

Facilitating a description of intercultural conversations: the Hong Kong Corpus of Conversational English

Winnie Cheng and Martin Warren

Department of English, The Hong Kong Polytechnic University

1 Background

The relative difficulty with which spoken corpora can be compiled by the researcher compared with written discourses, coupled with the time needed to fully transcribe spoken data, to say nothing of the additional expenses involved, inevitably has made large spoken corpora a far rarer entity than written corpora. And yet, if we are to further unravel the intricacies of spoken discourse, we need to work with large, representative corpora of spoken discourses. The two largest corpora of the English language, the Bank of English and the British National Corpus, devote approximately 10 per cent of their corpora to spoken English. This is not to say that spoken discourse has been forgotten or ignored by corpus linguists. The London-Lund Corpus, for example, has been an invaluable source of data for many years, and projects such as COLT (see eg Andersen and Stenström 1996) and CANCODE (see eg Carter and McCarthy 1997; McCarthy and Carter 1997) are evidence of more recent initiatives to compile spoken corpora.

Conversations are the most prevalent form of discourse, accounting for more than 90 per cent of all spoken language according to Svartvik (1980) and yet, compared to more specialised spoken discourse types and written discourses, they do not receive proportionately anything like that amount of attention from researchers investigating language use. As mentioned above, only a relatively small proportion of the Bank of English and the British National Corpus is made up of naturally-occurring conversations, as opposed to other types of spoken discourse. Almost two decades after Svartvik and others had made the London-Lund Corpus available in scholarly form, Svartvik (1996: 10) was still bemoaning the fact that ‘conversation – the quintessence of spoken language – is either missing or seriously underestimated in most existing corpora’. However, while there have been relatively few studies on native speaker conversations (see eg Stenström 1994; Tsui 1994; Eggins and Slade 1997), naturally-occurring non-

native/non-native or non-native/native speaker conversations have received even less attention.

In Hong Kong, little attempt has been made to date to offer systematic descriptions of naturally-occurring English conversations involving Hong Kong Chinese participants (ie mother tongue speakers of Cantonese). To embark on such descriptions, a corpus of naturally-occurring conversations was greatly in need and, until recently, no suitable corpus had existed. The Hong Kong component of the International Corpus of English Project (see eg Bolt 1994: 15) includes recordings of private conversations, but these discourses fall short of being naturally-occurring, due to the constraints imposed by the coordinators of the overall project in London. The compilation of the Hong Kong Corpus of Conversational English (HKCCE), housed in the English Department of The Hong Kong Polytechnic University is, therefore, the creation of a unique and useful resource. The HKCCE is what Sinclair (1991: 23–26) terms a ‘sample corpus’ and does not claim to hold the state of conversational English language in Hong Kong in the way that a ‘monitor corpus’, such as the Bank of English, might be said to contain the state of the English language.

2 Why a corpus of Hong Kong conversational English?

Interest in analysing and describing the English produced by Hong Kong Chinese in different contexts for different purposes has been growing in recent years. There have been studies of classroom talk and academic discourses (eg Flowerdew 1994; Lockhart and Ng 1994; Pennington 1995; Tsui 1985, 1987, 1991). Others have examined spoken business, professional, public and broadcast discourses (eg Leung 1996; Bilbow 1997a, 1997b; Flowerdew 1997; Scollon and Flowerdew 1997). There have also been studies adopting a conversation-analytic approach to examine an aspect of the spoken discourse produced by local Hong Kong Chinese as, for example, the study of responses to compliments between British and Chinese speakers (Loh 1993) and conversational sequences in English and Chinese (Luke 1996). Regarding the nature of the English spoken by Hong Kong people, Bolton and Kwok (1990) provide a sketch of the Hong Kong accent, which is based on recordings of radio and television news, Legislative Council speeches (ie parliamentary speeches) and interviews with employees of various companies and with university students. However, none of the studies to date have been based on a representative corpus of naturally-occurring conversations. We are of the opinion that conversations are a benchmark for other spoken discourses, and by more fully describing con-

versational English in Hong Kong we will better understand the ways in which other spoken discourses differ from it.

3 What is a conversation?

A major problem in compiling a corpus of conversations is to have a clear idea of just what is and what is not to be included. To date, there has not been any generally accepted definition of a conversation as a speech event. Definitions given differ greatly, and conversations are often described as ranging somewhere between casual talk in everyday settings and being synonymous with spoken interaction in general. This use of the term ‘conversation’ as a catchall for any type of spoken discourse is quite commonplace. McGregor (1984: 72–73) states that although researchers of various disciplines, including ethnomethodologists, insist on using and examining “‘everyday’ language use’, their data of ‘conversation’ rarely meets the requirements of their own terms of reference. The sources of data which are described as conversations include ‘job interviews, telephone calls, joke and story telling episodes, teacher-pupil exchanges, ritual insults, and a range of reported and recorded psychiatric, psychotherapeutic and psychological investigative activities’ (McGregor 1984: 73). In Ricento’s (1987: 757–758) study on multiparty conversation, for instance, ‘spontaneous conversations’ are on a par with ‘conversational re-telling’ of a film. Francis and Hunston (1992: 123–124) equate telephone conversation to everyday conversation. Eggins and Slade (1997: 20) use ‘conversation’ to encompass both ‘pragmatic conversation’ and ‘casual conversation’.

For our purposes, such non-specific use of the term ‘conversation’ is unsuitable for two main reasons. First, we were aiming to collect 50 hours (approximately 500,000 words) of data, and not to be specific in terms of the discourse type would render the corpus useless in terms of serious analysis, as it may have ended up as a collection of very disparate discourses. Second, we believe that conversations do constitute a distinctive form of spoken discourse and that to conflate them with spoken discourse in general is to suggest that there are no criteria for distinguishing conversations from what we term ‘specialised discourse types’ after Abercrombie (1965). Swales (1990: 43) argues that conversation, as the pre-eminent form of language, is a pre-genre in the sense that all genres, both spoken and written, are derived from it. Similarly, Fillmore (1981) argues that conversational language constitutes the benchmark against which other forms of the language can be compared and contrasted and that ‘once the syntax, semantics, and pragmatics of these basic types of discourse have been mastered, other types of discourse can be usefully described in terms of their deviation

from such a base' (1981: 165). For our purposes, we have used the definition of a conversation given by Warren (1993: 8) as a set of guidelines for the research team to determine the contents of the HKCCE. This definition is given and explained below:

A speech event outside of an institutionalised setting involving at least two participants who share responsibility for the progress and outcome of an impromptu and unmarked verbal encounter consisting of more than a ritualised exchange.

An essential element in a conversation is the active involvement of at least two participants outside of an institutionalised setting. Levinson (1983: 284) defines conversation as 'that familiar pre-dominant kind of talk in which two or more participants freely alternate in speaking, which generally occurs outside specific institutional settings like religious services, law courts, classrooms and the like'. Abercrombie (1965: 2–9) distinguishes between conversation and other kinds of discourse, which he describes as 'specialised discourse types'. For the purposes of this project, the term 'specialised discourse types' is used to describe institutionalised spoken discourses that have been excluded from the HKCCE.

'Institutionalised setting' is used to describe the backdrop for specialised discourse types in which the equality of speaking rights (Wilson 1987, 1989), which is a characteristic of conversations, is suspended. Indeed, Wilson (1987: 96) defines a conversation solely in terms of the 'equal distribution of speaker rights', in recognition of the fact that, in conversation, speakers have equal rights in terms of initiating talk, interrupting, responding, or deciding not to do any of these. Equality of speaking rights is however absent from specialised discourse types, in which certain designated speakers control the speaking rights of the other participants.

Linked to the notion of equality of speaking rights is the relative status of the participants in conversations. Cheepen and Monaghan (1990: 16) include the equality of status among the participants in their definition of conversation. For the duration of a conversation the external status of the participants is set aside, and for the purposes of conducting the conversation the participants are deemed to be of equal status. In this way the participants perceive themselves to be of equal status for the purposes of holding a conversation. Again, this sets conversations apart from specialised discourse types in which the status of the participants is unequal, which in turn has consequences for the resulting discourse.

In the definition of conversation employed in this project, the perceived equality of status between the participants is subsumed within the notion of

‘shared responsibility’ (Warren, 1993: 314). In conversation, the negotiated beginning, progress, and ending of the discourse are the shared responsibility of all of the participants. Even if, in reality, a particular conversation is dominated verbally by one or more of the participants, the responsibility for the discourse is still shared. This is however not the case in specialised discourse types, in which it is the speaker who is designated as dominant has the ultimate responsibility for the discourse. Shared responsibility, therefore, is considered to be a basic ingredient in conversations that contributes to many of the distinctive characteristics found in them. The shared responsibility between the participants is closely connected with the ‘progress and outcome’ (Warren 1993: 8) of a conversation, which is related to the open-endedness of what is said, when and by whom. The unpredictability in topic changes (Sinclair and Coulthard 1975: 4) and ‘topic shift without the context of a predetermined topic schedule’ (Cheepen and Monaghan 1990: 19) are characteristics of conversation.

Conversations are defined as ‘impromptu’ speech events, in which the agenda is not fixed, predetermined or controlled by any designated participants. Instead, as stated by Biber (1988: 71), they are characterised by a high degree of interaction and goal negotiability, considerable effort at maintaining a relationship, and considerable shared background knowledge. Conversations are also regarded as ‘unmarked’ verbal encounters, so as to set them apart from discourse types which are not institutionalised discourses and which are similar to conversation in certain respects. For instance, telephone calls are marked in their highly predictable opening and closing sequences (see eg Schegloff 1972; Schegloff and Sacks 1973) and in that the caller, at least, usually has a reason for making the call. In fact, a caller who has no specific reason for making his/her call invariably makes this fact known early on in the discourse, which demonstrates that this is not the norm. Another example of a marked discourse is arguments, which are marked deviations from conversation as they lack some basic properties of conversations, such as the politeness formulas used in the maintenance of face and the open-endedness of topic.

Lastly, the length and content of the discourse is also an important consideration in our definition of a conversation. A verbal exchange needs to go beyond the mere exchanging of greetings to be classified as a conversation. This definition does not include ritualised exchanges such as *hello/hello* and/or politeness formulas such as *how are you/fine thanks*. Goffman (1971: 17) describes ritualised exchanges such as these as ‘minimal’ conversations, which he claims function to show that nothing has changed since the last encounter.

4 Methodology

4.1 Naturally-occurring data

Throughout the data collection stage, we placed great importance on the collection of spontaneous naturally-occurring conversations in which the participants could be assumed to be behaving normally. By emphasising the nature of the data, it was believed that the credibility of the findings would be enhanced and the extent to which they may be generalised increased. Roger (1989: 71–2) maintains that the main difference between experimental data collection and naturalistic methods is the reason for the talk. In the former, it is because the subjects are requested to talk, while in the latter it is of their own volition, and this major difference must have an impact on the resulting discourse. It was for this reason that we avoided the contrived nature of speech situations that are specifically arranged for data collection. This position is supported by Taylor and Cameron (1987: 52), who state that there is no guarantee that data obtained from experimental contexts is in any way representative of talk in the real world.

In order to record naturally-occurring English conversations in Hong Kong, the presence of a native speaker, or a speaker of a language other than Cantonese, is necessary to ensure that it is English and not Cantonese that is spoken. To comply with the ethical standards stipulated by our university's Ethics Committee for any research project involving human subjects, certain conditions had to be met. All the participants in the conversations were asked to sign a consent form, which clearly states the purposes of the research and the likelihood that the data could be made available in electronic format or on CD-ROM. No names of individuals were transcribed in order to ensure anonymity of the participants and others. Similarly, the names of the participants have not been and will not be made public.

4.2 Recording equipment

All of the conversations were digitally audio-recorded on mini-discs (MDs) using Sony MD recorders (model nos MZ-R2, MZ-R3 and MZ-R50) with a Sony stereo microphone (model no ECM-717). This equipment was used for the superior quality compared with audio or digital audio tapes, and because it is digitally recorded, the data are more readily accessible for computer-aided analyses. The MD recorders we used are very small (approximately 12x8x2 cm), and so are the microphones, making the recording equipment unobtrusive from the participants' points of view.

4.3 Conversational participants

There was a danger that the corpus could have ended up with a preponderance of workplace conversations, involving expatriate and Hong Kong Chinese colleagues based in the English Department of our university, which would have made our task relatively easy, but would not have captured ‘typical’ Hong Kong conversational English. From the outset, we decided not to collect data in which the non-native speaker was a teacher of English. In order to obtain a better representation of Hong Kong conversational English, the research team have used as wide a range of contacts as possible, to gather suitable data from a variety of settings. Student research assistants were also employed to collect data randomly around Hong Kong along with members of the research team, colleagues, friends, etc. Those responsible for collecting the data did not participate in the conversations and, once consent was given by the participants, they retreated to allow the conversations to proceed as naturally as possible. The recordings were made in a wide variety of locations in Hong Kong, including homes, restaurants, coffee shops, pubs, clubs, workplaces, public parks, and social gatherings.

In addition to providing the written consent of the conversational participants, the consent form consisted of a number of questions that have furnished us with useful background information of the participants. This information is detailed below:

- Gender
- Age group
- First language
- Place of birth
- Occupational group
- Periods of residency overseas in excess of six months
- Educational background

We are now in the process of entering the above information into the HKCCE against each of the participants in a machine-readable form, to enable the user to effectively create mini-corpora from the main corpus. In this way, a user could, for example, access all of the utterances made by young female professional Hong Kong Chinese who have studied overseas for more than five years.

4.4 Transcription conventions

Basically, the data is orthographically transcribed, but interactional features like overlaps, latched utterances, pauses, non-linguistic features, and indecipherable

utterances are indicated as accurately as possible. Notation conventions first developed by Gail Jefferson (1984) are used and these are given as follows:

Symbol	Meaning
...	Intervening utterances which have been omitted.
[Simultaneous or overlapping utterances.
[
end of line =	Speech delivered in quick succession to one another with no
=beginning	silence, ie latched utterances.
of line	
	or
	An intervening interruption of a single speaker's utterance.
(.)	A short, untimed pause or a gap of silence (a micro-pause), equal to a one-syllable-length pause.
(pause)	Longer pause.
heh, hah, huh	Laughter.
((sniff))	Non-linguistic features: vocalisations and non-verbal aspects
((door bell rings))	of the speech situation.
()	Indecipherable utterances.
(class)	Doubtful transcription, as in the case of softly spoken or overlapped utterances.
A/a:	Speakers are identified by letters. The first speaker is A/a; the
B/b:	second B/b. The upper case stands for native speakers; the lower case Hong Kong Chinese.

The excerpt below represents a typical extract of transcription taken from the HKCCE:

a: female Hong Kong Chinese

B: female English

- 308 B: [((laughs)) studying
 309 a: then when I come back I have to work on my PhD[(.) so erm so they say
 310 B: [yea
 311 a: [that I am always in front of the eh eh eh eh TV screen but actually that's
 312 B: [(yea)
 313 a: not TV screen that is a [computer ((laughs))

314 B: [computer ((laughs))
315 B: oh that's very admirable that is very good
316 (pause)

We decided against punctuating the transcripts and writing phonetic transcription for words or groups of words which are spoken informally and quickly ([d̥əz] for *does*) or pronounced in a reduced way (*wanna* for *want to*) (see Carter and McCarthy 1997: 20). We felt that a simple and standard transcription would provide the level of detail that is necessary for the purposes of analysing and describing the conversational data. A lack of sufficient funding has also prevented us from adding any form of prosody indication to our orthographic transcription. Most of the initial transcriptions were done by student research assistants with the aid of Sony transcribers (model no BM-77T). These transcriptions were then checked by full-time research personnel and finally double-checked by a member of the project team.

5 Contents of the HKCCE

The HKCCE comprises 50 hours of conversations made up of 130 separate conversations involving a total of 341 participants. The conversations average about 23 minutes in length, with the longest conversation lasting 1 hour 15 minutes and the shortest 2 minutes 49 seconds. The total number of words is approximately 500,000, and we have found that conversations seem to be fairly consistent in generating about 10,000 words per hour. In the following, the HKCCE is described in terms of gender distribution across native speakers (NS) and non-native speakers (NNS), age distribution, educational background, and occupations.

Figure 1 shows the proportions of female and male NS and NNS among the participants:

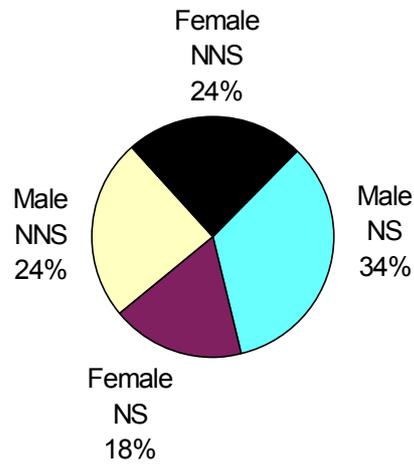


Figure 1: Overall percentage of female and male NS and NNS in the HKCCE

Figure 2 gives a breakdown in terms of the number of participants in each of the conversations collected. Over half of the conversations are dyadic, followed by 23 per cent with three speakers, 10 per cent with four speakers, and 9 per cent more than four speakers participating in the conversations.

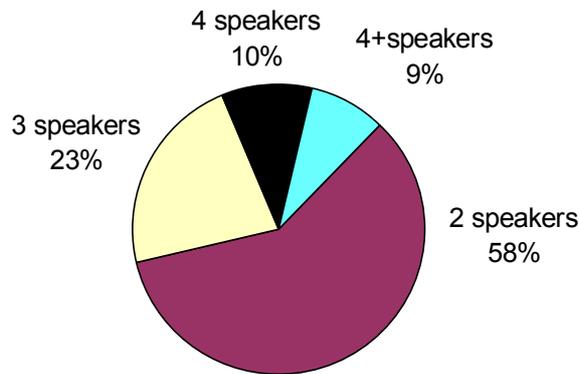


Figure 2: Number of participants in the conversations in the HKCCE

Concerning age distribution of the participants in HKCCE (see Figure 3), 40 per cent are within the range of 18–29 years of age; similar percentages are within the age ranges 40–49 and 30–39.

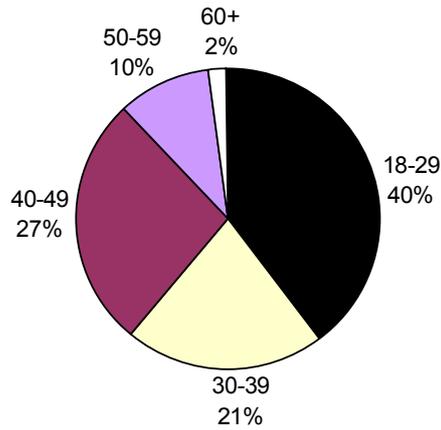


Figure 3: Age distribution of participants in the HKCCE

Figure 4 shows the educational background of all the participants in the Corpus. The great majority of them have studied to university level.

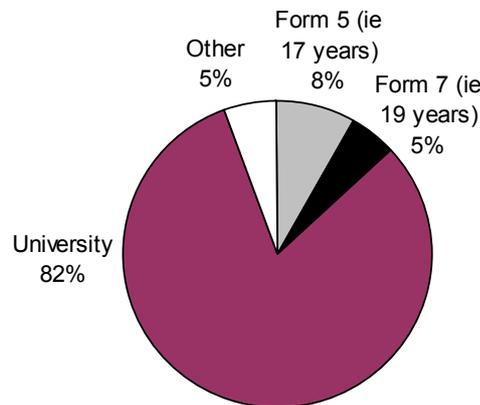


Figure 4: Distribution of educational levels of participants in the HKCCE

The participants in the HKCCE are found to be quite diverse in their occupations (see Figure 5) but the percentages of participants working in the fields of Education (35%) and Business (23%) are relatively higher than those in occupations such as Administration, Engineering, Service, Arts, and so on.

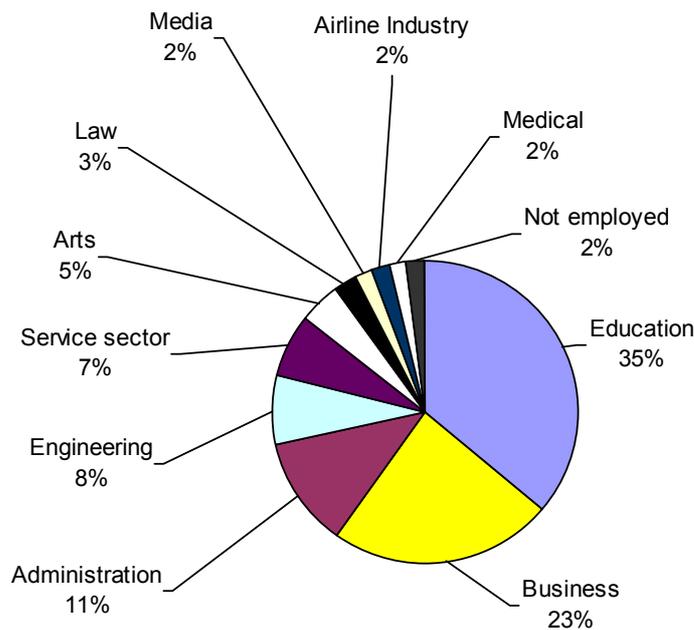


Figure 5: Distribution of occupations of participants in the HKCCE

6 Applications of the HKCCE

The combination of NNS and NS conversing in similar contexts means that the HKCCE is a source of intercultural studies. We have already conducted research based on the HKCCE which has revealed interesting similarities and differences between the two sets of speakers in terms of the use discourse markers (see eg Cheng and Warren 1998), tag questions (Cheng 1996), vague language (Cheng et al 1996/97), and the level of inexplicitness employed (Cheng and Warren forthcoming). The possibilities for exploiting the corpus in English language learning and teaching materials have been explored (Cheng and Warren 1998),

and it is hoped these will be extended if we are able to make the HKCCE available on the WWW. Undergraduate and postgraduate students have made use of the corpus in their projects and theses, and again we hope that the wider research community will find our corpus of value. In our teaching of academic subjects such as Discourse Analysis, Pragmatics, and Intercultural Communication, we have found the existence of a corpus of conversational data very useful.

7 Future developments

We are hoping to obtain funding in the future to add intonation markings to the transcripts of the conversations along the lines suggested by Brazil (1997). This would mean marking tone units, tone, prominence, key and termination. We are also planning that the HKCCE would become a sub-corpus of a Hong Kong Corpus of Spoken English, which would also contain a sub-corpus of academic discourses and a sub-corpus of business discourses. We also hope to put our corpus (corpora) on the WWW and to make it available as a CD-ROM, whereby the user could both hear and read the corpus contents.

Acknowledgements

The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administration Region (Project No G-S642). Our thanks go to Ms Alice Lo for helping to prepare the charts.

References

- Abercrombie, David. 1965. *Studies in phonetics and linguistics*. London: Oxford University Press.
- Andersen, Gisle and Anna-Brita Stenström. 1996. COLT: A progress report. *ICAME Journal* No 20: 133.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Bilbow, Grahame. 1997a. Spoken discourse in the multicultural workplace in Hong Kong: Applying a model of discourse as 'impression management'. In F. Bargiela-Chiappini and S. Harris (eds) *The languages of business: an international perspective*, 21–48. Edinburgh: Edinburgh University Press.

- Bilbow, Grahame. 1997b. Cross-cultural impression management in the multi-cultural workplace: The special case of Hong Kong. In R. Scollon and J. Flowerdew (eds) *Journal of Pragmatics, special issue on 'Language and discourse issues in Hong Kong's change of sovereignty'* 28: 461–488.
- Bolt, Philip. 1994. The International Corpus of English Project: the Hong Kong experience. In U. Fries, G. Tottie and P. Schneider (eds) *Creating and Using English Language Corpora: Papers from the Fourteenth International Conference on English Language Research on Computerized Corpora, Zurich 1993*, 15–24. Amsterdam: Rodopi.
- Bolton, Kingsley and Helen Kwok. 1990. The dynamics of the Hong Kong accent: social identity and sociolinguistic description. *Journal of Asian Pacific Communication* 1(1): 147–172.
- Brazil, David. 1997. *The communicative value of intonation in English*. Cambridge: Cambridge University Press.
- Carter, Ronald and Michael McCarthy. 1997. *Exploring spoken English*. Cambridge: Cambridge University Press.
- Cheepen, Christine and James Monaghan. 1990. *Spoken English: A practical guide*. London: Pinter.
- Cheng, Winnie. 1996. *We've heard this before, isn't it?* Tag questions in Hong Kong conversational English. Paper presented at 'Worlds of Discourse', Twenty-first International Conference of the Applied Linguistics Association of Australia, Sydney, Australia, 3–6 October 1996.
- Cheng, Winnie, Warren, Martin, Bradshaw, Julie, Greaves, Chris, Pellowe, John and Jane Setter. 1997. Compiling a Hong Kong corpus of conversational English. *HKPU Working Papers in ELT and Applied Linguistics*, 2(2): 1–20.
- Cheng, Winnie and Martin Warren. 1998. Exploiting the Hong Kong Corpus of Spoken English for research, learning and teaching purposes. *Proceedings of Teaching and Language Corpora 98*. Keble College, Oxford University, 33–41.
- Cheng, Winnie and Martin Warren. Forthcoming. Inexplicitness: What is it and should we be teaching it? *Applied Linguistics*.
- Eggins, Suzanne and Diana Slade. 1997. *Analysing casual conversation*. London and Washington: Cassell.
- Fillmore, Charles. 1981. Pragmatics and the description of discourse. In P. Cole (ed) *Radical pragmatics*, 143–166. New York: Academic Press.
- Flowerdew, John. (ed). 1994. *Academic listening*. Cambridge: Cambridge University Press.

- Flowerdew, John. 1997. Competing public discourses in transitional Hong Kong. In R. Scollon and J. Flowerdew (eds) *Journal of Pragmatics, special issue on 'Language and discourse issues in Hong Kong's change of sovereignty'* 28: 533–554.
- Francis, Gill and Susan Hunston. 1992. Analysing everyday conversation. In M. Coulthard (ed) *Advances in spoken discourse analysis*, 123–61. London: Routledge.
- Goffman, Erving. 1971. *Relations in public*. New York: Harper and Row.
- Jefferson, Gail. 1984. On the organization of laughter in talk about troubles. In J. M. Atkinson and J. Heritage (eds) *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press, 346–369.
- Leung, May C.M. 1996. The use of the discourse markers *you know, well* and *so on* in Hong Kong English and British English. Unpublished MPhil thesis. The Hong Kong Polytechnic University.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Lockhart, Charles and Peggy Ng. 1994. *Analyzing talk in ESL peer response groups: Stances, function, and content*. Hong Kong: Department of English, City University of Hong Kong.
- Loh, Teresa. 1993. *Responses to compliments across languages and cultures: A comparative study of British and Hong Kong Chinese*. Hong Kong: Department of English, City University of Hong Kong.
- Luke, Kang Kwong. 1996. The structure of conversational sequences in Chinese and English (ms).
- McCarthy, Michael and Ronald Carter. 1997. Grammar, tails, and affect: Constructing expressive choices in discourse. In T. van Dijk (ed) *Text: An interdisciplinary journal for the study of discourse* 17(3): 405–429. Berlin: Mouton Publishers.
- McGregor, Grahame. 1984. Conversation and communication. *Language and Communication* 4(1): 71–83.
- Pennington, Martha. 1995. *Eight case studies of classroom discourse in the Hong Kong secondary English class*. Hong Kong: Department of English, City University of Hong Kong.
- Ricento, Thomas. 1987. Clausal ellipsis in multi-party conversation in English. *Journal of Pragmatics* 11: 751–775.
- Roger, Derek. 1989. Experimental studies of dyadic turn-taking behaviour. In D. Roger and P. Bull (eds) *Conversation: An interdisciplinary perspective*, 75–95. Philadelphia: Multilingual Matters Ltd.

- Schegloff, Emanuel. 1972. Sequencing in conversational openings. In J. Gumperz and D. Hymes (eds) *Directions in sociolinguistics*, 346–380. New York: Holt, Rinehart & Winston.
- Schegloff, Emanuel and Harvey Sacks. 1973. Opening up closings. *Semiotica* 7(4): 289–327.
- Scollon, Ronald and John Flowerdew. 1997. Public discourse in Hong Kong and the change of sovereignty. In R. Scollon and J. Flowerdew (eds) *Journal of Pragmatics, special issue on 'Language and discourse issues in Hong Kong's change of sovereignty'* 28: 417–426.
- Sinclair, John and Malcolm Coulthard. 1975. *Towards an analysis of discourse*. Oxford: Oxford University Press.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stenström, Anna-Brita. 1994. *An introduction to spoken English*. London: Longman.
- Svartvik, Jan. (ed). 1980. *The London-Lund Corpus of Spoken English: Description and research*. Sweden: Lund University Press.
- Svartvik, Jan. 1996. Corpora are becoming mainstream. In J. Thomas and M. Short (eds) *Using corpora for language research*, 3–13. London: Longman.
- Swales, John. 1990. *Genre analysis*. Cambridge: Cambridge University Press.
- Taylor, Tracy and Deborah Cameron. 1987. *Analysing conversation: Rules and units in the structure of talk*. Oxford: Pergamon Press.
- Tsui, Amy B.M. 1985. Analyzing input and interaction in second language classrooms. *RELC Journal* 16(1): 8–32.
- Tsui, Amy B.M. 1987. An analysis of different types of interaction in ESL classroom discourse. *IRAL* XXV(4): 337–353.
- Tsui, Amy B.M. 1991. Learner involvement and comprehensible input. *RELC Journal* 22(2): 44–60.
- Tsui, Amy B.M. 1994. *English conversation*. Oxford: Oxford University Press.
- Warren, Martin. 1993. Towards a description of the features of naturalness in conversation. Unpublished Ph.D. thesis, University of Birmingham.
- Wilson, John. 1987. On the topic of conversation as a speech event. *Research in language and social interaction* Vol. 21. pp 93–114.
- Wilson, John. 1989. *On the boundaries of conversation*. Oxford: Pergamon Press.