

# Chat, Chunk and Topic in Casual Conversation

**Emer Gilmartin**  
ADAPT Centre  
Trinity College Dublin  
gilmare@tcd.ie

**Carl Vogel**  
School of Computer Science and Statistics  
Trinity College Dublin  
vogel@tcd.ie

## Abstract

Casual conversation does not follow a highly interactive ‘ping-pong’ structure but rather develops as a sequence of ‘chat’ and ‘chunk’ phases. Chat phases are highly interactive stretches amenable to Conversation Analysis, while chunk phases present as longer monologic stretches where one speaker dominates. Chunk phases can often be described in terms of genre, often manifesting as narrative. We describe an annotation scheme for chat and chunks, the chat and chunk annotation performed on six long (c. 1 hour) multiparty casual conversations, and outline plans to compare these chat and chunk annotations to topic annotations on the same data.

## 1 Introduction

Casual conversation, where participants talk ‘for the sake of talking’ rather than using spoken interaction to perform a practical task such as the purchase of a pizza, is fundamental to human social life. Dialog technology, particularly in the domains of healthcare, education, and entertainment is starting to focus on modelling such conversation in order to create more realistic dialog agents. Successful applications will require knowledge of the characteristics of such talk and annotated data to underpin modelling. Practical, ‘task-based’ or instrumental talk has been well studied and there exist several corpora of such interactions, including the ICSI and AMI meeting corpora, and the MapTask and Lucid Diapix ‘knowledge gap task’ corpora (Janin et al., 2003; McCowan et al., 2005; Anderson et al., 1991; Baker and Hazan, 2011). Casual social or interactional talk is less well studied, and existing corpora of such interactions tend to be collections of relatively short dyadic conversations or first encounters, such as the Cardiff Conversational Database, and the Spontal and Nomco corpora (Aubrey et al., 2013; Edlund et al., 2010; Paggio et al., 2010). However, much casual talk is multiparty and can last for extended periods. Casual talk has been described in terms of multi-turn phases of chat/chunk, and in terms of successive topics. Both descriptions provide a view of casual conversation above the utterance level – a level of granularity useful in understanding casual talk and in designing artificial talk. In this paper, we describe ongoing work on multiparty casual conversation. We briefly overview casual conversation in terms of its form and function, describe the annotation of chat and chunk phases in a dataset of such conversations, and discuss plans to contrast these annotation with annotations of topic made on the same data.

## 2 Casual Conversation - Chat, Chunks, and Topic

Casual conversation occurs whenever humans gather (Malinowski, 1923). Examples include short conversations when people meet, intermittent talk between workers on topics unrelated to the job, or longer dinner table or pub conversations. The duration of such interactions can vary from short ‘bus stop’ conversations to ongoing interactions which lapse and start again over several hours. Such talk is thought to build social bonds and avoid unfriendly or threatening silence, rather than simply to exchange information (Jakobson, 1960; Brown and Yule, 1983; Dunbar, 1998). In task-based or instrumental encounters, participants assume clear pre-defined roles (‘customer-salesperson’, ‘teacher-student’) which can strongly influence the timing and content of their contributions, even if roles vary over the course of an encounter. However, in casual talk, participants have equal speaker rights and can contribute at any time (Cheepen, 1988, p. 90). The form of such talk is also different to that of task-based exchanges –

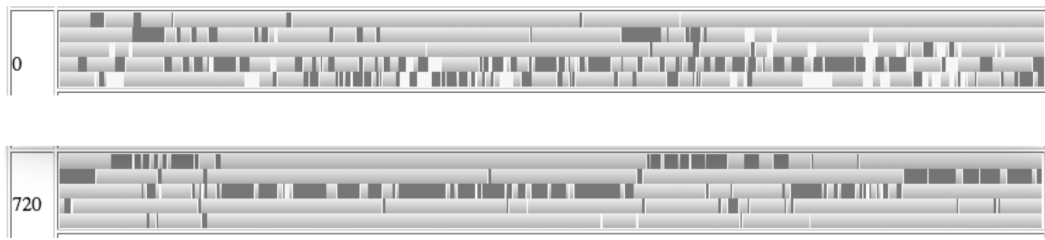


Figure 1: Examples of chat (top) and chunk (bottom) phases in two stretches from a 5-party conversation in the dataset. Each row denotes the activity of one speaker across 120 seconds. Speech is dark grey, and laughter is white on a light grey background (silence). The chat frame, taken at the beginning of the conversation (running from 0 to 120 seconds), can be seen to involve shorter contributions from all participants with frequent laughter. The chunk frame, taken from later in the conversation, running from 720 to 830 seconds, shows longer single speaker stretches.

with less reliance on question-answer sequences and more on commentary, storytelling, and discussion (Thornbury and Slade, 2006; Wilson, 1989). Instrumental and interactional exchanges differ in duration; task-based conversations are bounded by task completion and tend to be short, while casual conversation can go on indefinitely. In terms of structure greater than the utterance level, casual conversation seems to proceed in phases (Laver, 1975; Ventola, 1979), often beginning with ritualised opening greetings, light uncontroversial small talk, and in longer conversations more informative centre phases (consisting of sequential but overlapping topics), and then back to ritualised leave-takings (Ventola, 1979). Slade and Eggins describe sequences of ‘chat’ and ‘chunk’ elements in casual conversation (Eggins and Slade, 2004, p. 230). Chunks are segments where (i) ‘one speaker takes the floor and is allowed to dominate the conversation for an extended period’, and (ii) the chunk appears to move through predictable stages – that is, it is generic. ‘Chat’ segments are highly interactive and managed locally, turn by turn, as several participants contribute utterances with many questions and short comments. Chat frequently occurs at the start of a conversation and is often used to ‘break the ice’ among strangers involved in casual talk, (Laver, 1975). As the conversation progresses, chat phases are interspersed with chunk phases. The ‘ownership’ of chunks seems to pass around the participants in the talk, with chat linking one chunk to the next (Eggins and Slade, 2004). In a study of conversational data collected during workplace coffee breaks, Slade found that around fifty percent of all talk could be classified as chat, while the rest comprised longer form chunks from the following genres: storytelling, observation/comment, opinion, gossip, joke-telling and ridicule. Figure 1 shows examples drawn from our data of typical chat and chunk phases in a 5-party conversation, produced from annotations made using the scheme described in Section 3.1 below.

Topic is often used to segment conversations at a level above the utterance. The notion of topic in discourse has been studied extensively but a concise definition is difficult to find. Topic has been described at sentence level and discourse level (Lambrecht, 1996; Van Dijk, 1977), and as a manifestation of speakers’ intentions (Passonneau and Litman, 1997). For the present work, we consider topic in the sense of the ‘aboutness’ of a stretch of talk or coherent segment of discourse about the same thing – what is being talked about. In terms of topic, task-based interactions often have the contents of the conversation pre-defined. A meeting has an agenda and it is perfectly normal for the chairperson to impose topics for discussion, while the task of buying a pizza imposes what is talked about during the transaction (toppings, price). In casual conversation there is no chairperson and topics are often introduced by means of a statement or comment by a participant which may or may not be taken up by other participants. Thus topic shifts or changes may succeed or fail and cannot be predicted as easily prior to the conversation or mandated by any one participant. The goal of this work is to explore how casual talk can be segmented into stretches of chat and chunk phases or into topics, how easily these phenomena can be annotated, and any correlations between the two.

Corpus	Participants	Gender	Duration (s)
D64	5	2F/3M	4164
DANS	3	1F/2M	4672
DANS	4	1F/3M	4378
DANS	3	2F/1M	3004
TableTalk	4	2F/2M	2072
TableTalk	5	3F/2M	4740

Table 1: Source corpora and details for the conversations used in dataset

### 3 Data and Annotation

We base our work on a set of 6 long (c. 1 hour) multiparty conversations, from the D64, DANS and TableTalk corpora (Oertel et al., 2010; Hennig et al., 2014; Campbell et al., 2006). Table 1 shows details of participant numbers, gender, and conversation duration for each of the six conversations, which were chosen to have a balance of participant numbers.

All of the conversations were recorded in a living room setting or around a table, where participants were instructed to speak or not as the mood took them. As these conversations were face to face and multiparty, automatic segmentation or transcription was not possible, and thus the recordings were hand segmented at the intonational phrase level and transcribed. Details of the dataset and annotation can be found in (Gilmartin et al., 2018b). Below we describe the annotation of chat and chunk phases in some detail and more briefly describe the topic annotation being carried out on the data.

#### 3.1 Annotation of Chat and Chunk Phases

Chat and chunk phases were marked using an annotation scheme devised from the definitions of chat and chunk phases given in (Eggs and Slade, 2004, p. 230), and the descriptions given by Slade of her PhD work on transcripts of chat and chunk phases in casual talk (Slade, 2007). For an initial classification, conversations were divided by first identifying the ‘chunks’ and considering everything else ‘chat’. This was done using the first, structural part of Slade and Eggs’ definition of a chunk as ‘a segment where one speaker takes the floor and is allowed to dominate the conversation for an extended period.’ (Eggs and Slade, 2004). The following guidelines were created to aid in the placing of chat/chunk boundaries.

- Start     A chunk starts when a speaker has established himself as leading the chunk. So, in a story, the speaker will be starting the story, and the story will be seen to go on. This guideline is necessarily vague as Slade and Eggs’ definition does not give a figure for ‘an extended period’
- Stop      To avoid orphaned sections, a chunk is ended at the moment the next element (chunk or chat) starts.
- Aborted   In cases where a story is attempted, but aborted before it is established, this is left as chat. In cases where there is a diversion to another element mid-story, for example, and a return later, all three elements are annotated as though they were single chunks/stretches of chat but the chunks are given the same title with indices to distinguish them. For example, if Speaker A was midway through a story about a football match she’d seen in Barcelona, and a short chat developed around the city of Barcelona, after which Speaker A resumed and completed her story, the segments could be marked FootballStoryA1, BarcelonaChat, FootballStoryB
- Overlap   Sometimes a new chunk begins where a previous chunk is still tailing off. In these cases the new chunk onset is placed where the new chunk starts, and serves as the closure of the old chunk.

Generic labels	
Type	Chunk: x, Chat: o
Ownership	mark with corresponding speaker code from transcript - a, b,... for chat - z
Story	S Can be further marked with codes denoting subtypes: Narrative (N), Anecdote (A), Exemplum (E), Recount (R)
Observation/Comment	C
Opinion	O
Gossip	G

Table 2: Labelling scheme for chunks

Once the chunk was identified, it could be classified by genre. This process was aided by the taxonomy of the stages of various types of story, given by Eggins and Slade in (Eggins and Slade, 2004) For annotation, a set of codes for the various types of chunk and chat was created. Each code is a hyphen-separated string containing at least a Type signifier for chat or chunk, an Ownership label, and optional subelements further classifying the chunks as shown in Table 3.1.

A Conversational Analysis (CA) style transcription for each conversation was generated from the Praat textgrids used in the segmentation and transcription process. These transcriptions had different coloured text for utterances produced by each participant. Annotators found this presentation easy to read, and found the colour coding useful in distinguishing turns. The resulting transcriptions were read by two annotators and onset and offset of chunks roughly marked. The type of chunk was then decided. The temporal boundaries of each chunk were then marked off more exactly in Praat. Intervals were labelled using the code shown in Table 3.1, marking type of phase (chat - o or chunk - x), subtype of phase (narrative, story, discussion...), name of phase (roughly equivalent to the topic under discussion), and phase ‘owner’ (main speaker in chunks and everyone in chat phases). As an example, the code **x.s.g.cats** would denote a chunk phase where the main speaker is **g** and the chunk, which was in story form, was about cats. A total of 213 chat and 358 chunk phases were identified across the six conversations. The topic annotation is ongoing. For the first pass of the topic annotation, two annotators are asked to mark the conversations in terms of ‘aboutness’, and advised to mark both central zones where topics are clear and points are topics changed. This methodology is based on work demonstrating that largely untrained or naive annotators demonstrate high levels of agreement when asked to segment by topic in meetings and conversations (Ries, 2001; Hearst, 1997), and on Schneider’s observations on topic recall in casual conversation (Schneider, 1987). Below we briefly review the results of our analyses of chat and chunk phases in the data, which are more fully described in (Gilmartin et al., 2018a). We discuss how these results may be contrasted with topic data, and our plans for completion of this ongoing work.

#### 4 Analysis of Chat and Chunk Phases

We contrasted the chat and chunk data in terms of phase duration, distribution of laughter and overlap, and how chat and chunk phases were distributed across conversations. Preliminary inspection of chat and chunk duration data (in seconds) showed that the distributions were unimodal but heavily right skewed, with a hard left boundary at 0. Log durations were near normal. The geometric means for duration of chat and chunk phases in the dataset were 28.1 seconds for chat and 34 seconds for chunks. The chat and chunk phase durations (raw and log) are contrasted in the boxplots in Fig 2, where it can be seen that there is considerably more variability in chat durations. Wilcoxon Rank Sum tests showed significant differences in the distributions of the untransformed durations for chat and chunk ( $p < 0.01$ ). A Welch Two Sample t-test showed significant difference in log duration distributions ( $p < 0.01$ ).

The findings on chunk duration are not speaker, gender or conversation specific in our preliminary experiments. Both laughter and overlap were found to be far more prevalent in chat than in chunk phases, reflecting the light and interactive nature of chat phases. We also investigated the frequency and

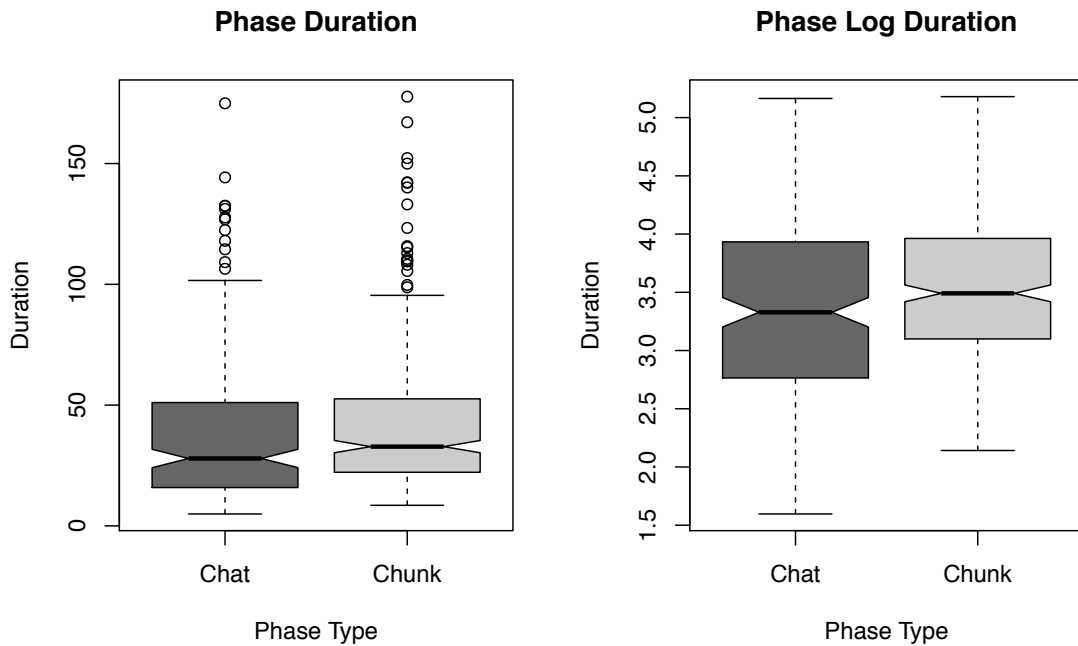


Figure 2: Boxplots of phase duration in Chat vs Chunk in raw and log transformed data

distribution of laughter in chat and chunk phases, finding that laughter accounts for approximately 9.5% of total duration of speech and laughter production in chat phases and 4.9% of total duration of speech and laughter production in chunk phases. Overlapping speech was also found to be more common in chat than in chunk phases. Table 3 shows the occupancy of the conversational floor for all conversations in chat and chunk phases. The number of speakers ranges from 0 (global silence), 1 (single speaker), 2 (2 speakers in overlap) to n-1 where n is the number of speakers in the conversation.

No. Speaking	Chat	Chunk
0	25.75%	22.14%
1	61.58%	72.27%
2	11.88%	5.25%
3+	0.73%	0.42%

Table 3: Floor occupancy in percentage (seconds in brackets) in chat and chunk for all conversations

We observed more chat at conversation beginnings, with chat predominating for the first 8-10 minutes of conversations. As the conversation develops, chunks start to occur much more frequently, and the structure is an alternation of single-speaker chunks interleaved with shorter chat segments. Figure 3 shows the probability of a chunk phase being followed by chat or by chunk as conversation continues (only chunk phases can be followed by chat). There is a greater tendency for the conversation to go directly from chunk to chunk the longer the conversation continues, perhaps reflecting ‘story swapping’.

## 5 Discussion and Work in Progress

The chunk duration distribution may indicate a natural preference of around half a minute for the time one speaker should dominate a conversation. It will be very interesting to compare this with the topic annotations, to see if topic length is comparable. Laughter has been shown to appear more often in social talk than in meeting data, and to happen more around topic endings and topic changes (Gilmartin et al.,

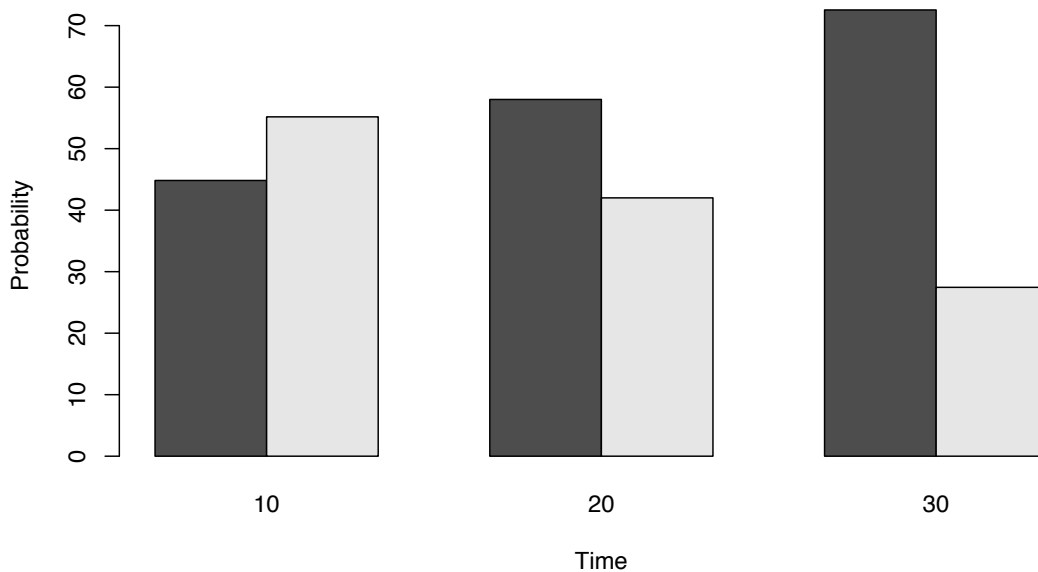


Figure 3: Probability of chunk-chunk transition (black) and chunk-chat transition (grey) in ten-minute bins for first 30 minutes of conversation

2013; Bonin et al., 2012). This is consistent with observations on chat and chunk phases – laughter is more common in chat phases which often provide a ‘buffer’ between single speaker chunks. The structure of chat following chunks is also consistent with Schneider’s description of ‘idling’ phases at the end of topics, where speakers contribute short utterances referring back to previous content rather than committing to starting a new topic. In terms of the distribution of phases across conversations, chat was more common at the start of the multiparty conversations studied, in line with descriptions of casual conversation in the literature, where the initial stages comprise light interactive talk. Chunk phases become more prevalent as conversation continues. These observations lead to the interesting question of whether, once the conversation is well established, the extent of topics may correlate with the stretch from the beginning of one chunk to the beginning of the next chunk, with any intervening chat reflecting the decay of one topic and the process of starting the next. We intend to investigate this question using the topic annotation currently under way on the data.

## 6 Conclusion

We have described ongoing research into the ‘superstructure’ of multiparty casual talk. Our preliminary chat and chunk annotations show that these phases are significantly different in structure, and raise interesting questions on how chat/chunk structure may link to topic distribution. We are hoping to answer these questions in the near future. In addition to the better understanding of conversational structure we hope to gain, this work has practical implications. Automatic annotation of topic without transcripts is not possible, and transcripts of multiparty dialog are often of poor quality. It is often practically and ethically difficult to produce or obtain accurate knowledge of what is being said, but it may be possible to record that speech is happening. As chat and chunk phases may be largely described in terms of the temporal distribution of speech and silence among participants, if there are correlations between chat/chunk phases and topic distribution, such knowledge could be used to better follow the progress of casual talk, and ultimately model such interactions for a number of useful applications.

## Acknowledgements

This work is supported by Science Foundation Ireland (Grant 13/RC/2106) and the ADAPT Centre (www.adaptcentre.ie) at Trinity College, Dublin.

## References

- A.H. Anderson, M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, et al. 1991. The HCRC map task corpus. *Language and speech*, 34(4):351–366.
- A. J. Aubrey, D. Marshall, P. L. Rosin, J. Vandeventer, D. W. Cunningham, and C. Wallraven. 2013. Cardiff Conversation Database (CCDb): A Database of Natural Dyadic Conversations. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, pages 277–282.
- Rachel Baker and Valerie Hazan. 2011. DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior research methods*, 43(3):761–770.
- Francesca Bonin, Nick Campbell, and Carl Vogel. 2012. Laughter and topic changes: Temporal distribution and information flow. In *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on*, pages 53–58. IEEE.
- Gillian Brown and George Yule. 1983. *Teaching the spoken language*, volume 2. Cambridge University Press.
- N. Campbell, T. Sadanobu, M. Imura, N. Iwahashi, S. Noriko, and D. Douchamps. 2006. A multimedia database of meetings and informal interactions for tracking participant involvement and discourse flow. In *Proceedings of LREC 2006*, pages 391–394, Genoa, Italy.
- Christine Cheepen. 1988. *The predictability of informal conversation*. Pinter London.
- R. Dunbar. 1998. *Grooming, gossip, and the evolution of language*. Harvard Univ Press.
- Jens Edlund, Jonas Beskow, Kjell Elenius, Kahl Hellmer, Sofia Strömbergsson, and David House. 2010. Spontal: A Swedish Spontaneous Dialogue Corpus of Audio, Video and Motion Capture. In *Proceedings of LREC 2010*, pages 2992–2995, Malta.
- S. Eggins and D. Slade. 2004. *Analysing casual conversation*. Equinox London.
- Emer Gilmartin, Francesca Bonin, Carl Vogel, and Nick Campbell. 2013. Laughter and Topic Transition in Multiparty Conversation. In *Proceedings of the SIGDIAL 2013 Conference*, pages 304–308, Metz, France. Association for Computational Linguistics.
- Emer Gilmartin, Christian Saam, Carl Vogel, Nick Campbell, and Vincent Wade. 2018a. Just talking - modelling casual conversation. In *Proceedings of the SIGDIAL 2018 Conference*, pages 51–59, Melbourne, Australia. Association for Computational Linguistics.
- Emer Gilmartin, Carl Vogel, and Nick Campbell. 2018b. Chats and Chunks: Annotation and Analysis of Multiparty Long Casual Conversations. In *Proceedings of LREC 2018*, pages 1964–1970.
- Marti A Hearst. 1997. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational linguistics*, 23(1):33–64.
- Shannon Hennig, Ryad Chellali, and Nick Campbell. 2014. The D-ANS corpus: the Dublin-Autonomous Nervous System corpus of biosignal and multimodal recordings of conversational speech. In *Proceedings of LREC 2014*, pages 3438–3443, Reykjavik, Iceland.
- Roman Jakobson. 1960. Closing statement: Linguistics and poetics. *Style in language*, pages 350–377.
- Adam Janin, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, and Andreas Stolcke. 2003. The ICSI meeting corpus. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, volume 1, pages 1–364.
- Knud Lambrecht. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, volume 71. Cambridge University Press.
- John Laver. 1975. Communicative functions of phatic communion. In *Organization of behavior in face-to-face interaction*, pages 215–238. Mouton The Hague.

- B. Malinowski. 1923. The problem of meaning in primitive languages. *Supplementary in the Meaning of Meaning*, pages 1–84.
- Iain McCowan, Jean Carletta, W. Kraaij, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, and V. Karaiskos. 2005. The AMI meeting corpus. In *Proceedings of the 5th International Conference on Methods and Techniques in Behavioral Research*, volume 88, page 100.
- Catharine Oertel, Fred Cummins, Jens Edlund, Petra Wagner, and Nick Campbell. 2010. D64: A corpus of richly recorded conversational interaction. *Journal on Multimodal User Interfaces*, pages 1–10.
- Patrizia Paggio, Jens Allwood, Elisabeth Ahlsen, and Kristiina Jokinen. 2010. The NOMCO multimodal Nordic resource - goals and characteristics. In *Proceedings of LREC 2010*, pages 19–21, Malta.
- Rebecca J. Passonneau and Diane J. Litman. 1997. Discourse segmentation by human and automated means. *Comput. Linguist.*, 23(1):103–139.
- Klaus Ries, 2001. *Segmenting Conversations by Topic, Initiative, and Style*, page 51–66. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.
- Klaus Peter Schneider. 1987. Topic selection in phatic communication. *Multilingua*, 6(3):247–256.
- Diana Slade. 2007. *The texture of casual conversation: A multidimensional interpretation*. Equinox London.
- Scott Thornbury and Diana Slade. 2006. *Conversation: From description to pedagogy*. Cambridge University Press.
- Teun A Van Dijk. 1977. Sentence topic and discourse topic. *Papers in Slavic philology*, 1:49–61.
- Eija Ventola. 1979. The structure of casual conversation in English. *Journal of Pragmatics*, 3(3):267–298.
- John Wilson. 1989. *On the boundaries of conversation*, volume 10. Pergamon Oxford.