% < 31.9%), similar kinds of discourse items are also used in nearly one fifth of the instances of Chinese simultaneous speech. According to the criteria used to identify the five sets of discourse items in the subcorpus of LLC, it can be shown that the Chinese discourse items which are used to introduce the overlappers'

speech also include the initials of agreement, disagreement, hesitation, exclamation

and continuity. eg.:

[14]
GF:　最緊要　　　就喺 你哋 有興趣　個的 題目 %*比較好 啵*
　　　most important just is you　interested those topics %*better　{bo}*

M: &*喺呀　叫 老豆 * 講 啊 睇吓　　就慘囉　　十次　　呢
　　&*yes{ya} ask father* talk {ah} have a look just terrible{lo} ten times {ne}
　　講　十次...
　　talk　ten times ...

[GF:the most important thing is the topic you're interested in *(that would be)
　　　better PT*
 M: *yes PT ask father* to talk PT just a look at it it would be terrible PT ten
　　　times PT talk ten times ...]　　　　　　　　　(F1: 0091-0092)
Note: {bo} is a Cantonese particle indicating persuasion.

[15]
GM: ..乜嘢　都 講 . %* 又 講 佢病　呢:: *
　　...anything　all　talk . %*also talk she's ill {ne::}*

GF: &*耶: 着 -- 等　等 *等 我 着番 件衫　　　　　先...
　　&* ee: wear -- wait　wait* wait　me put on　an item of clothing first ...

[GM: ...(she) talked about everything *also talked about her illness PT*
　GF: *@: (you should) wear --(you should say) wait me wait me* until I get
　　　dressed]　　　　　　　　　　　　　　　　　　(F1: 0190-0191)

[16]
M: ...重要　　唔 %*喺 toll call　黎喎*
　　...moreover not %* is toll call　{lai wo}*

GM: &* 嘩 :: 你 真　真喺*　個晚　　潘潘瑞琴　　　打電話
　　&*{wa}:: you really really are*that night Pan Pan Rui Qin　made phone call
　　俾我...
　　to me...

[M: ...moreover (it's)* not a toll call PT*
GM:*wow:: you really really (don't know)* that night Pan Pan Rui Qin gave me a
　　call...]　　　　　　　　　　　　　　　　　　(F1: 0172-0173)
Note:
{wa} is a Cantonese exclamation word, {lai wo} is a sentence-final particle of
emphasis.

[17]
M: ... 年年 　 都 去 打 個個 master |* (<u>不清</u>　　　　) * <u>噉咪 我唸</u>
　　...every year all go play that master|*(inaudible words)*so that I think
　　年年　　　都...
　　every year all ...

GM: ||*<u>但 喺</u> 你　打::　　喺邊度*
　　||* **but** you play:: in which part*

[M: ...every year they all go to play that master (cup) *(inaudible words) so I think
　　　every year they all...
GM: *but where do you play*　　　　　　　　　　　(F4: 0128-0130)

The four examples above illustrate that the hearers use discourse items to initiate
their turn-biddings. In [14], the mother's STB is introduced by an agreement "
喺呀 (yes)"; in [15], the grandfather's STB is introduced by a prolonged hesitator
"耶:: (@::)"; in [16], the grandmother's STB is initiated by an exclamatory "嘩
(wow)" and in [17], the grandmother's UTB is initiated by a conjunction "但 喺
(but)". Out of the 266 TB in this data, 18.1% are introduced by such discourse
items. They function in a similar way to their equivalents in the LLC data. For
example, the turn-bidders use agreements as a trigger to gain the floor; or use a
hesitation marker as a hedge to soften the abruptness of insertion; or use an
exclamation to draw the current speaker's attention and use a connector to make
the turn transition cohesive and coherent. The present data also reveals that
among these discourse items, connectors have a much more frequent occurrence
than the others. They take up about 10% of the 266 TB cases. This may be due to
the frequent occurrence of one Cantonese initial particle "噉" whose English
equivalent can be "so" or "then" or "well" or "but", all playing a connecting
function in the discourse. eg.:

[18]
GF: 咳:: 唔知 %*乜嘢 有趣*
　　{hai::} don't know %*what's interesting*

M: &*噢 就 唔喺 我 㗎嘑*我 哋家庭歷史 無 乜嘢 有趣 㗎吓
　　&* so just is not mine{ga la}*our family story no anything interesting{ga ha}

[GF: PT (signing) don't know what is interesting
 M: so that's not my case PT my family story has nothing interesting PT]
　　　　　　　　　　　　　　　　　　　　　　　　(F1: 0235-0237)
Note:
{ga la} and {ga ha) are all mood auxiliaries indicating an emphatic ending.

[19]
M:....你 自己 揾 佢 |*佢 最初 佢 最初 *問 我 㗎啦 最初 填 form
　　...you yourself found him |*he at first he at first*asked me{ga la}at first fill form

GF: ||* 噢 你 哋幫 佢 揾吓*
　　||*then you help him look for*

[M: you found him yourself *at first he at first he* asked me PT to fill the form at
　　first
 GF: *then you help him to find someone*]
　　　　　　　　　　　　　　　　　　　　　　　(F4: 0119-0120)

[20]
F: 即喺 乜嘢 呀即 |*喺 最 最 有趣 * 嘅家庭歷史 呀...
　　that means what{ya}that|*means most most interesting*{ge}family story {ya}...

GF: ||*噢 就 試 吓 嘑*
　　||*well just try once {na}*

[F: that means that *means the most interesting* family story PT
GF: *well (in that case) let's have a try*　　　(F1: 0234-0235)

In [18], the mother starts her TB with "噢" which means "so", as a consequential
response to the grandfather's statement. In [19] the grandfather's starting word
"噢" is a continuant, which is equivalent to the English "then". In [20] the
grandfather's TB is introduced by another meaning of "噢", which could serve the

function of a turn initiator like "well". Such a connecting particle is quite popularly used in spoken Cantonese.

5.3.2 REPETITION STRATEGIES

Table 38 shows that both Chinese and English samples have about a quarter of turn-bidding instances with repetition patterns (22.9% in Chinese and 26.7% in English). According to the position of repetition in the overlap, there are also three types of repetition identified in the Chinese data. The first type occurs at the beginning of the overlap, ie. the turn-bidder repeats the current speaker's last word or pattern and then continues the new turn (see example [21]); the second type of repetition occurs within the overlap, ie. the turn-bidder repeats his or her own words or pattern during the overlap (see example [22]); the third type of repetition occurs immediately after the overlap, ie. the turn-bidder repeats his or her own words or pattern after gaining the floor (see example [23]):

[21]
F: ...<u>第二次</u>　　都 喺 自由 %*自由*
　　...the second time also is free %* free*

M: &*自由*　不過 唔 喺 好　硬性　　硬性　規定　　　嘅 ...
　　&* free*　but is not very strictly strictly prescribed {ge} ...

F: ...the second time is also free *free*
M: *(it's) free* but still it is not very strictly strictly prescribed PT]

(F1: 0121-0122)

[22]
M: 你 唔去 %*着 件衫 出來*
   you not go %*wear an item of clothing come out*

F: &*你 叫 佢. 你 叫 佢 唔好講住 *等我 等我 去
   &*you asked her. you asked her not speaking* wait me wait me go
      着番 件衫 先...
      put on an item of clothing first...

[M: (why) don't you go to *get dressed (and then) come out*
 F: *you (should have) asked her you (should have) asked her stop talking* wait
     wait until I get dressed]   (F1: 0180-0181)

[23] (after laughter)
GS: <*嗽 人哋 問 我 呀嘛*
     <*but he asked me {ya ma}*

M:<<* 嗽佢問 你 需唔需要 *嗽佢 問 你 需唔需要
  <<* so he asked you need not need *so he asked you need not need
     transportation
     transportation

[GS: *but he asked me (first) PT
 M: *so he asked if you need or not*so he asked if you need transportation or
     not]   (F4: 0097-0099)
Note:
{ya ma} is a Cantonese auxiliary indicating emphasis and signalling a sentence ending.

In [21], the mother's STB, which starts by repeating the father's last words "自由 (free)" before the overlap, could serve as an agreement to or a continuant of the father's speech. In [22], the father repeats his own pattern "你叫佢 (you asked her)" during the overlap to reiterate his statement. [23] is a turn-competition case between the mother and the son. As the mother gains the floor after the overlap, she restarts her new turn by repeating her pattern "嗽佢問你需唔需要... (so he asked if you need ...)", which can be regarded as a recycled turn-beginning. All these examples indicate that when two speakers' voices are overlapped, the clearness of voice quality could be affected or their expression could be hampered.

Therefore the turn-bidders need to reinforce their speech by repetition so as to express their continuing wish for a turn.

In summary, Chinese simultaneous speech tends to have similar linguistic strategies in the use of discourse items and repetitions as those in the English LLC data. The discourse items and repetitions of both languages appear in similar positions of the overlap and display the same discourse functions. Yet native speakers may not be consciously aware of such strategies, though they may use them habitually in their mother tongue. On the other hand, when learning a foreign language, learners may not pay enough attention to such strategies in the target language, and hence there is a lack of positive transfer from L1 to L2. As noted above linguistic analysis has been carried out at discourse and pragmatic levels. Any further analysis at the prosodic level is hampered by the lack of fully transcribed Chinese data especially in an electronic corpus. Were it otherwise, the data could be investigated further to see if the difference between the sentence-stress timed English rhythm and the syllable-timed Chinese rhythm might have a bearing on the difficulty of Chinese learners of English when involved in turn-taking and turn-bidding. The rhythmic unit in English (ie. a beat or a foot) is composed of one stressed syllable with or without unstressed syllables around it (see O'Connor:1967) while a Chinese rhythmic unit is a combination of one or more stressed words with or without unstressed words as company (see Liu Guanghui & Shi Peiwen: 1988). It may be difficult for Chinese learners of English to get used to the rhythmic pace in English conversation, a factor which may increase the difficulty they have in overlapping the ongoing turns of native English speakers, and thus taking the floor.

5.4 SOCIOLINGUISTIC ASPECTS OF CHINESE SIMULTANEOUS SPEECH

The limited number of 22 BC tokens identified in the Chinese data is insufficient to warrant further analysis of BC in respect of its conventional social variables such as age, gender and status. However, an exploration of the 266 turn-bidding tokens (including STB, UTB, STC and UTC) reveals interesting data about the social constraints on Chinese speakers' turn-bidding behaviour, especially in terms of gender and age.

5.4.1 TURN-BIDDING AND GENDER

The Chinese data, which involve family members of three generations are all mixed-gender conversations. The present study analyses both the percentage of TB produced by speakers who are within the same gender (ie. female-to-female or male-to-male) and those who are different in gender (ie. female-to-male or male-to-female). The former is defined as "same gender TB" and the latter as "cross-gender TB". The results are shown in **Table 39**.

**Table 39   Distribution of Turn-bidding by Males and
Females in the Chinese Sample**

Gender type	Tokens	%	Frequency/1000 w.
Same-gender	123	46.2	11.8
Female-to-male	70	26.4	6.7
Male-to-female	73	27.4	7.0
Total	266	100	25.6

Table 39 shows that nearly half (46.2%) of the Chinese TB tokens occurred between speakers within the same gender, while the rest of the TB tokens are quite evenly distributed among speakers different in gender (26.4% are made by female

to male and 27.4% are made by male to female). As the Chinese sample has 12 female speakers and 11 male speakers while the English sample extracted from the LLC for comparison has only 2 female speakers but 7 male speakers, it is invalid to compare the same-gender TB and cross-gender TB in the two samples. Yet the above results from the Chinese data may be interpreted to show that simultaneous speech in the Chinese data not only has a lower frequency than that of the English data (see 27.7/1000 w. < 36.6/1000 w. in **Table 36**), but also occurs mainly among speakers within the same gender. If we compare the females' and males' cross-gender TB frequency per thousand words in Chinese data with the corresponding ones in the LLC subcorpus (see **Table 32** (3) in **4.4.4**), the former is much lower than the latter (for female: 6.7 < 13.4, for male: 7.0 < 10.4).

To have a closer examination of the gender effect on turn-bidding in both samples, the present study sorted out all the cross-gender TB tokens from the two samples for comparison. The focus was on the frequency of turn-bidding per thousand words and the proportion of successful cases (ie. STB and STC) and unsuccessful cases (ie. UTB and UTC).

Table 40   Cross-gender TB in Chinese & English Samples

	Chinese		English	
	female	male	female	male
(1) Frequency distribution:				
Total words	7553	2832	6282	4175
TB tokens	70	73	138	88
frequency /1000 w.	9.3	25.8	22	21.1
(2) Proportion of successful & unsuccessful turn-bidding:				
Successful	45.7%	46.6%	63.8%	68.2%
Unsuccessful	54.3%	53.4%	36.2%	31.6%
	100%	100%	100%	100%

Table 40 shows that in terms of frequency of simultaneous speech in our data, the Chinese females have a much lower frequency than Chinese males (9.3/1000 w. < 25.8/1000 w.) while English females have a marginally higher frequency than English males (22/1000 w. > 21.1/000 w.). As these frequencies are obtained by dividing the total tokens of females' TB by the total word count of females themselves (the same is done in case of males'), one possible interpretation of the results is that Chinese females have obtained the floor more than males because their total word count is much larger than male's (7553 > 2832). Therefore they need not make as much turn-bidding or turn-competing as males do. Another possible explanation may be that Chinese females are not likely to overlap male's speech until it is really necessary. This again may be related to Chinese speech etiquette.

In comparing the proportion of successful and unsuccessful turn-bidding, the second section of **Table 40** shows that Chinese females and males have a very similar percentage (45.7% : 46.6% in successful TB and 54.3% : 53.4% in unsuccessful TB). Yet in the English data, male's turn-bidding has a slightly larger proportion of successful cases than female's (68.2% > 63.8%) but a smaller proportion of unsuccessful cases than female (31.8% < 36.2%). This may suggest that in casual conversation although Chinese females make much less turn-bidding when males are speaking, when a turn-bidding situation really occurs, they seem to have no fewer opportunities to win the floor than males. Both males and females have similar chances to win or lose the floor after the overlap. However, in the English sample, although both males and females have a larger proportion of successful TB than unsuccessful TB, males tend to win the floor more than females.

5.4.2   TURN-BIDDING AND AGE

In looking for an association between turn-bidding and age, the Chinese data cannot be compared with the data from the LLC. However the Chinese data reveals that age difference may affect the occurrence of turn-bidding and the outcome of overlapped speech in cross-generation conversation.

Table 41   Distribution of Chinese TB in Terms of Age

Same-generation TB and cross-generation TB						
Age type	Successful		Unsuccessful		Total	
	tokens	%	tokens	%	tokens	%
same generation	42	34.4	47	32.6	89	33.4
old-to-young	47	38.5	29	20.1	76	28.5
young-to-old	33	27.1	68	47.2	101	38.1
Total	122	100	144	100	266	100
Cross-generation TB						
	tokens	%	tokens	%	tokens	%
old-to-young	47	61.8	29	38.2	76	100
young-to-old	33	32.7	68	67.3	101	100

Note:
The percentage is based on the total tokens of each type of TB.

Table 41 shows that 33.4% of TB occurs between speakers who belong to the same generation such as grandfather to grandmother, father to mother, or grand daughter to grandson. Only 28.5% of TB is made by the older generation to the younger generation such as grandparents to their children or to their grandchildren, and middle-aged parents to their children. Yet 38.1% of TB is made by the younger generation to the older, ie. grandchildren to their parents or to their grandparents, and middle-aged parents to their elderly parents. However, this does not mean that the younger speakers are always successful turn-bidders. A further examination of TB in terms of successful and non-successful outcomes shows that there is a more subtle difference between the older and the younger. With regard to the percentage of cross-generation TB, out of the 76 TB tokens made by the older

to the younger, 61.8% are successful cases and only 38.2% are unsuccessful cases. However, out of the 101 TB instances made by the younger to the older, only 32.6% are successful but 67.4% are unsuccessful. It could be interpreted that in a turn-bidding or turn-competing situation, the older people are more likely to gain the floor (ie. continue to speak after the overlap) than the younger. It could be a result of either the younger having yielded the floor to the older, or the older having been more assertive in striving for a turn. One interpretation of this may lie in the characteristics of Chinese speaking etiquette which emphasise older people's authority and the requirement for youth to respect age. Its impact in conversation may be firstly, that the younger tend to avoid interrupting the older; and secondly when an interruption does occur, the younger tend to give the floor to the older for the sake of politeness and respect.

We may describe the picture in this way: on the one hand, simultaneous speech seems unavoidable in conversation because the next speakers may mistime the start of their speech or they may sometimes feel it necessary to bid for a turn. On the other hand, when such turn-bidding or turn-competing actually happens, who will yield the floor and who will dominate the other speaker and continue to speak can depend on various factors. Some are related to the speaker's turn-bidding strategies or turn-holding strategies; some are related to the speakers' own value of turn-bidding or competing in that specific situation, that is, whether it is necessary, worthwhile or polite to go on speaking, or whether it is wiser to give up the floor. The findings from the Chinese conversations indicate that in such a situation, the younger people are more likely to give the floor to the older people. This may be another explanation of why Chinese learners of English may feel embarrassed in turn-bidding when they are communicating with native English speakers, especially when the interlocutor is older or of higher social status. The reasons for this are discussed further below.

Culturally speaking, Chinese and English speech etiquette, although sharing the same general principles of cooperation and politeness, may have different emphases. China has a history of more than five thousand years, during which agriculture was the major production activity, an activity which is oriented to the expectation of the elderly and experienced people. The hierarchical clan and family network was the major social administration system, within which everybody had to know his or her place in the social order and be aware of what to say and how to behave. Chinese culture has thus inherited a tradition of respect for seniority and authority, based on moral doctrine advocated by Confucianism, which emphasises a hierarchical order between the old and the young. This is evident in Chinese speech etiquette, which takes special account of differences in age and status, valuing modesty, inwardness and even quietness --- as expressed in the Chinese proverb: "kwai yahn yu chí (a noble person would speak less and keep his ideas till later)". An essential aim of social communication is to establish, to maintain or to coordinate the participants' inter-relationship so as to keep the hierarchical system working. A traditional norm, based on status or age difference, was that children should not interrupt adults when speaking, and this was extended broadly to similar situations such as student not interrupting teacher, employee not interrupting employer, and even female not interrupting male, particularly in public. Therefore in a Chinese conversation, a quiet child who sat to the side listening to adults' talk, not speaking until being asked, was regarded as well behaved. The same was the case for a student who kept quiet while listening attentively to the teacher; or for a wife who just kept smiling and listening while her husband talked with visitors. Nowadays, the traditional hierarchical system has long disappeared and people's values have changed a lot. Equity and equal opportunity are greatly enhanced, especially in terms of the relationship between old and young, men and women, "high ranking" authorities and common citizens, and the like. However, the awareness of age difference and status difference is still guiding the direction of etiquette in many speaking situations. In the following Chinese data, which was

extract from an inter-generational conversation amongst members of a Chinese family living in New Zealand, we can still find the cases of traditional Chinese expectations:

[24]
(In a three-generation conversation recording situation, all family members (including the young children) were encouraged to speak while being recorded. Yet when adults were talking, the 9-year-old grandson was singing and speaking to himself for a while, and the mother and grandmother tried several times to stop him:)

M: 你 做 乜嘢 呀 人哋 聽 你 唱歌 咩:
　　you doing what {ya} others　hear you singing {me:}
[M: what are you doing (here) PT you think others would hear you singing PT]

(a few seconds later, mother said again)
M: 正經啲　　呀嘛
　　formal a bit {ya ma}
[M: can you behave yourself PT]

(but the boy was still speaking to himself, grandma said)
GM 唔好　　　嘈
　　no good　make noise
[GM: don't make any noise]

(a minute later, grandma said again)
GM:停下 講嘢　　　　先

　　stop　talking　things first

This example may show that traditional Chinese speaking etiquette is still deeply rooted in the older family members' evaluation of the younger ones' behaviour, despite the fact that in this case, the family has been settled in Wellington, New Zealand for a dozen years and the young children were born and grew up in this Western society.

5.5 SUMMARY

Although the available Chinese case study data is not able to be directly compared with the quantitative analysis of the English conversation in the LLC in terms of all the multi-level linguistic and sociolinguistic features, there is still enough evidence of certain similarities and differences between the Chinese and English simultaneous speech. The results of the case study show:

(a) The Chinese speakers also make frequent simultaneous speech in casual conversation, though the frequency is not as high as for the English speakers in the LLC because the Chinese speakers make many fewer backchannels than the English speakers.

(b) Chinese speakers also choose similar linguistic unit boundary signals to make turn-bidding and backchannelling and use similar discourse and pragmatical strategies to make successful turn-bidding. Many discourse items have similar functions and positions to their equivalents in English, although some discourse items have no equivalents in English.

(c) Chinese turn-bidding speech acts occur most frequently between speakers within the same gender and age range. The Chinese females, though they make less frequent turn-bidding on males' speech, do not have fewer chances to gain the floor than males if they actually attempt to bid for a turn. The younger Chinese, though they attempt more frequently to bid for a turn than the older Chinese, are more likely to yield the floor when the situation of simultaneous speech actually occurs.

These findings may help to answer the question why Chinese learners of English may face difficulties in turn-bidding and turn-taking in English conversation. The reasons are multi-dimensional and complex. Accounting for the difference according to social factors such as age and gender, we may say different cultural norms and speech etiquette is one of the important reasons, because every culture

has its own ways to express politeness and cooperation in conversation. However, accounting for the fact that simultaneous speech occurs frequently in both languages and that native Chinese and English speakers share many similar skills and strategies in carrying out such speech acts, the learners' English proficiency and experience in English conversation is another major reason for the difficulty of making simultaneous speech in English. On the one hand, Chinese learners of English may not be aware of the features and devices of their own language; and on the other hand, they may not be familiar with the corresponding ones in English. They will find it easy to bid for a turn in their own language, because they are so familiar with Chinese boundary signals and contextual cues that they may do so unconsciously. But if they have not had practical experience in English conversation, such Chinese knowledge cannot be automatically transferred into their spoken English performance. It is at this point that English teachers may need to assist their students by offering them more exposure to those features or devices in the target language and by making comparisons between the two languages as well.

The lack of English communicative competence on the part of Chinese or other non-native speakers can be seen in their withdrawal from conversation because they just do not know what are appropriate topics for discussion with foreign speakers nor how to start a conversation. This can be supported by the common feeling that young children may interrupt adults at an inappropriate time or in an improper manner, thus failing to gain the floor in conversation. This is due to the young children's lower language proficiency in their mother tongue and their lack of language performance experience. The same could be said of a non-native English speaker. The difference is that while young children can naturally acquire such a competence as they grow up, adult learners have to be taught this in effective ways and in a shorter time by more conscious learning and practice.

CHAPTER VI

CONCLUSIONS AND IMPLICATIONS

In this chapter, the major findings of the research are summarised in relation to the five research questions set out in Chapter I. Firstly, there is a summary of the findings in the analysis of the English spoken corpus in relation to the first four research questions; and secondly, a summary of the findings of the case study of Chinese conversation in relation to the fifth research question. Finally, implications from this study are discussed.

6.1 SUMMARY OF FINDINGS IN THE ANALYSIS OF THE SPOKEN CORPUS OF ENGLISH

6.1.1 HOW OFTEN SIMULTANEOUS SPEECH OCCURS

The present study retrieved 2011 instances of simultaneous speech from a 91802-word corpus of English conversation which formed part of the London-Lund Corpus. The general frequency over three speech domains is almost 22 occurrences of simultaneous speech in every thousand words. This means that simultaneous speech occurs quite frequently in English conversation. Five categories of simultaneous speech were identified in terms of their pragmatic functions. They include turn-bidding speech acts, ie. TB (which is subcategorised as STB, UTB, STC and UTC) and non-turn-bidding speech acts, ie. BC (backchannels). The turn-bidding cases take up 65% of the total overlaps and thus

form the bulk of simultaneous speech. Moreover, about 70% of TB were identified as successful turn-bidding (STB) and successful turn-competing (STC). STB occurred most frequently with nine occurrences per thousand words. As the corpus is a collection of naturally occurring English conversations which cover casual conversations, radio discussions and telephone conversations, and as it involved 193 speakers (113 females and 80 males), a conclusion can be drawn that turn-bidding, especially successful turn-bidding, is a frequent and normal phenomenon in English conversation, even though it used to be thought to be abnormal, interruptive or even offensive. It is normal in the sense that it reflects the natural operation of the human speech mechanism --- the mechanism of anticipation and prediction in particular. It is normal also in the sense that its occurrence is not random, but rule governed and context constrained. Gumperz (1982: 101) once described conversation as a dynamic process full of anticipation and inference, which both interprets what has been said and generates expectations about what is to come. He noted that:

> The process is always situated or context bound. It begins with informed guessing based on what we know about the physical setting, the participants and their backgrounds, and how we relate the situation at hand to other known activities. These initial hypotheses are subjected to constant modification by our perception of information signalled in both the form and the content of speech. ... Knowledge of the conversational activity entails expectation about possible goals or outcomes for the interaction, about what information is salient and how it is likely to be signalled, about relevant aspects of interpersonal relations, and about what will count as normal behaviour.

The anticipating and inferring mechanism, the context constraints and the linguistic signalling system mentioned by Gumperz above were all found in the present study.

6.1.2   WHEN AND WHERE SIMULTANEOUS SPEECH USUALLY OCCURS

A three-level linguistic analysis was made of the environment in which simultaneous speech is most likely to occur. The findings show that hearers who initiate simultaneous speech usually make use of the unit boundaries in the current speakers' turn, or even anticipate a possible completion point of a turn unit boundary, because a unit boundary can be regarded as a transition-relevance place to start a new turn. Such transition-relevance places in the current speaker's turn are usually marked by prosodic, lexical and grammatical features.

(1)   Prosodic Features

61% of the simultaneous speech instances occur at or near a tone unit boundary. Such boundaries are signalled either by a pause or a nuclear pattern, especially when the nuclear pattern is a falling tone. There is a higher frequency of prosodic boundary markers occurring before STB than UTB.

(2)   Lexical Features

Nearly 24% of the simultaneous speech instances occur when the current speakers utter some specific words or sounds which are called as discourse items. They include: address tags (such as tag questions, name addressing and phrases like "you know" or "you see"), hesitation markers (such as "well", "sort of", "@m:"), termination tags (such as "...that sort of thing", "... and so on and so forth"), emphasizers (such as "yes", "right", "absolutely", "anyway") and connectors (such as "and", "so", "or", "but", "because"). The first two sets can function as an appeal to a hearer's response and the other three sets can function to terminate an utterance. They can all be regarded as transition-relevance places for the hearer to

start a new turn. Turn-bidding has a higher frequency of this lexical feature than backchannelling does.

(3) Grammatical Features

77.4% of the simultaneous speech instances occur at a syntactic boundary, ie. when the current speaker has just completed a sentence, or a clause or a phrase structure or a one-word utterance such as "yes". STB, UTB and BC all have a similar proportion of this feature. Although 21.5% of the simultaneous speech occurs at a non-syntactic boundary such as a word boundary or a syllable boundary, hearers can also anticipate the coming grammatical boundary by using their knowledge of grammatical and lexical structure. This is particularly so in the cases of TB.

Finally the three levels of unit boundary signals --- prosodic, lexical and syntactic --- usually occur together in the current speaker's turn. 97.5% of the tone unit boundaries are coincident with syntactic boundaries and 91% of the lexical boundaries (which refer to a number of discourse items) are coincident with syntactic boundaries as well. Moreover, 62.2% of the syntactic boundaries and 65.1% of the discourse items are accompanied by salient prosodic features such as word stress or a nuclear pattern. As a result, the co-occurrence of signals at three levels emphasises a unit boundary and makes it easier for the hearer to recognise it as a transition-relevance place for turn-bidding or backchannelling.

6.1.3 LINGUISTIC DEVICES AND STRATEGIES WHICH ARE USED TO INITIATE SIMULTANEOUS SPEECH

Linguistic analysis was also undertaken of the strategies and devices being used in simultaneous speech. It was found that particular prosodic devices, discourse items and repetition strategies are frequently used by the person who is the overlapper in simultaneous speech.

(1) <u>**Prosodic Means**</u>

(a) The overlapper tends to use prominent sound features to attract the current speaker's attention, because more than two thirds of simultaneous speech is introduced either by a stressed initial or a nuclear patterned initial or a single-word TU initial. This is particularly the case in BC. Comparing STB and UTB, the former has a higher frequency of this feature than the latter.

(b) Generally, successful turn-bidders and turn-competitors have a faster speech rate than the other speakers during the overlap, because the former have a larger average word count in the overlap than the latter. Yet this is not the case for unsuccessful turn-bidders and turn-competitors, nor is it so for backchannel makers.

(c) As for the amplitude of the overlappers, successful turn-bidders and turn-competitors generally have a higher degree of clarity during the overlap than other speakers, because STB and STC all have a lower occurrence of unclearly-heard words or syllables during the overlap. However this is not the case with unsuccessful turn-bidding and turn competing, for UTB and UTC have a much higher frequency of unclearly-heard words or syllables in their overlap. Moreover, the clarity has a tendency to decrease at the end of UTB and UTC, but it has a tendency to increase at the end of STB and STC. Very little difference was observed between the overlapper's and the current speaker's amplitude in backchannels.

(d) At the end of the overlap, STB and STC have a higher percentage of incomplete tone unit structures than UTB, UTC and BC, which indicates that the successful turn-bidders and turn-competitors tend to have a stronger intention to continue their speech after the overlap.

(2) Discoursal Means

54.5% of simultaneous speech instances in this corpus data are introduced by five
sets of discourse items in descending order of frequency as follows:

(a) agreement initials such as "yes", "yeah", "yup", "@m", "m" (in falling tone),

(b) hesitation initials such as "well", "@", "@m", "mhm" (unstressed or in level
  tone),

(c) exclamation initials such as "oh", "ooh", "ah",

(d) continuity initials such as "and", "but",

(e) disagreement initials such as "no".

Comparing the five categories of simultaneous speech, BC has the highest
percentage of agreement initials; STB has the highest percentage of hesitation
initials; STC has the highest percentage of exclamation and continuity initials; UTB
and UTC have the highest percentage of disagreement initials. Moreover, more
than two thirds of all these discourse items are marked with prominent prosodic
features such as a stress or nuclear tone, but the disagreement initials tend to have a
softened intonation contour such as a rising/level nuclear tone or even unstressed
words.

(3) Pragmatic Means

17.3% of simultaneous speech instances have at least one of the three types of
repetition. In the type 1 repetition, the overlapper repeats the current speaker's
incomplete sentence pattern or words so as to take over the current speaker's
unfinished turn. This only appears in STB and UTB. In the type 2 repetition, the
overlapper repeats his or her own words or pattern during the overlap. This
appears in all five categories of simultaneous speech. In the type 3 repetition, the
overlapper repeats his or her own words or pattern immediately after the overlap so
as to enhance the new turn opening. This only appears in STB and STC.

STB has the highest frequency of the three types of repetition, which makes up about one third of its total 830 tokens. Moreover, the successful turn-bidders and turn-competitors tend to retain or upgrade intonation contours more often than unsuccessful turn-bidders and turn-competitors. In contrast, the UTB and UTC repetitions have a higher percentage of less prominent intonation contours.

Comparing the linguistic devices and strategies used in successful turn-bidding instances (ie. STB and STC) and unsuccessful turn-bidding instances (ie. UTB and UTB), STB and STC generally have a faster speech rate, a higher amplitude, and a higher percentage of prominent onset of the simultaneous speech and incomplete TU endings than UTB and UTC. STC and STC also employ more discourse items to initiate the turn-bidding, as well as more repetition. As for the BC instances, because of their briefness in form and supportiveness in function, the devices are characterised by frequent use of prominent starting sounds, agreement initials and exclamation initials. All these can be regarded as hearers' effective strategies in active and cooperative communication.

6.1.4 HOW SIMULTANEOUS SPEECH IS CONSTRAINED BY SOCIAL FACTORS

The present research analysed the association between simultaneous speech and its social constraints in terms of :

(a) the formality of speech domains, ie. casual conversation vs. public discussion vs. telephone conversation;

(b) the degree of familiarity between interlocutors, ie. equal conversation vs. disparate conversation;

(c) the status difference between interlocutors, ie. speakers with higher status vs. speakers with lower status,

(d) the gender difference between speakers and hearers.

The findings are summarised as follows:

(1) Formality

Firstly, the degree of formality in different speech domains has a strong association with the frequency of TB and BC occurrence, ie. the more formal the situation, the lower the occurrence of simultaneous speech. For example, casual conversation has much higher frequencies of both TB and BC than formal public discussion. Secondly, formality can alter the frequency of hearer's choice of unit boundary at which simultaneous speech is made. The present research finds that the more formal the situation, the more frequent the occurrence of boundary markers before simultaneous speech. For example, hearers in public discussion are shown to make TB and BC more frequently at an actual unit boundary which is prosodically or grammatically marked than hearers in casual conversation. Thirdly, formality in conversational settings can make the hearer prefer some TB or BC devices and strategies more than others. For example, the TB tokens in public discussion have a tendency to have fewer stressed initial sounds and fewer initials of agreement. The BC tokens in casual conversation have a higher frequency of agreement initials which are accompanied by prominent prosodic stress as well. And in the domain of telephone conversation, TB has the highest percentage of prominent initial sounds and agreement initials, but the lowest percentage of repetition.

(2) Familiarity or Intimacy

The degree of familiarity or intimacy between interlocutors has an obvious association with the frequency of TB although it is not so obvious in cases of BC. In general the more familiar the interlocutors, the more frequent the occurrence of simultaneous speech. For example, equal conversation has a much higher frequency of TB and a slightly higher frequency of BC than disparate conversation. However, there is no significant difference between familiar and unfamiliar interlocutors in accounting for the choice of unit boundary to start TB

and BC. As for the preferred devices, TB among "disparates" tends to have a higher amplitude and to be introduced more by discourse items signalling an agreement and a continuity. However the TB tokens in equal conversation are more frequently initiated by exclamation markers. The BC tokens in disparate conversation are more frequently initiated by words or sounds of agreement and hesitation, while in equal conversation they are more frequently introduced by exclamation markers.

(3)   <u>Social Status</u>

Difference in interlocutors' social status (which is assumed to be related to professional ranking and age in this study) shows an interesting relationship with the occurrence of TB and BC. The subcorpus shows that in disparate conversation higher status speakers (H speakers) generally have a larger word count than lower status speakers (L speakers). As a result, L speakers' speech has more occurrence of TB and BC than H speakers in terms of per-thousand-word frequency. With regard to the environmental features, no significant difference is found between H and L speakers in terms of using unit boundary markers for TB, but L speakers tend more frequently to choose prosodic and syntactic boundaries when they insert BC into H speakers' turns. Moreover, the turn-bidding produced by L speakers tends to have a lower amplitude, but more initials of disagreement. Nevertheless, the backchannels produced by L speaker are introduced by more initials of agreement, which are prosodically stressed.

(4)   <u>Gender</u>

Gender is also an important factor affecting the frequency of TB and BC and the use of particular linguistic devices or strategies. The present study shows that in single-gender conversation, males have a higher frequency of TB and BC than females. However in mixed-gender conversation, females, though they seem to speak less than males, have a higher frequency of TB and BC than males.

Moreover, females' TB occurs more frequently at a lexical boundary, whereas female BC occurs more at a prosodic or syntactic boundary. In terms of the preferred devices and strategies, females' TB has a higher amplitude during the overlap, while males' TB has a stronger tendency towards incomplete tone unit structure at the end of the overlap. Moreover, males' TB has a more frequent use of repetition than that of females.

The above summary indicates that simultaneous speech such as turn-bidding and backchannelling are constrained by different social variables. Though it could be assumed that people can overlap each other's speech at any time they want, the present findings show that people in fact tend to look for an appropriate time or place to initiate the overlap, particularly when they attempt to bid for a turn. They prefer to make TB and BC at a unit boundary, particularly at a prosodic and syntactic boundary. This is more the case in formal speech domains, in disparate conversation and in low status speakers' speech. As a whole, the overlappers in each social category tend to use similar linguistic devices in all categories of simultaneous speech. They differ only in showing a stronger or weaker preference for some devices or strategies in particular. However, due to the limited number of tokens in the data, some of the findings discussed above do not have statistical significance.

6.2 SUMMARY OF FINDINGS IN THE CASE STUDY OF CHINESE CONVERSATION

In a complementary study to the analysis of the English spoken corpus, a 10385-word sample of Chinese casual conversation was studied in comparison with a 16098-word sample of English casual conversation from the subcorpus of LLC. The study of 288 instances of Chinese simultaneous speech has found many

similarities but also some differences between the Chinese and English languages in the domain of casual conversation. It should be noticed that the Chinese data did not include telephone conversation and public discussion.

(1) Frequency of Distribution

The Chinese data (which is described in **3.5**) shows that simultaneous speech occurs frequently in Chinese casual conversation, though its frequency per thousand words (ie. 27.7/1000 w.) is not as high as that of the English sample (ie. 36.6/1000 w.). The same five categories of simultaneous speech were identified in the Chinese data as in the English data, amongst which TB (including STB, UTB, STC and UTC) accounted for 92.4% and BC accounted for only 7.6%. Compared with the English sample whose TB accounted for 60.5% and BC accounted for 39.5%, the Chinese TB makes up a greater bulk of simultaneous speech while the backchannels in this sample of Chinese simultaneous speech have a much lower frequency than in English.

(2) Environmental Features

In terms of environmental features, the most frequently occurring features before Chinese TB and BC are unit boundary markers as they are in English. Although no prosodic data is available for the Chinese case study, the syntactic and lexical boundary markers are frequently identified before the overlap. The Chinese sample reveals that 67.4% and 27.4% of the simultaneous speech tokens actually start after a syntactic boundary and a lexical (discoursal) boundary respectively in the current speaker's turn. Moreover, there is evidence that some instances of TB and BC occur in an environment where it is possible for the hearer to anticipate a syntactic and discoursal boundary. Compared with the English sample, the Chinese casual conversation has a higher frequency of discourse items occurring immediately before the overlaps because Chinese speakers tend to use more frequently the "mood auxiliaries", a type of utterance final particle which can signal the boundary

of a syntactic structure or an information "chunk". Most of the Chinese mood auxiliaries have no exact equivalents in English.

(3)   Linguistic Devices

Although the Chinese data was not prosodically transcribed, the same five sets of discourse items and the same three types of repetition were also found frequently in Chinese turn-bidding as they are in the English data. 18.1% and 22.9% of the 266 TB tokens were identified as using the discourse items and repetition strategies respectively. The difference between the two samples is that the Chinese data has a lower frequency of discourse items as initials than the English data, but the initials of continuity in Chinese TB have a higher percentage of occurrence than the other four sets of discourse items. This may be due to the pervasive usage of Chinese connective auxiliaries, since Chinese is a non-inflected language.

(4)   Social Constraints

The case study of Chinese casual conversation shows that more than one third of TB tokens occur between speakers who are of the same gender or in the same age group. In mixed-gender conversation, the Chinese females have a much larger total word count than males but a much lower frequency of TB than males. Yet gender difference does not appear to be associated with the outcome of turn bidding, ie. who wins the floor after the overlap. Both the Chinese females and males have a very similar proportion of successful TB and unsuccessful TB in their total number of tokens of cross-gender turn-bidding respectively. This is different from the results of the English sample in which the English females, though they also have a larger total word count than males, still get a similar frequency of TB to males. Yet the proportion of successful TB by English females is not as high as that of English males.

Age difference is another factor associated with turn-bidding in Chinese casual conversation. In terms of cross-generation TB, although the older speakers have a lower frequency of TB than the younger speakers, the older have a much larger proportion of successful TB than the younger. This indicates that when a turn-bidding situation occurs between an older speaker and a younger speaker, the younger speaker is more likely to yield the floor after the overlap while the older is more likely to gain the floor. As seniority is closely related to status in Chinese society, such an age difference in turn-bidding may contrast notably with the findings from the LLC subcorpus.

So far the findings show that English and Chinese simultaneous speech share many similar kinds of features and devices, though some of them are subject to different social constraints. This leads us to re-address the problem of why Chinese learners of English find it difficult to bid for a turn in English conversation. This research suggests that the reason may be both linguistic and cultural. On the one hand, this is a problem of the proficiency in English of Chinese learners. The problem refers both to their knowledge of the English language and to their practical experience in English conversation. On the other hand there are cultural differences between the two speech communities, especially in terms of speech etiquette. Both of these factors have implications for the teaching and learning of English in China.

6.3 IMPLICATIONS

The findings of the present study show how successful communication depends on the participants' shared system for signalling or negotiating shifts or transitions from one activity to another, or from one utterance to another, or even from one word to another. Chinese learners of English, because of the level of proficiency required to use such a system in English, may find it difficult to anticipate and

predict oncoming forms and content, and hence there is a lack of confidence in turn-bidding and turn-taking.

The corpus-based analysis of the LLC subcorpus in the present research provides rich information about this turn-taking-signalling system in English conversation, which can be transferred into rich input for English teaching and learning. Many applied linguists such as Long (1991) and Ellis (1993) have strongly suggested a form-focused pedagogy in foreign language teaching and learning. They advocate that learners must attend to specific linguistic features in the input and be ready to incorporate these into their interlanguage. Previous research has found clear evidence that focus on form in language teaching has some beneficial effects:

> (1) It speeds up the rate of learning (for review, see Long, 1983). (2) It affects acquisition processes in ways possibly beneficial to long-term accuracy (Lightbown, 1983; Pica, 1983). ...(3) It appears to raise the ultimate level of attainment.
>
> (Long, 1991: 45)

Yet focusing on form in language teaching and learning does not mean simply addressing a series of isolated linguistic forms such as sound contrasts, lexical items, structures, speech acts, etc. It also aims at taking the formal teaching in a direction of "consciousness-raising".

> Consciousness-raising refers to a deliberate attempt on the part of the teacher to make the learners' aware of specific features of the L2; it entails an attempt to instill an understanding of the formal and functional properties of these features by helping the learners develop a cognitive representation of them.
>
> (Ellis, 1993: 109)

The present research has described a number of linguistic features and devices of English simultaneous speech both in terms of form and function. Thus the

pedagogical implication of this study for English teaching in China is that it can enhance our efforts to

(a) select relevant materials for teaching spoken English;

(b) focus on relevant spoken English features and strategies; and

(c) practise conversational skills.

6.3.1   THE SELECTION OF MATERIALS FOR TEACHING SPOKEN ENGLISH

English teaching in China has long been concentrated on written English and has long emphasised grammar, intensive reading and memorising and reciting vocabulary. It needs to improve listening and speaking skills and train communicative strategies. The first step required is to select authentic spoken English as teaching material, because it reflects the natural use of the language. The traditional English text books in China, which concentrated mainly on written English, or artificially constructed English dialogues, have resulted in learners developing bookish, pattern-like or pre-fabricated English speech. That may be why many Chinese learners, after learning many English grammatical rules and a lot of vocabulary, still find it difficult to listen to or engage in a simple conversation with native speakers of English. Audio-video teaching materials in authentic or semi-authentic spoken English which uses data covering different fields of life in English speaking countries are often completely absent from the curriculum. Such materials could be topic-based or related to a series of themes or connected stories. At the post-intermediate level, extracts from films or videos can also be used as teaching material, because they can demonstrate polished language skills and develop meaningful ideas, which provide authentic and interesting exposure for non-native English learners.

6.3.2   THE FOCUS ON SPOKEN LANGUAGE FEATURES AND STRATEGIES

As more and more imported English spoken materials are now being used in China for language teaching, there is a need to explore the value of these materials. Listening materials used at present and audio-visual materials imported from abroad are mostly at elementary level and mainly used for word learning, pattern drilling or listening comprehension. The ways these materials reflect grammatical, lexical and prosodic features of spoken English, or cultural norms, pragmatic elements and communicative strategies, are generally neglected. For example, the discourse markers which have been shown in the present research to play an important role in turn-bidding are not likely to be handled well in most dictionaries and grammars, and they are not likely to be involved in traditional English teaching. Thus one consequence is learners' unnaturalness in English. If the goal of English teaching is to build up the learners' communicative competence, the relevant syllabus design should cover both grammatical, lexical and prosodic knowledge as teaching items, not only at the single word or single sentence level but also at a discourse level which can reveal discourse items or pragmatic strategies. As suggested by Ellis (1993: 109), the teacher should focus the learners' attention on the meaning(s) performed by specific grammatical properties so as to help their intake --- a necessary step for internalisation of the features as implicit knowledge. Pedagogy may involve activities which induce the learners to notice and understand the feature in the input. It may also involve problem-solving tasks which supply the data needed to discover the rule for themselves. Such activities and tasks will raise their consciousness about a specific language item and at the same time provide opportunities for communicating in the target language. Moreover, efforts need to be made to design motivating learning tasks or to select situations (including interesting discourse or text) which provide repeated exposure to salient, useful, frequently occurring linguistic and pragmatic elements of the language (see

Kennedy, 1992: 357). In this regard, descriptions based on the distributional analysis of large English spoken corpora like the LLC can contribute valid and reliable input for teaching.

6.3.3   THE PRACTICE OF CONVERSATIONAL SKILLS

Although the present study is on simultaneous speech, its descriptive findings imply that learners' mechanisms of anticipation and prediction not only contribute to turn-bidding and backchannelling in conversation, but are also available for reading and listening comprehension. As Chinese learners are typically taught English starting from isolated speech sounds, words and sentences, they tend to regard English as a static combination of segments of sounds and words, thus forming a habit of dealing with reading or listening materials sound by sound, or word by word. Efforts should be made by teachers to draw the attention of learners to the dynamic occurrence of coherent or cohesive "chunks" of language items such as the collocation of words, transitional devices of sentence and paragraphs, and intonation contours of discourse. Nowadays, cohesive devices have been studied a great deal in written discourse particularly in terms of lexicon and syntax, but not much in spoken discourse, and even less in terms of prosody. The present research shows that cohesive devices at prosodic, lexical, syntactic and discoursal levels play an important role in the turn-taking system of conversation, and many of them are predictable in the context as well. Therefore in training English learners' reading and listening skills, teachers may design anticipation exercises to encourage learners to predict the coming content as well as language forms by looking for relevant linguistic features or contextual cues. This can be helpful for those intermediate and advanced level English learners such as teacher trainees. For example, in an audio-visual lesson, by building their understanding of the main content of the material, the trainees should be encouraged to explore the language more deeply by picking up those salient and useful linguistic devices or pragmatic

strategies of native speakers. These may include linguistic signals of turn-taking, turn-yielding, turn-bidding, turn-competing, turn-holding or backchannelling strategies. They may also include stylistic features in communication such as the discourse intonation features and lexical features in formal and informal speaking styles, or in such styles as narration or sports commentary. They can be discussed in terms of their formal codes, pragmatic functions and cultural significance. Although evaluation of the effect of such a formal-functional learning is beyond the scope of the present study, it has already aroused the interest of some Chinese teachers and learners of English, because it makes good use of the imported teaching materials and has the potential to improve the classroom atmosphere, enrich the variety of learning exercises and stimulate the students' motivation for learning.

6.3.4   THE AWARENESS OF CULTURAL DIFFERENCES

Although cultural difference is not the only factor influencing the learners' study of English, it should not be neglected by teachers and learners. Take speech etiquette as an example. In the classroom of native English speakers, students are often encouraged to "think aloud" by using speaking to develop their ideas. Yet in Chinese classrooms, students are told not to speak in haste until they have thought about what they want to say or what they should say, because in traditional Chinese culture, to raise a question indicates that one is a poor or naive learner, and making mistakes is a matter of "losing face". Therefore teachers, especially native English teachers who are teaching in China, need to be sensitive to such cultural differences. As learning and teaching a foreign language is a process of cross-cultural communication, mutual respect and mutual understanding of cultural differences is very necessary. In other words, one object of foreign language learning is to arouse transcultural awareness in the learners, thus developing their understanding of and respect for the foreign culture, as well as their own.

The present study has not been undertaken to advise the Chinese learners of English to give up their traditional speech etiquette and become active turn-bidders in English conversation, nor does it attempt to persuade the native English teachers to get used to the silent behaviour of Chinese learners in the classroom. This study however suggests that knowledge of how native English speakers produce different kinds of simultaneous speech in conversation, particularly successful turn-bidding and appropriate backchannels, can be derived from a corpus-based analysis. Teachers and learners of English can thus contribute to inter-cultural communication through gaining greater insight into how English is used, what conversational devices and strategies are used most frequently by native speakers of English, and how they are similar or different from their own.

BIBLIOGRAPHY

Aijmer, K. 1984. "*Sort of* and *kind of* in English conversation". Studia Linguistica 38: 118-128.

Aijmer, K. 1985. "What happens at the end of our utterances? The use of utterance-final tags introduced by *and* and *or*". In O. Togeby (ed.) Papers from the English Scandinavian Conference of Linguistics: 117-127. Institute for Nordisk Filologi, University of Copenhagen.

Aijmer, K. and Altenberg, B. 1991. (eds.) English Corpus Linguistics: Studies in Honour of Jan Svartvik. London: Longman.

Altenberg, B. 1986. "Contrastive linking in spoken and written English". In G. Tottie and I. Bäcklund (eds.) English in Speech and Writing: 13-40. Stockholm: Almquist and Wiksell.

Altenberg, B. 1987. Prosodic Patterns in Spoken English. Studies in the Correlation between Prosody and Grammar for Text-to-speech Conversion. Lund Studies in English 76. Lund: Lund University Press.

Altenberg, B. 1991. "A bibliography of publications relating to English computer corpora". In S. Johansson and A-B. Stenström (eds.) English Computer Corpora: Selected Papers and Research Guide: 355-396. Berlin: Mouton de Gruyter.

Altenberg, B. 1993. ICAME Bibliography 3. Department of English. Lund University.

Atkinson, J. M. and Herritage, J. (eds.) 1984. Structures of Social Action: Studies in Conversation Analysis. Cambridge: Cambridge University Press.

Beattie, G. 1977. "The dynamics of interruption and the filled pause". British Journal of Social and Clinical Psychology 16: 283-284.

Bernstein, B. 1962. "Linguistic codes, hesitation phenomena and intelligence". Language and Speech 5: 31-46.

Biber, D. 1986. "On the investigation of spoken/written differences". Studia Linguistica 40: 1-21.

Bilous, F. R. and Krauss, R. M. 1988. "Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads". Language and Communication 8: 183-194.

Brazil, D. C., Coulthard, M. and Johns, C. 1980. Discourse Intonation and Language Teaching. London: Longman.

Brown, G. 1977. Listening to Spoken English. London: Longman.

Bublitz, W. 1988. Supportive Fellow-Speakers and Co-operative Conversations. Amsterdam: John Benjamins Publishing Company.

Carlson, L. 1984. *Well* in Dialogue Games: A Discourse Analysis of the Interjection *"well"* in Idealised Conversation. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Coates, J. 1989. "Gossip revisited: language in all-female groups". In J. Coates and D. Cameron (eds.) Women in Their Speech Communities: A New Perspective on Language and Sex: 94-122. London and New York: Longman.

Coulthard, M. and Brazil, D.C. 1979. Exchange Structure: Discourse Analysis Monographs 5. Birmingham: Birmingham University. (Reprinted in M. Coulthard & M. Montgomery (eds.) Studies in Discourse Analysis London: Routledge & Kegan Paul. 1981: 82-106).

Crosby, F. 1976. The Effect of Mode of Interaction, Sex and Acquaintance on Conversation Management. PhD Dissertation, Boston University.

Crystal, D. and Davy, D. 1969. Investigating English Style. London: Longman.

Cutler, A. and Pearson, M. 1986. "On the analysis of prosodic turn-taking cues". In C. J. Lewis (ed.) Intonation in Discourse: 139-155. London: Croom Helm.

Duncan, S. 1972. "Some signals and rules for taking speaking turns in conversations". Journal of Personality and Social Psychology 23: 283-292.

Duncan, S. 1973. "Toward a grammar for dyadic conversation". Semiotica 9: 29-46.

Duncan, S. 1974. "On the structure of speaker-auditor interaction during speaking turns". Language in Society 2: 161-180.

Duncan, S. 1981. "Conversational strategies". In T. A. Sebeok and R. Rosenthal (eds.) The Clever Hans Phenomenon: Communication with Horses, Whales, Apes and People. New York: New York Academy of Science.

Duncan, S. 1985. Interaction Structure and Strategy. New York: Cambridge University Press.

Duncan, S. and Fiske, D. W. 1977. Face-to-Face Interaction: Research, Methods and Theory. Hillsdale, N.J.: Erlbaum Associates.

Dunne, M. and Ng, S. H. 1994. "Simultaneous speech in small group conversation: all-together-now and one-at-a-time?" Journal of Language and Social Psychology 13: 45-71.

Edelsky, C. 1981. "Who's got the floor?" Language in Society 10: 383-421.

Ellis, R. 1993. "The Structural syllabus and second language acquisition". TESOL Quarterly 27-1: 91-113.

Ferguson, N. 1977. "Simultaneous speech, interruptions and dominance". British Journal of Social and Clinical Psychology 16: 295-302.

Ford, C. E. and Thompson, S. A. 1992. "Interactional units in conversation: syntactic, intonational and pragmatic resources for turn management". to appear in E. Ochs, E. Schegloff & S. A. Thompson (eds.) Interaction and Grammar. 1996, New York: Cambridge University Press.

Francis, W. N. and Kučera, H. 1964. Manual of Information to Accompany a Standard Sample of Present-Day Edited American English for Use with Digital Computers. Department of Linguistics, Brown University, Province, RI.

French, N. R., Carter, C. W. and Koenig, W. 1930. "The words and sounds of telephone conversation". Bell Syst. Tech. J. 9: 290-324.

French, P. and Local, J. 1986. "Prosodic features and the management of interruptions". In C. J. Lewis (ed.) Intonation in Discourse: 157-180. London: Croom Helm.

Goldberg, J. A. 1990. "Interrupting the discourse on interruption: an analysis in terms of relationally neutral, power-and rapport oriented acts". Journal of Pragmatics 14: 883-904.

Greenbaum, S. 1991. "The development of the International Corpus of English". In K. Aijmer and B. Altenberg (eds.) English Corpus Linguistics: Studies in Honour of Jan Svartvik: 83-91. London: Longman.

Greenbaum, S. and Svartvik, J. 1990. "The London-Lund Corpus of Spoken English". In Svartvik (ed.) The London-Lund Corpus of Spoken English: Description and Research: 83-91. Lund Studies in English 82. Lund: Lund University Press.

Grice, H. 1975. "Logic and conversation". In P. Cole and J. Morgan (eds.) Syntax and Semantics 3: Speech Acts: 41-58. New York: Academic Press.

Gumperz, J. 1982. Discourse Strategies. Cambridge: Cambridge University Press.

Halliday, M. A. K. 1970. A Course in Spoken English: Intonation. Oxford: Oxford University Press.

Halliday, M. A. K. 1985. An Introduction to Functional Grammar. London: Edward Arnold..

Halliday, M. A. K. 1989. Spoken and Written Language. Hong Kong: Oxford University Press.

Hatch, E. M. 1992. Discourse and Language Education. New York: Cambridge University Press.

Hess, L. J. and Johnston, J. R. 1988. "Acquisition of back channel listener responses to adequate messages". Discourse Processes 11: 319-335.

Holmes, J. 1991. "Language and gender". Language Teaching 24: 207-220.

Jaffe, J. and Feldstein, S. 1970. Rhythms of Dialogue. New York: Academic Press.

James, D. and Clarke, S. 1993. "Women, men, and interruption: a critical review". In D. Tannen (ed.) Gender and Conversational Interaction: 231-280. New York: Oxford University Press.

Jefferson, G. 1973. "A Case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences". Semiotica 9: 47-96.

Jin, L. 1992. Academic Cultural Expectations and Second Language Use: Chinese postgraduate students in the UK --- a cultural synergy model. Unpublished doctoral dissertation, University of Leicester.

Jin, L. and Cortazzi, M. 1995. "A cultural synergy model for academic language use" in P. Bruthiaux, T. Boswood & B. D. Babcock (eds.) Explorations in English for Professional Communication: 215-244. Hong Kong: City University of Hong Kong.

Johansson, S., Leech, G. N. and Goodluck, H. 1978. Manual of Information to Accompany the Lancaster-Oslo/Bergen Corpus of British English, for Use with Digital Computers. Department of English, University of Oslo.

Kendon, A. 1967. "Some functions of gaze-direction in social interaction". Acta Psychologica 26: 22-63.

Kendon, A. 1977. Studies in the Behaviour of Social Interaction. Bloomington: Research Centre for Language and Semiotic Studies.

Kendon, A., Harris, R. M. and Key, M. R. (eds.) 1975. Organisation of Behaviour in Face-to-Face Interaction. The Hague: Mouton Publishers.

Kennedy, C. W. and Camden, C. T. 1983. "A new look at interruptions". The Western Journal of Speech Communication 47: 45-58.

Kennedy, G. 1992. "Preferred ways of putting things with implications for language teaching". In J. Svartvik (ed.) Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82: 335-373. Berlin: Mouton de Gruyter.

Kennedy, G. 1996. "The corpus as a research domain". In S. Greenbaum (ed.) Comparing English Worldwide: The International Corpus of English: 217-226. Oxford: Oxford University Press.

Kjellmer, G. 1992. "Comments". In J. Svartvik (ed.) Directions in Corpus Linguistics Proceedings of Nobel Symposium 82: 374-378. Berlin: Mouton de Gruyter.

Lancashire, I. 1995. "Review of directions in corpus linguistics: proceedings of Nobel Symposium 82". ICAME Journal 19: 93-109.

Leech, G. 1983. Principles of Pragmatics. London: Longman.

Leech, G. 1991. "The state of art in corpus linguistics". In K. Aijmer and B. Altenberg (eds.) English Corpus Linguistics: Studies in Honour of Jan Svartvik: 8-29. London: Longman.

Levinson, S. 1983. Pragmatics. New York: Cambridge University Press.

Li, C. N. and Thompson, S. A. 1981. Mandarin Chinese: A functional reference grammar. Berkeley: University of California Press.

Lightbown, P. M. 1983. "Exploring relationships between developmental and instructional sequences". In H. W. Seliger and M. H. Long (eds.) Classroom-Oriented Research on Second Language Acquisition: 217-243. Rowley, MA: Newbury House.

Liu Guanghui and Shi Peiwen. 1988. A Practical Course in Chinese Phonetics. Beijing: Economy Administration Press.

Liu Yuehua, Pan Wenyu and Gu Hua. 1983. Chinese Grammar. Beijing: Foreign Language Teaching and Studying Press.

Long, M. H. 1983. "Does instruction make a difference? A review of research". TESOL Quarterly 17-3: 359-382.

Long, M. H. 1991. "Focus on form: a design feature in language teaching methodology". In Rees de Bot, R. B. Girsberg and C. Kramsch (eds.) Foreign Language Research in Cross-Cultural Perspective: 39-52. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Luke, K. K. 1990. Utterance Particles in Cantonese Conversation. Amsterdam: John Benjamins Publishing Company.

Markel, N. 1975. "Coverbal behaviour associated with conversation turns". In A. Kendon, R. M. Harris & M. R. Key (eds.) Organisation of Behaviour in Face-to-Face Interaction: 189-197. The Hague Mouton; Chicago: distributed in the USA and Canada.

Maynard, S. 1986. "On back-channel behaviour in Japanese and English casual conversation". Linguistics 24: 1079-1108.

McLachlan, A. 1991. "The effects of agreement, disagreement, gender and familiarity on patterns of dyadic interaction". Journal of Language and Social Psychology 10-3: 205-212.

Meltzer, L., Morris, W. N. and Hayes, D. P. 1971. "Interruption outcomes and vocal amplitude: explorations in social psychophysics". Journal of Personality and Social Psychology 18: 392-402.

Miller, G. A. 1963. Language and Communication. New York: McGraw-Hill.

Mishler, E. G. and Waxler, N. E. 1968. Interaction in Families: An Experimental Study of Family Processes and Schizophrenia. New York: Wiley.

Murata, K. 1994. "Intrusive or co-operative? A cross-cultural study of interruption". Journal of Pragmatics 21 :385-400.

Ng, S. H., Bell, D. and Brook, M. 1993. "Gaining turns and achieving influence ranking in small conversational groups". British Journal of Social Psychology 32: 265-275.

O'Connor, J. D. 1967. Better English Pronunciation. Cambridge: Cambridge University Press.

Persson, G. 1974. Repetition in English Part I: Sequential Repetition. Uppsala: Almqvist and Wiksell.

Pica, T. 1983. "Adult acquisition of English as a second language under different conditions of exposure". Language Learning 33-4: 465-497.

Psathas, G. 1995. Conversation Analysis: The Study of Talk-in-Interaction. Qualitative Research Methods Series Vol. 35. Thousand Oaks, California: Sage Publications.

Roach, P. 1983. English Phonetics and Phonology: A Practical Course. Cambridge: Cambridge University Press.

Roger, D. 1989. "Experimental studies of turn-taking behaviour". In D. Roger and P. Bull (eds.) Conversation: An Interdisciplinary Perspective: 75-95. Philadelphia, PA: Multilingual Matters.

Roger, D., Bull, P. and Smith, S. 1988. "The development of a comprehensive system for classifying interruptions". Language and Social Psychology 7-1: 27-34.

Roger, D. and Schumacher, A. 1983. "Effects of individual differences on dyadic conversational strategies". Journal of Personality and Social Psychology 45: 700 -705.

Sacks, H. 1989. Lectures. 1964-65 (G. Jefferson, ed.; E. A. Schegloff, "Introduction/Memoir") Human Studies 12: 3-4.

Sacks, H., Schegloff, E. and Jefferson, G. 1974. "A simplest systematics for turn-taking for conversation". Language 50: 696-735.

Schegloff, E. 1981. "Discourse as an interactional achievement: some uses of "uh huh" and other things that come between sentences". In D. Tannen (ed.) Analysing Discourse: Text and Talk: 71-93. (Georgetown University Round Table on Languages and Linguistics 1981) Washington DC: Georgetown University Press.

Schegloff, E. 1987. "Recycled turn beginning: a precise repair mechanism in conversation's turn-taking organisation". In G. Button and J. R. E. Lee (eds.) Talk and Social Organisation: 70-85. Avon: Multilingual Matters.

Schegloff, E. and Sacks, H. 1973. "Opening up closings". Semiotica 8: 289-327.

Schiffrin, D. 1987. Discourse Markers. Cambridge: Cambridge University Press.

Schiffrin, D. 1994. Approaches to Discourse. Oxford, UK; Cambridge, Mass., USA: B. Blackwell.

Schourup, L. C. 1985. Common Discourse Particles in English Conversation: like, well, y'know. New York: Garland.

Sinclair, J. M. and Coulthard, R. M. 1975. Towards an Analysis of Discourse: the English Used by Teachers and Pupils. London: Oxford University Press.

Sinclair, J. M and Renouf, A. 1988. "A lexical syllabus for language learning". In R. Carter and M. McCarthy (eds.) Vocabulary and Language Teaching: 140-160. London: Longman.

Stenström, A-B. 1984. "Discourse Tags". In Aarts-Meijs (eds.) Corpus Linguistics: Recent Development in the Use of Computer Corpora in English Language Research: 65-81. Amsterdam: Rodopi.

Stenström, A-B. 1990. "Lexical items peculiar to spoken discourse". In J. Svartvik (ed.) The London-Lund Corpus of Spoken English: Description and Research: 137-176. Lund Studies in English 82. Lund: Lund University Press.

Stenström, A-B. 1994. An Introduction to Spoken Interaction. London and New York: Longman.

Street, R. L., Jr. and Murphy, Thomas, J. 1987. "Interpersonal orientation and speech behaviour". Communication Monographs 54: 42-62.

Stubbe, M. 1991. Talking at Cross-purposes? The Effect of Gender on New Zealand Schoolchildren's Interaction Strategies in Pair Discussion. MA Thesis, Victoria University, Wellington, New Zealand.

Svartvik, J. (ed.) 1990. The London-Lund Corpus of Spoken English: description and research. Lund Studies in English 82. Lund: Lund University Press.

Svartvik, J. and Quirk, R. (eds.) 1980. A Corpus of English Conversation. Lund: Lund University Press.

Tannen, D. 1983. "When is an overlap not an interruption in cross-sex conversation? One component of conversational style". In R. J. DiPietro, W. Frawley and A. Wedel (eds.) The First Delaware Symposium on Language Studies: 119-129. Newmark, DE: University of Delaware Press.

Tannen, D. 1984. Conversational Style : Analysing Talk among Friends. Norwood, N.J.: Ablex.

Tannen, D. 1986. That's not what I meant! How Conversational Style Makes or Breaks Your Relation with Others. New York: William Morrow.

Tannen, D. 1989. Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse. Cambridge: Cambridge University Press.

Tannen, D. 1990. You Just Don't Understand: Women and Men in Conversation. New York: William Morrow.

Tao, H. and Thompson, S. A. 1991. "English backchannels in Mandarin communication : a case study of superstratum pragmatics interference". Journal of Pragmatics 16-3: 209-223.

Taylor, T. J. and Cameron, D. 1987. Analysing Conversation. New York: Pergamon Press.

West, C. 1979. "Against our will: male interruptions of females in cross-sex conversations". in J. Orasanu, M. K. Slater & L. L. Adler (eds.) Language , Sex and Gender: 81-97. New York: Annals of the New York Academy of Sciences.

West, C. and Zimmerman, D. H. 1983. "Small insults: a study of interruptions in cross-sex conversation between unacquainted persons". In B. Thorne., C. Kramarae and N. Henley (eds.) Language Gender and Society: 103-117. Rowley, MA: Newbury House.

White, S. 1989. "Backchannels across culture: a study of Americans and Japanese". Language and Society 18: 59-76.

Wilson, T. P., Wiemann, J. M. and Zimmerman, D. H. 1984. "Models of turn-taking in conversation interaction". Journal of Language and Social Psychology 3.3:159-183.

Yngve, V. H. 1970. "On getting a word in edgewise". In Papers from the Sixth Regional Meeting of the Chicago Linguistic Society: 567-578. Chicago: University of Chicago.

Zimmerman, D. H. and West, C. 1975. "Sex roles, interruptions and silence in conversation". In B. Thorne and N. Henley (eds.) <u>Language and Sex: Difference and Dominance</u>: 105-129. Rowley, MA: Newbury House.

Zimmerman, D. H. and West, C. 1977. "Doing gender". Paper given at the meeting of the American Sociological Association. September. Chicago.

Zipf, G. K. 1935. <u>The Psycho-biology of Language: an Introduction to Dynamic Philology</u>. Boston: Houghton Mifflin.