Multi-modal corpus pragmatics: the case of active listenership [*]

Dawn Knight and Svenja Adolphs
School of English Studies, The University of Nottingham, UK

1. Abstract

The difficulties associated with the development of spoken corpora large enough to yield stable analytical results have meant that much of corpus linguistics has focused on the analysis of written discourse. However, alongside the large-scale studies of lexico-grammar on the basis of mainly written corpora, there has been a consistent effort in the exploration of spoken discourse through a corpus linguistic approach (e.g. McCarthy 1998, Leech 2000). Spoken corpora provide a particularly valuable resource for both quantitative and qualitative types of analysis of specific pragmatic functions and thus help generate new empirical insights of patterns of language usage (Aijmer 1996, DeClerck 2004, Tsui 1994). Such studies help us to re-evaluate claims and concepts that originate in more philosophical traditions where the conceptualisation of pragmatic functions has arguably received most attention (Austin 1962, Searle 1976, Searle and Venderveken 1985).

However, one of the key differences between written and spoken corpus analysis, is that the spoken corpus is a mediated record, a textual rendering of an event which is multi-modal in nature, and thus capturing only a limited and limiting aspect of the reality of that event. As a result, analyses of pragmatic functions in spoken corpora tend to exclude the exploration of the interplay between gesture and language and therefore neglect a core element in the construction of meaning in interaction (Kendon 1994, Kress and van Leeuwen 2001).

The recent development of multi-modal corpus resources allows for a fuller representation of aspects of the context and co-text in which human communication takes place, in terms of the types of gestures that accompany verbal signals. Through advances in data collection and representation of this kind, it has now become possible to study pragmatic functions in a multi-modal environment which includes both gesture and language, and to study patterns between the two that arise through the analysis of a corpus of human interaction.

This paper starts to examine how a multi-modal corpus analysis can be employed to facilitate the exploration of pragmatic categories in discourse. Using as an example the category of active listenership in conversation, it aims to focus on the way in which descriptions of pragmatic functions have to be reconsidered in the light of accompanying gestures. The gestures that will be examined are movements of the head that occur when the hearer signals active listenership. Drawing on a multi-modal corpus of an academic supervision session, different existing categorisations of back-channels will be discussed in terms of their validity and applicability to multi-modal data. The paper will end with a brief exploration of the place of multi-modal corpora in pragmatics research.

2. Corpus pragmatics

The systematic exploration of utterance function has long been a key concern in the area of pragmatics. Morris (1946: 287) defines pragmatics as 'the science of the relation of signs to their interpreters' which implies that this discipline is concerned with the relationship between language form and its interpretation by language users. In human communication, the interpretation of language–in-use is partly based on the meaning of the actual language form or 'sign' used, and partly on other sources of knowledge, such as knowledge about the context of the situation, knowledge about other speakers or listeners and knowledge about culturally recognised norms and activities. Pragmatics, then, does not assume a one-to-one relationship between language form and message function. Rather, it tries to account for the reasons and processes behind the phenomenon that certain linguistic forms might be interpreted as carrying a particular function in a particular context.

In doing so, a range of different methods and approaches have been developed which come under the very broad heading of pragmatics. Spencer-Oatey and Zegarac (2002:84) distinguish between the cognitive-psychological approach which is marked by the analysis and discussion of decontextualised utterances and their possible interpretations, and the social-psychological approach. The latter tends 'to focus on the ways in which particular communicative exchanges between individuals are embedded in and constrained by social, cultural and other contextual factors.' (ibid.).

Depending on which perspective is taken, the methods that are being used to explain the relationship between 'signs' and 'users' vary. They range from philosophical explorations of the 'conditions' that have to be met for the achievement of a particular interpretation of a message; as advanced by the speech act theorists in the 1960s to corpus-based explorations of recurrent matches between particular forms and functions (see DeClerck 2004) to the use of discourse completion tasks, detailed fields notes and interviews relating to a particular communicative event and/or function.

Corpus-based research into utterance function is a relatively under-explored area at present. This is despite the fact that a number of key journals in the area of pragmatics now feature corpus-based research, including, for example, a special issue on corpus linguistics in the Journal of Pragmatics in 2004. This is partly due to problems related to assembling and accessing spoken corpora and partly due to the inevitable focus on lexical rather than functional units, concordance lines rather than extended discourse stretches, in corpus research. In addition, the fact that spoken language is multi-modal in nature, and that meaning is created through an interplay of a range of semiotic modalities, contributes further to the unease that often accompanies this particular line of enquiry.

The number of studies which embrace the painstaking annotation of spoken data with information on prosody and gesture, and which align audio and video streams with the transcript is limited. Yet, there are substantial benefits to carrying out a multi-modal corpus-based investigation of language function and this paper aims to make a start at outlining some of the key issues involved in such a study.

In order to illustrate these issues we draw on the function of signalling active listenership, an area which has remained largely under-explored in the area of pragmatics. Signals of active listernership, or back-channels, can be verbal or visual or both and are thus an ideal function to explore in a multi-modal corpus.

3. Multi-modality and spoken corpora

It is widely understood that the generation of meaning in face-to-face discourse does not rely purely upon verbalisations. Instead, 'face-to-face interaction is filled with various verbal and non-verbal signals which interact and sometimes even counteract' (Maynard, 1987: 590). Discourse is therefore 'multi-modal', comprising of both the vocalised characteristics of talk and accompanying meaning, iconic gestures such as hand movements (Beattie & Shovelton, 1999 & Rimé & Schiaratura, 1991 & Thompson & Massaro, 1986), gaze (see Cerrato & Skhiri, 2003 & Griffin & Bock, 2000), body movement, head nods and facial expression.

There is a long history of experimental research into the multimodal nature of discourse, research which is generally designed to be either 'observational' or 'experimental' (Kendon, 1994: 182), and is designed to answer one or both of the following questions (177):

1- If recipients are offered utterances which include gestures, if they are permitted to see these gestures, do they interpret these utterances differently than when they are not permitted to see them? (examples of such studies include Rogers, 1978, Riseborough, 1982)
2- If recipients are asked to make judgements about the gestures of others in the absence of speech to which they were related, do they make such judgements in a consistent way, and if they do, do these judgements show that they have some understanding of the utterance of which they were a part?

The present study, however, adopts a more pragmatic approach to the study of gesture-in-talk, by examining the communicative roles and functions of a specific gestural form: head-nod behaviour. Head nods are a salient form of 'non-verbal communication', an umbrella term that includes 'expressive signs, signals and cues that are used to send and receive messages apart from manual sign language and speech' (Kapoor & Picard, 2001:1). As outlined above we are particularly interested in head nods that function as backchannels in conversation. Backchannels (also known as 'minimal responses' (Fellegy, 1995), 'accompaniment signals' (Kendon, 1967), 'listener responses' (Roger, Bull & Smyth, 1988) and 'assent terms' (Schegloff, 1972) are best understood as 'vocal, verbal and non-verbal non-floor-holding devices that a listener may use to respond to the floor-holding message in a conversation' (O'Keeffe & Adolphs, 2007).

Backchannels exist to provide feedback in communication, which, as Allwood et al. point out (1993: 4) 'enable[s] the participants of a conversation to unobtrusively exchange information about four basic communicative functions: contact, perception, understanding and attitudinal reactions'. Duncan & Neiderehe outline five 'types' of backchannel behaviour, comprising of vocal, verbal and gestural forms: (1974)

1- Readily identified, verbalised signals such as *yeah, right, mmm*
2- Sentence completions
3- Requests for clarification
4- Brief restatements
5- Head Nods and shakes (referred to as 'non-verbal' backchannels in this study)

Whilst verbal forms of backchannels, such as *mmm* and *yeah,* have been widely researched by the linguistic community (see, for example, Gardner, 1998, Oreström, 1983 & O'Keeffe & Adolphs, 2007 for more details), there is limited detailed linguistic research into the various forms and discourse functions of head nod behaviour and its relationship to verbalised elements of backchanneling phenomena.

Dobrogaev (1929, reported in McClave, 2000: 856) conducted one the earliest explorative studies that looked at head nods specifically, questioning whether effective communication is possible when all head movements are suppressed. Kendon developed this study further and was one of the first to propose that a direct link exists between lexical units and head nod behaviour in discourse (1972: 195), an observation that has been the premise of many studies since. Other studies have focused on the physical generation of backchanneling head nods, such as in HCI based studies (Human-Computer Interaction, for example Kapoor & Picard, 2001).

Previous research in this arena has tended to focus either on forms, positions and functions of head nods, with few looking at the relationship between these features and how they contribute to the meaning of discourse. This study aims to explore in more detail the closeness of the verbal and non-verbal relationship highlighted by Kendon (1972), exploring in detail backchanneling head nod behaviour and verbal backchannels in discourse, as evidenced in naturally occurring multi-modal corpora.

4. Data and study design

The study presented in this paper draws on a sub-corpus of a larger multi-modal database which is currently being developed at the University of Nottingham (the NMMC: Nottingham Multi-Modal Corpus, see Adolphs & Carter, forthcoming 2007, Knight et al., 2005 & Knight, 2006). We focus on an in-depth analysis of a ten minute excerpt of a face-to-face, one-to-one conversation (circa 2200 words). The section has been extracted from a forty-five minute video of an MA supervision session involving a male supervisor and female supervisee, both of whom are British nationals. This is the first MA dissertation supervision between the participants who, prior to this meeting, had only met on a few, short occasions. The supervision was randomly selected from five hours of video that have been collected for the NMMC, and the excerpt was extracted from the middle of the session, between the 15.00 and 25.00 minutes.

When recording the data, the participants were sat face-to-face, each with a partially hidden camera facing them, and a high quality microphone between them. This allowed for the recording of close-up images of each individual, which were later synchronised using Adobe Premiere, so that during analysis both the activity of both participants could be observed at the same time.

In order to analyse both language and gesture in a systematic way the following phases of exploration were applied:

- Defining and classifying verbal backchannel behaviour
- Defining and classifying non-verbal backchannel behaviour
- Combining verbal and non-verbal and highlighting the potential for exploring patterns and relationships between the two

When backchannels were detected in both modes of data, in order to make them 'usable' for analysis, they were marked-up and coded, allowing us to map their distribution throughout the discourse, and to group and compare them accordingly (see Knight et al., 2005).

The definition of backchannels that was used as the basis for their identification throughout this study was of 'any short response item that did not appear to take over a speaker turn, and was not a response to a question' (O'Keeffe & Adolphs, 2007). The following example taken from our data illustrates this definition:

> **<$1>** I mean what do you want to do you know you <\$=> obviously you keep all the data you get a richer source of stuff <$=> to <\$=> to draw from. Erm but you've really got to be self disciplined then of ignoring things that might be really interesting+
>
> **<$2>** Yeah.
>
> **<$1>** +if they don't fit the parameters <$=> of <\$=> of what you're looking at. Or you do it the other way and slim the data down to look at just one or two things but do quite a detailed almost qualitative analysis of them.

In this case the example of the use of *yeah* by speaker <$2> is classed as a backchannel as it provides feedback to the speaker without interrupting the flow of the discourse. In some instances the backchannels are less salient and more difficult to define, such as with the following example, which is also taken from the excerpt data:

> **<$1>** But you could do that with the doctor ones.
>
> **<$2>** Yes.
>
> **<$1>** <$=> That's <\$=> That's why I'm wondering whether <$=> the two looking at the two types <\$=> It suggests it's a contrastive study.

In this case the use of *yes* by speaker <$2> is a more engaged mark of agreement, a direct response to the statement that follows. It also appears to interrupt the conversation, as speaker <$1> requires the feedback of affirmation in order to continue. As a result although short, this response is not defined as a backchannel.

In order to minimise the potential problems of definition it was important that all instances were consistently analysed to determine whether, for example, they may impose on a turn, or perhaps answer a question, and so whether or not that they confidently be classified as a verbal backchannel.

Following the identification of all back-channels in the transcript they were classified according to their discourse functions (see Maynard, 1987, Schegloff, 1982, Gardner, 1997 and Tottie, 1991 for accounts of backchannel functions). We used the following framework developed by O'Keeffe and Adolphs (2007).

- CONTINUERS (CON):

    - The most basic form of backchannel which is used to maintain the flow of discourse, and to provide feedback on how the message is being received.
    - Continuers act as floor-yielding tokens signalling that the addressee is listening, desiring the speakers floor holding narrative to continue.
- CONVERGENCE TOKENS (CNV):
    - Convergence Tokens have a 'higher relational value' than Continuers, as they are used to mark agreement / convergence.
    - They are used to help maintain good relations, by reinforcing commonality throughout the discourse.
- ENGAGED RESPONSE TOKENS (ER):
    - These are more affective response tokens, communicating emotive signals and opinions to the speaker without taking over the turn in the discourse.
    - They can highlight, for example, the addressee's anger, shock, surprise, disgust, sympathy, empathy and so on.
- INFORMATION RECEIPT TOKENS (IR):
    - IR response tokens are highly organisational, associated with asymmetrical discourse, where one speaker has control over the flow of discourse.
    - They are rare in casual conversations in familiar settings.
    - They can assume the role of a discourse marker, signalling the close or shift of a topic (so are usually marked by falling pitch).

The mark-up and encoding of backchanneling head nods is not as straightforward as this, since gestures belong to a different semiotic system to that of spoken language, and are not always discrete and freely identifiable units of behaviour in discourse. The development of computer vision techniques to track head movement in multi-modal corpora is currently underway (Knight et al., 2006) but was not used for the current study. Head nods were defined on the basis of the video and grouped in the following way:

Type 1: Small (nonchalant) nods with a short duration
Type 2: Small (nonchalant), multiple nods with a longer duration than type 1
Type 3: Intense nods with a short duration
Type 4: Intense and multiple nods with a longer duration than type 3
Type 5: Multiple nods, comprising of a combination of types 1 and 3, with a longer duration than types 1 and 3.

5. Analysis

Table 1 depicts the frequencies of backchannels in the data according to speaker and whether they co-occur with nods. All of the instances have been grouped according to speaker. There are a total of 68 verbal backchannels in this excerpt, circa 59% of which are used by speaker <$1>, the supervisor. It is important to note that of the 2156 words contained within the extract, 1401 were spoken by speaker <$1>, whilst only 755 were spoken by speaker <$2>, which possibly accounts for the disparity between the frequencies of verbal backchannel behaviour.

	TOTAL	<$1>	<$2>
B-C Total:	68	40	28
B-Cs with Nod:	41	18	23
B-Cs without Nods:	27	22	5

Table 1: *Table showing the breakdown of frequency counts of verbal backchannels in the excerpt*

The most common forms of verbal backchannel behaviour seen here is *yeah*, occurring a total of 29 times (<$1> = 10, <$2> = 19), 7 without nods (<$1> = 6, <$2> = 1), and 22 with nods (<$1> = 4, <$2> = 18). In other words 10% (1 out of 10) of the *yeah's* spoken by <$1> occur with nods, whereas for <$2> it is 95% (18 out of 19), so with only 1 instance where *yeah* is used without a nod. *Yeah* is most frequently used with Convergence Tokens and Continuers. *Yeah* is used as a Convergence Token, and with a nod a total of 11 times, with 2 from <$1> and 9 from <$2> and is used without a nod a total of only 4 times (3 from <$1>, 1 from <$2>). It is used as a Continuer with a nod a total of 11 times (<$1> = 2, <$2> = 9), and without a nod on only 3 (<$1> = 3, <$2> = 0) occasions. So <$2> uses *yeah* a lot more across the data, but more so with nods, whereas <$1> uses it more frequently without nods. This pattern and these results will be discussed further in section 6.

The most common discourse functions of the backchannels in the extract were Continuers and Convergence Tokens. Both accounted for 45 of the 68 instances of verbal backchanneling. There is no real marked difference between the use of these tokens between the speakers, especially when taken in relation to the total amount of occasions in which backchannels are used by respective speakers. There is a more marked difference in the usage of Engaged Response tokens, with <$1> using the token 6 (86%) times and <$1> using it only on 1 (14%) occasion out of the 7 total occurrences.

The use of Information Receipts is even more marked across the speakers. In total 16 Information Receipts were found, and it is interesting to note that of which, 14 (88%) were spoken by <$1>, accounting for 35% of the total occasions on which this speaker backchannels. Figure 2, below, shows the total frequency counts for each backchannel function (occurring with and without nods):

	TOTAL	<$1>	<$2
CON	23	11	12
CNV	22	9	13
IR	16	14	2
ER	7	6	1
	68	**39**	**28**

Table 2: *Table showing frequency counts of functions of verbal backchannels in the excerpt*

Table 3 below shows the frequencies of the various types of backchanneling nods as evidenced by the excerpt, noting where they occur with and without nods. This table shows that there are a total of 113 backchanneling head nods in the excerpt. The data shows a reversal of what was seen with the verbal backchannels, with <$2> nodding

more frequently than <$1>, with a total of 70 nods, amounting for 62% of the total, whilst <$2> only nods a total of 19 times, amounting to 38% of the total.

	TOTAL	<$1>	<$2>
Nod Totals:	113	43	70
Nods without B-Cs (verbal backchannels):	72	24	48
Nods with B-Cs:	41	19	22

Table 3: *Table showing frequency counts of non-verbal backchannels (nods) in the excerpt*

From a closer inspection of the data it was discovered that <$2> nods without backchannels twice as many times as <$1>. In addition to this, nods co-occurring without backchannels were most likely to be of type A or C, as seen in figure 4 below:

	TOTAL	<$1>	<$2
A	32	7	25
B	14	8	6
C	21	6	15
D	1	1	0
E	4	2	2
	72	**24**	**48**

Table 4: *Table depicting the frequencies of nods types- occurring without backchannels*

When we combine our analysis of verbal and non-verbal backchannels we find a number of interesting patterns. Tables 5 and 6 below show the frequency of backchannel functions where they co-occur with head nods and where they are used by spoken language alone.

	TOTAL	<$1>	<$2
CON	13	4	9
CNV	16	4	12
ER	3	3	0
IR	9	7	2
	41	**18**	**23**

Table 5: *Table showing backchannels that occur with head nods- breakdown of functions*

	TOTAL	<$1>	<$2
CON	10	7	3
CNV	6	5	1
ER	4	3	1

IR	7	7	0
	27	**22**	**5**

***Table 6**: Table showing backchannels that occur without head nods- breakdown of functions*

These tables show that backchannels (68) co-occurred more frequently with nods (41) than without nods (27) across the data. Table 7 below shows the different types of head nods that co-occurred with the verbal backchannels in the data:

	TOTAL	**<$1>**	**<$2**
A	12	6	6
B	7	4	3
C	17	6	11
D	2	1	1
E	2	1	1
	41	**18**	**23**

***Table 7:** Table showing the frequency counts of different nod types, co-occurring with listener backchannels*

The type A nods were most likely to co-occur with Information Receipt backchannels. 5 of the 12 (42%) nods of type A were of this nature, consisting of 3 by speaker <$1> and 2 by speaker <$2>. In comparison to this, type B nods were more likely to co-occur with Convergence Tokens. 5 of the 8 (63%) nods of type B were of this nature, with 2 by speaker <$1> and 3 by speaker <$2>. Type D nods were as likely to co-occur with Continuers and Convergence Tokens, and Type E nods were as likely to co-occur with either Continuers or Engaged Response Tokens, although at 2 instances of each, the frequency of these types was relatively small.

6. Discussion

The data on which the analysis above is based forms part of a larger multi-modal corpus of spoken academic discourse. As such it is only meant to serve as an illustration of what we may be able to achieve with a larger scale corpus in terms of analysing the interplay between gesture and language in relation to a particular discourse function.

The analysis above highlights a number of interesting issues which will have to be taken into account if we are to extend a corpus pragmatics approach to include gestures as well as linguistic choices. In the data that we have analysed there seems to be a relationship between the type of head-nod and the discourse function. While this relationship is tenuous at this stage and in need of future validation we see for example that half of the small nods of short duration co-occur with the information receipt function while half of the small nods of longer duration (type B) co-occurred with the function of a convergence token. There also appears to be a distinction between the two speakers in terms of gesture and functions of back-channels that are being used where, for example, all of the type C nods (i.e. short and intense nods)

used by the supervisor are accompanied by a verbal signal which has been classified as carrying either the continuer or convergence function.

The brief analysis above has shown that functions of backchannels as determined through linguistic analysis might have to be modified and re-evaluated when we take an integrated multi-modal approach to discourse function. Rather than using discrete categories such as convergence or continuer function it may be more appropriate to conceive of backchannel functions as a cline that moves from a simple continuer function to an engaged function as one of the possible axes. The intensity and duration of accompanying head-nods adds to the description of the overall function and placement of a backchannel on this cline. The results presented above are preliminary and more data needs to be analysed to see whether there is any stable relationship between head-nods and linguistic signal.

7. Conclusion

Corpus analysis has traditionally focused on discrete items, such as individual words or grammatical categories. Pragmatics, on the other hand, is concerned with language functions which are not discrete in nature and which require different descriptive frameworks. Adding the complexities of gesture and movement to the analysis further complicates matters, but also affords new, systematic ways of describing functions in language. Baldry and Thibault (2006: 181) point out that it is 'critically important [..] that corpus-based approaches to text engage with the level of discourse analysis and discourse-level meaning relations on various scalar levels of textual organisation'. While the integration of scalar levels and discrete categories is likely to cause problems in the development of an integrated framework, it also promises to lead to a much richer description of patterns in social interactions. This should apply to both the pragmatic analysis of traditional spoken and written corpora and the analysis of multi-modal corpora.

Acknowledgement

The research on which this article is based is funded by the UK Economic and Social Research Council (ESRC), e-Social Science Research Node *DReSS* (www.ncess.ac.uk/nodes/digitalrecord), and the ESRC e-Social Science small grants project *HeadTalk* (Grant N$^{o.}$ RES-149-25-1016).

References

Adolphs, S. & Carter, R. (forthcoming 2007). Beyond the word: New challenges in analysing corpora of spoken English. *European Journal of English Studies 11*, 2. (18 pages).

Aijmer, K. (1996). *Conversational Routines in English*. London: Longman.

Allwood, J. & Nivre, J. & Ahlsen, E. (1993). On the semantics and pragmatics of linguistic feedback. *Journal of semantics 9*: 1-26.

Austin, J.L. (1962). *How to Do Things with Words*. Oxford: Clarendon Press.

Baldry, A. & Thibault, P.J. (2006). *Multimodal Transcription and Text Analysis: A multimedia toolkit and coursebook*. London: Equinox.

Beattie, G.W. & Shovelton, H.K. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica 123*, 1/2: 1 – 30.

Cerrato, L., & Skhiri, M. (2003). Analysis and measurement of head movements signalling feedback in face-to-face human dialogues. *Proc. of AVSP 2003, St. Jorioz*, France, pp. 251-256

DeClerck, B. (2004). On the pragmatic functions of *let's* utterances. In Aijmer, K. and Altenberg, B. (eds) Advances in Corpus Linguistics. Amsterdam and New York: Rodopi. 213-236.

Duncan, S. & Niederehe, G. (1974). On signalling that it's your turn to speak. *Journal of experimental social psychology 23*, 283-292.

Fellegy, A.M. (1995). Patterns and functions of minimal response. *American Speech 70*, 186-199.

Gardner, R. (1997). The listener and minimal responses in conversational interaction. *Prospect 12*, 2: 12-30.

Gardner, R. (1998). Between speaking and listening: The vocalisation of understanding. *Applied Linguistics 19*, 2: 204-224.

Griffin, Z. M. & Bock, K. (2000). What the eyes say about speaking. *Psychological Science: A journal of the American Psychological Society 11*, 4: 274-9.

Kapoor, A. and Picard, R. (2001). "A real-time head nod and shake detector", *Proceedings of the 2001 ACM Workshop on Perceptive User Interfaces*, Orlando: http://vismod.media.mit.edu/tech-reports/TR-544.pdf

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psycholoigia 20*, 22-63.

Kendon, A. (1972). Some relationships between body motion and speech. In Seigman, A. & Pope, B. (Eds.), *Studies in Dyadic Communication*. Elmsford, New York: Pergamon Press. pp.177-216.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction 27*, 3: 175-200.

Knight, D. (2006). Case Study: Corpora: The Next Generation. Part of the AHRC funded online *Introduction to Corpus Investigative Techniques*, The University of Birmingham. http://www.humcorp.bham.ac.uk/

Knight, D. & Bayoumi, S. & Mills, S. & Crabtree, A. & Adolphs, S. & Pridmore, T. & Carter, R. (2006). Beyond the Text: Construction and Analysis of Multi-Modal

Linguistic Corpora. Paper delivered at the *2nd annual international eSocial Science conference*, June 2006, University of Manchester.

Knight, D., Carter, R. & Adolphs, S. (2005). *The linguistic coding of backchannels: A proposed methodological approach*. Unpublished NCeSS Node Meeting Report. University of Nottingham, December 2005.

Kress, G. and van Leeuwen, T. (2001). *Multimodal Discourse*. London: Hodder Arnold.

Leech, G. (2000). Same grammar or different grammar? Contrasting approaches to the grammar of spoken English discourse. In S. Sarangi and M Coulthard (eds), *Discourse and social life*. Harlow: Longman. 48-65.

Maynard, S.K. (1987). International functions of a nonverbal sign head movement in Japanese dyadic casual conversation. *Journal of Pragmatics 11*, 589-606.

McCarthy, M (1998). *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.

McClave, E. (2000). "Linguistic functions of head movements in the context of speech", *Journal of Pragmatics, vol. 32 (7),* pp. 855-878.

Morris, C. (1946). *Signs, Language and Behaviour.* Englewood-Cliffs: Prentice Hall.

O'Keeffe, A. & Adolphs, S. (2007). Using a corpus to look at variational pragmatics: Response tokens in British and Irish discourse. in K.P. Schneider and A. Barron, ed., *Variational Pragmatics*. Amsterdam, Netherlands: John Benjamins.

Oreström, B. (1983). *Turn-taking in English Conversation.* Lund, Sweden: LiberFörlag Ltd.

Rimé, B. & Schiaratura, L. (1991). Gesture and speech. In Feldman, R. & Rimé, B. (Eds.). *Fundamentals of nonverbal behaviour*. Cambridge: Cambridge University Press. pp. 239-281.

Riseborough, M.G. (1982). Meaning in movement: An investigation into the interrelationship of physiographic gestures and speech in seven-year-olds. *British Journal of Psychology 73*: 497-503.

Roger, D., Bull, P. & Smyth, S. (1988). The development of a comprehensive system for classifying interruptions. *Journal of Language and Social Psychology 7*: 27-34.

Rogers, W.T. (1978). The contribution of kinesic illustrators towards the comprehension of verbal behavior within utterances. *Human Communication Research 5*: 54-62.

Schegloff, E. (1972). Notes on a conversational practice: formulating place. In Sudnow, D.N. (Ed). *Studies in social interaction*. New York: Free Press.

Schegloff, E. (1982). Discourse as interactional achievement: some uses of “uh huh” and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse, text, and talk*. Washington, DC: Georgetown University Press. 71-93.

Searle, J. (1976). The classification of illocutionary acts. *Language and Society* 5 (1), 1-24.

Searle, J. and Venderveken, D. (1985). *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Spencer-Oatey, H. & Zegarac, V. (2002). Pragmatics. In Schmitt, N., (ed.). *Applied Linguistics*. London: Arnold. pp.74–91.

Thompson, L.A. & Massaro, D.W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42,* 144-168.

Tottie, G. (1991). Conversational style in British and American English: The case of backchannels. In Aijmer, K. & Altenberg, B. (Eds.), *English corpus linguistics*. London: Longman. 254-271.

Tsui, A. (1994). *English Conversation*. Oxford: Oxford University Press.