TURN-TAKING STRATEGIES IN COOPERATIVE TASK DIALOGUE

Štefan Beňuš

Abstract

We examine turn-taking in collaborative dyadic conversations in which one player described the position of a target object with respect to other fixed objects on her laptop screen, while the other tried to move his representation of the target object to the same position on his own screen. We concentrate on two issues: the role of filled pauses (FPs) such as */um/* or */uh/* in the system of turn-taking, and the strategies for establishing dominance in the dialogues. A quantitative analysis of FP use supports the descriptive observations in the literature that filled pauses mostly function as pre-starts, floor-holders, and to some extent also as floor-yielders. Turn-taking behavior quantified with turn-latencies and the distribution of turn-types also varies with the gender of the interlocutors and the role they perform in the communicative task, and may signal dominance in the conversations.

Key words

filled pause, latency, dominance, gender, turn-taking, task-oriented dialogues

1 Introduction

Turn-taking is a cognitive system that determines who speaks when. It is a prime example of a linguistic system because it has several crucial characteristics. First, taking turns in conversations is governed by rules. It is easily observable that stopping in mid-sentence and expecting interlocutor(s) to jump in, or conversely, jumping in and interrupting somebody before s/he is finished is 'against the rules'. Even though it might be easy to judge whether a turn was initiated appropriately or not, our turn-taking behavior and the rules that underlie it do not lend themselves easily to introspection. For example, it is difficult to say why, in a particular situation, a speaker started speaking exactly 0.2 seconds after the interlocutor finished speaking and waiting more would be awkward while waiting less would be rude.

Secondly, turn-taking behavior is cognitively real. It is acquired unconsciously and without explicit instructions just like many other facets of language. It has been reported that the norms and signals for turn-taking are acquired when children are 2 years old, even before they enter a 2-word stage (e.g. Donahue 1978). Additionally, patients with some forms of aphasia (Wernike aphasia) are able to retain the rules of turn-taking despite the fact that their utterances make

no sense and are not pragmatically cohesive (e.g. Schienberg & Holland 1980) while other patients with brain damage (non-fluent aphasia, Parkinson's disease) suffer from significant decrease in turn-taking competence. Furthermore, turn-taking behavior of children can also serve as an early indicator of dyslexia or special language impairment (SLI). For example, Smith et al. (2008) showed that 3-year old children later identified as having dyslexia were significantly more likely to wait for adults to finish their speaking turn than children who later acquired normal reading proficiency. Hence, turn-taking is closely linked to neuro-cognitive processes localized in various regions of our brains.

Finally, turn-taking behavior is clearly socially-determined. The rules and conventions differ from culture to culture. Moreover, choices that we make in selecting when we speak have clear implications as signals of our social roles, positions of power, or other typical socio-linguistic categories such as age, gender, or race.

From a formal point of view, there are two areas of turn-taking that are pursued in linguistic research. First, a broader goal is to design a testable formal model that describes and explains the observed patterns. A seminal paper attempting to provide a rudimentary model is Sacks et al. (1974). Their model is based on two (optional) rules stating that the conversation continues if the current speaker selects another speaker, and in the absence of this selection, any party of the conversation may self-select.

A more narrow, yet equally important and interesting enterprise, is the investigation of the cues that the participants in conversations may produce and perceive in order to signal the turn-taking organization. These signals can be varied. For example, Duncan (1972) identified signals that yield the turn to other speakers as rising or falling pitch, characteristic drawl as lengthening of speech sounds; body motion such as termination of hand gesture, or relaxation of a tensed hand position, "socio-centric sequences" (fixed expressions such as *or something, you know, but uh*); paralanguage (drops in pitch or loudness); or linguistic cues such as syntactic, semantic, or prosodic completion of a grammatical unit. Even from this early list we see that the modality of turn signals varies and includes visual, paralinguistic as well as linguistic signals. Hence, multiple signals in multiple modalities can be utilized as cues for 'floor-management' when interlocutors may signal that they intend to grab the floor from another speaker, yield the floor to somebody else, or hold the floor since they have something more to say.

In this paper, we concentrate on the potential of filled pauses such as *um* and *uh* (henceforth FPs) for signaling floor-management and turn-organization in dyadic task-oriented conversations in American English.

2 Previous research

Filled pauses have been shown to signal the length of the delay for upcoming speech (e.g. Smith & Clark 1993), and facilitate the perception of upcoming linguistic material (e.g. Fox Tree 2001). For example, *um* was found to signal longer pauses than *uh*, which may be linked to different reasons for hesitations. It was hypothesized that while *um* may signal deeper planning problems, *uh* may be linked to lexical retrieval problems. Also, FPs in general signal that a low-probability transition is going to occur, warning thus the listeners and ultimately facilitating the smoothness of linguistic parsing. Furthermore, FPs signal the strength of preceding intonation boundaries (Swerts 1998), correlate intonationally with surrounding material (Schriberg & Lickley 1993), and affect syntactic parsing (e.g. Ferreira et al. 2004). Moreover, FPs signal discourse structure in monologues (Swerts 1998) and distinguish between given and new referents (e.g. Barr 2001). In psycholinguistic research it was argued that the use of FPs may correlate with anxiety linked to observing one's speech, deception, or use of alcohol (e.g. Christenfeld & Creager 1996, Benus et al. 2006).

A recent review of research related to FPs as well as cross-linguistic comparison of the FP use in English, German, and Dutch can be found in Leeuw (2007). Leeuw divides the functions associated with the use of FPs into symptomatic and signal functions. Symptomatic use of FPs reflects cognitive processes of the speaker such as hesitations, planning difficulties, etc., while the signal use of FPs are deliberate cues aimed at the interlocutor(s) about the intended communicative functions such as floor-holding, or expressing uncertainty. Important novel result, reported in Leeuw, is that several arguments about the FP functions based on American English data such as the fundamental difference between nasal *um* and non-nasal *uh*, might not be straightforwardly extended to other Germanic languages, maybe not even into the British variety of English.

There are also several studies that investigated the relationship between the use of FPs and turn-taking. Sacks et al. (1974) pointed out that FPs belong among 'entry devices' that signal that a speaker is about to say something. They facilitate both production and perception of linguistic material because they allow speakers to think about and plan the intended message and they let listeners get ready to perceive important content. Since FPs are assumed to be lexically empty, if listeners miss the beginning of a turn, no crucial information is lost. In her analysis of the Lund corpus of English, Stenström (1990) found that in addition to assuming the floor in conversations, FPs also mark speakers' intentions to hold the floor in dialogues.

Local and Kelly (1986) argued that FPs differ phonetically based on their floor-management functions. Those FPs that signal floor-holding typically have glottal closure after the FP, and this closure is maintained throughout the following silent pause and released at the onset of the following word from the same speaker. Those FPs that signal floor-yielding have out-breathing at the end of the FP, and typically also contain a more centralized vowel than the floor-holding FPs.

Taboada (2008) investigated the use of FPs in relation to turn-taking in two conversational modes in Spanish: semi-spontaneous interactions similar to real-life dialogues, and non-spontaneous, one-way speech, with mechanical control of turns. She found that FPs were much more frequent in the semi-spontaneous mode, and concluded that FPs are necessary in managing spontaneous conversations, and that they fill mostly the floor-holding function.

Beattie (1982) analyzed televised pre-election interviews with prime minister candidates Margaret Thatcher and Jim Callaghan. He noted that the use of FPs in the interviews may be strongly linked with the public perception of the politicians. For example, despite the fact that Thatcher was interrupted twice as often as she herself interrupted, and Callaghan interrupted more often than he himself was interrupted, general public perceived Thatcher as domineering the interviews while Callaghan was perceived as a 'nice guy' in these interviews. However, Thatcher also used very few filled pauses compared with Callaghan, and thus she failed to signal the interviewer her desire to hold the floor and continue talking. In the absence of these signals, the interviewer butted in and wanted to ask a question, to which Thatcher reacted by increasing pitch and intensity and not letting the interviewer take the floor. Hence, the absence of floor-holding signals such as filled pauses, probably due to language coaching, may have facilitated negative perception of Thatcher by general public.

In addition to Beatie's approach, the relationship between dominance, use of FPs, and turn-taking organization has been investigated in socio-linguistics. One of the popular theories is that males produce more FPs than females (supported by many studies of American English), in order to maintain floor and thus assume a more dominant role in conversations than females (e.g. Coates 1997).

In summary, FPs may signal a wide variety of communicative functions, including the turn-taking organization, and both FP and turn-taking behavior may be strongly connected to establishing and maintaining a dominant role in conversations. In the current paper, we report on the investigation of the relationship between FPs and turn-taking in a large corpus of spontaneous speech of cooperative dialogues. Additionally, we explore the potential of turn-latency and turn-type as indicators of dominance in cooperative tasks.

3 Corpus

The material for this study comes from the Columbia Games Corpus (Gravano et al. 2007, Benus et al. 2007). The corpus consists of twelve dyadic spontaneous task-oriented conversations elicited from speakers of standard American English. Subjects were seated in a soundproof booth divided by a curtain to ensure that all communication was verbal, and played two types of collaborative games (Cards and Objects) using separate laptops. In this paper we analyze the language material elicited during the Objects games only. In these games, one player, henceforth the *Describer*, described the position of a target object with respect to other fixed objects on her screen, while the other player, henceforth the *Placer*, tried to move his representation of the target object to the same position on his own screen. Points were given based on the proximity of the target object to its correct location. The subjects switched roles repeatedly.

In each session, there were 14 changes of the screen, for half of them one subject had a role of Describer, while the other subject described for the remaining screens. There were 13 subjects in total (7 males and 6 females); eleven played with two different partners in two different sessions and two played a single session. All interactions were recorded, digitized, and downsampled to 16K. The recordings were orthographically transcribed, and words were aligned to the source acoustic signal by hand. There is the total of 36,503 words (tokens) and 1,484 unique words (types), which resulted in four hours and 19 minutes of speech.

We asked two labelers to annotate each switch between the speakers following a modified annotation scheme based on Beatie (1982). The scheme is illustrated in Figure 1 below. First, the presence of simultaneous speech between the speakers' turns was determined automatically. Then, the labelers proceeded in the following way. If a turn was judged as not intended for taking the floor, i.e. if the speaker signaled that s/he understands and prompted the other speaker to continue, it was labeled as a Backchannel (BC), or Backchannel with overlap (BC_O). If the turn was judged as intending to take the floor, turns without simultaneous speech were labeled as Smooth switches (S) if the preceding utterance was complete and as Pause Interruptions (PI), if the preceding turn was not complete. Finally, if the speaker intended to take the floor and simultaneous speech was present, the turn was labeled as Butting-in (BI) if the speaker did not succeed in grabbing the floor, and as Overlap (O) or Interruption (I) if the speaker did take the floor. Overlap was labeled if the previous speaker's utterance was complete and Interruption was labeled if it was syntactically, semantically or intonationally incomplete.

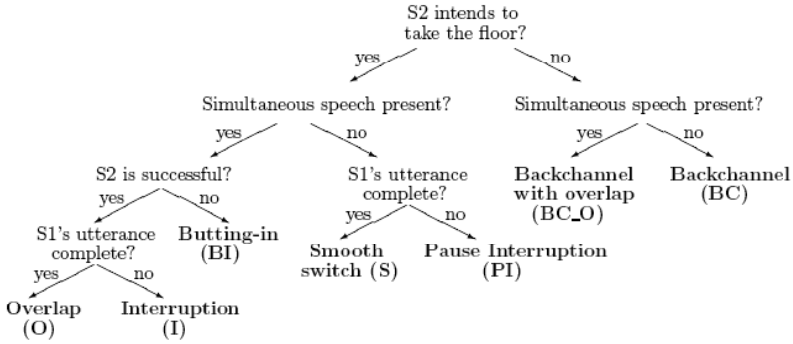

Figure 1: Scheme for labeling turn-exchanges; modified from Beattie (1982)

There were three additional special labels: X1, X2, and X3. X1 is a turn that begins a new task, that is, the first turn after the change on the laptop screens. X2 is a continuation of previous speech by the same speaker after a backchannel (BC, or BC_O) from the other speaker. Finally, X3 marks a simultaneous start. If two turns begin almost simultaneously (formally, within 210 ms of each other, see Fry 1975) then both speakers are most probably reacting to the preceding turn. For example, if turn A starts after 100ms after turn B, A does not respond to B but both A and B react to the preceding turn C.

After the two labelers annotated all the files, the remaining disagreements were discussed and if agreement could not be established, these turns were labeled as "?".

4 Results and discussion

Descriptive observations of the data reveal the following generalizations. The frequency of FP use in this corpus was 2.5 per cent. This is comparable with the FP rate found in the *Switchboard* corpus that contains dialogues between previously unknown persons (Shriberg 2001). However, Shriberg (ibid.) also reported the rate of 3 per cent for FPs in air-travel customer-agent dialogues (AMEX), and Benus et al. (2006) reported FP rate of 4.5 per cent in interviews aimed at determining whether people lie or not. It seems that the nature of the task in the Objects games of the Columbia Games Corpus is very close to natural spontaneous conversations, and more question-answer types of dialogues tend to have higher rates of FP use. Despite the 2.5 per cent FP rate among all words,

14 per cent of all turns had at least one FP. This finding suggests that FP use may provide useful information about turn-taking behavior.

In terms of gender, males used FPs more than females. The frequency of FP use for males was 3.35 per cent and for females 1.78 per cent. Pearson chi-square test showed that this difference was highly significant; $X^2(1, 36515) = 91.44$, $p < 0.0001$. This is mostly due to the higher frequency of *uhs* for males than for females. *Uhs* comprised 54 per cent of all male FPs and 1.8 per cent of all male words while they comprised only 22 per cent of all female FPs and 0.4 per cent of all female words. If these male *uhs* can be analyzed as floor-holders, it would be a support for the idea that FPs serve as one of the means to maintain a dominant role in conversations for male speakers.

	Mm		**uh**		**um**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
turn-final	11	23.4	26	6.8	46	9.8	**83**	**9.2**
turn-initial	28	59.6	67	17.5	124	26.4	**219**	**24.3**
turn-only	10	21.3	9	2.3	17	3.6	**36**	**4**
Total	47	100	383	100	470	100	**900**	**100**
chunk-final	24	51.1	145	37.9	347	73.8	**516**	**57.3**
chunk-initial	44	93.6	191	49.9	308	65.5	**543**	**60.3**
chunk-only	22	46.8	59	15.4	226	48.1	**307**	**34.1**
Total	47	100	383	100	470	100	**900**	**100**

Table 1: Position of filled pauses in turns and chunks

Table 1 illustrates the frequencies of the three types of FP depending on the position within individual turns or chunks. Chunks of speech were determined automatically as pause-defined units within a single turn with the duration of pause at least 50 milliseconds. We see that the peripheral position of FPs within a turn or chunk is dominant. More than one third of FPs starts or ends a turn, and more than half of FPs start or end a chunk. These peripheral positions of FPs suggest several floor-management functions. Turn-initial position is the most common. It suggests that FP initiated successful floor-grabbing or that FP has a pre-start function that allows speaker some time for planning and the listener for tuning in. Turn-final position is the second most frequent. FPs in this position probably signal the interlocutor that s/he may assume the floor. This is because the interlocutor initiated a new turn after these FPs. Turn-only FPs are rare but

point to an interesting function. They suggest unsuccessful floor-grabbing or floor-yielding hesitations that could serve as prompts for more input from the interlocutor.

Positions in chunks are immediately adjacent to silent pauses within a turn. FPs in these positions also have potential to signal floor-management. Chunk-final and chunk-only positions suggest floor-holding function because these FPs are followed by silent pauses during which the other interlocutor did not assume the floor. Chunk-initial position of FP suggest plain hesitation pause.

The table also shows that the position after a silent pause (turn/chunk initial) is dominant, and that the assumed differences between *um* and *uh* are minor. The only difference is in the distribution of *um* and *uh* in chunk-final and chunk-initial where *uh* is more common in chunk-initial while *um* in chunk-final position. However, if we group together the two nasalized FPs *um* and *mm*, this difference disappears.

	FP-initial		**non-FP-initial**	
T-label	**N**	**%**	**N**	**%**
BC	82	**24.48**	311	**8.11**
BC_O	16	**4.78**	106	**2.76**
BI	4	**1.19**	75	**1.96**
I	2	**0.60**	116	**3.02**
O	22	**6.57**	618	**16.11**
PI	4	**1.19**	204	**5.32**
S	139	**41.49**	1520	**39.62**
X1	4	**1.19**	168	**4.38**
X2	47	**14.03**	293	**7.64**
X2_O	2	**0.60**	38	**0.99**
X3	11	**3.28**	360	**9.38**
?	2	**0.60**	27	**0.70**
Total	335	100	3836	100

Table 2: Distribution of the turn types for turns that start with an FPs (on the left) compared with all other turns

Table 2 divides all turns in the corpus into those that start with an FP and the remaining ones, and shows the distribution of the turn types in these two groups. The first observation is that a turn-initial FP is a good predictor for the backchanneling function. We see that almost one third of FP-initial turns function as backchannels while only ten per cent on non-FP-initial turns have this function.

The second observation is that turn-initial FPs in this corpus do not seem to have floor-grabbing function. This is seen in the comparison of the Butt-ins (BI), Overlaps (O), Pause Interruptions (PI), Interruptions (I), and Simultaneous starts (X3). The rates of occurrence for each of these turn types is greater for turns that do not start with an FP than in the FP-initial turns. Hence, the possibility mentioned when discussing Table 1 that turn-initial FPs may have floor-grabbing function is not supported. Rather, the majority of the FP-initial turns are 'non-competitive' because they were labeled as Smooth switches (S, 41%), Backchannels (BC/ BC_O, 29%), and continuations after a backchannel (X2, 14%).

Finally, the third observation is that FP-initial turns tend to signal continuation after a backchannel (X2) more than non-FP-initial turns.

After examining the link between FP use and turn-taking, we explored turn-latency as a measure of conversational dominance in dialogues. Turn-latency is defined as the time between the end of the current speaker turn and the beginning of the new speaker turn; negative values indicate overlap. It is assumed that the lower the latency, the greater the pressure from the new speaker to control the floor. The left panel of Figure 2 shows mean turn-latency for turns spoken by males and females in the role of *Describer* and *Placer*.

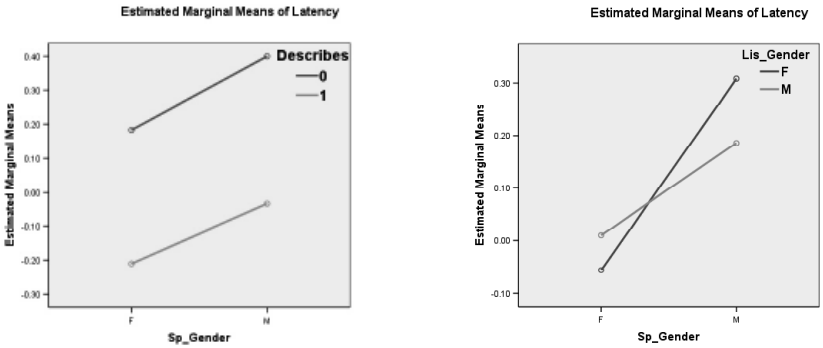

Figure 2: Turn-latency as a function of conversation role and the gender of the interlocutors

We see that *Describers* have shorter latencies than *Placers*; $F(1, 3919) = 61.08$, $p < 0.001$). This finding was expected since *Describers* are assumed to be 'in charge' and control the flow of the conversations. However, two comments to this finding are warranted. First, a general pattern that we observed in our data was that *Describers* controlled the early stages of the conversations since they provided necessary information for completing the task, while *Placers*

tended to be in control in the later stages of the conversation, during which they were eliciting more details about the precise placing of the objects. Second, the average turn-latency shown in Figure 2 pools all turn-types together, including backchannels.

The other observation in the left panel of Figure 2 is that females have shorter latencies than males; $F(1, 3919) = 15.96$, $p < 0.001$. Hence, the latency measure points to a more active and dominant role of females in our data. Finally, there was no significant correlation between the speaker gender and his/her role in conversation on turn latency.

Next we looked at the relationship between the gender of the speaker who starts a turn and the gender of his/her interlocutor, as a function of turn latency. This is illustrated in the right panel of Figure 2. We see that in addition to the main effect of speaker gender mentioned above, the gender of the listener does not significantly affect speaker latency; $F (1, 3919) = 0.256$, $p = 0.62$. However, the gender of interlocutors does play a role in the degree of overlap, as seen in the significant interaction between the two factors; $F (1, 3917) = 6.56$, $p = 0.01$. Hence, the difference between the latencies of males and females is much greater in the presence of a female interlocutor than a male one. Put in a different way, the conversation tends to be more overlapped for female speakers, especially if both the speaker and the interlocutor are female, and the flow is less overlapped if a male talks to a female.

So far, we looked at latency in all turns irrespective of the turn type. Now we explore the realization of dominance and the role in conversations through the turn-type. The left panel in Figure 3 shows the rates of individual turn types for *Describers* and *Placers*. The plot supports already mentioned low rates of interruptions in general. In addition to the expected bias related to backchannels – *Placers* backchannel more while *Describers* continue holding the floor after backchannels – *Describers* also have more overlaps (O), initiate more simultaneous responses (X3), and use smooth switches (S) less compared to *Placers*. Hence, the more dominant role of *Describers* shown with the latency measure can be attributed to the differences in the rates of these three turn types for *Describers* and *Placers*.

The right panel of Figure 3 pools together all interruption turn types (BI, I, PI), and tests their distribution with respect to speaker gender and role in conversation. We see that males tend to interrupt when they are placing objects, while neither females nor males describing objects tend to interrupt. This observation is supported statistically. Pearson chi-square tests showed that male *Placers* interrupt significantly more than male *Describers*; $X^2 (1, 1621) = 14.01$, $p < 0.001$, and that male *Placers* interrupt significantly more than

female *Placers*; $X^2(1, 1962) = 7.3$, $p = 0.007$. Hence, despite the seemingly less dominant turn-taking behavior of males in terms of turn latency, the analysis of turn-type showed that in the *Placer* role males tend to be more dominant and floor-competitive than females.

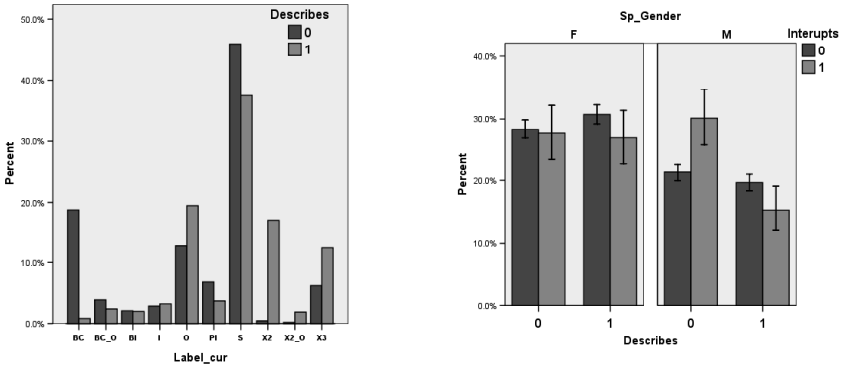

Figure 3: Distribution of turn-types in the speech of Describers and Placers (left). The rates of interruption as a function of speaker role and gender (right)

5 Conclusions and future research

Using the speech data from task-oriented dyadic conversations conducted in American English, we suggested that the features describing the use of filled pauses, latency between the end of the current turn and beginning of the next turn, and the turn type in terms of floor management, provide useful information about the system underlying turn exchanges in conversations. We observed that filled pauses in our corpus tend to start or end stretches of speech – turns or chunks – and thus have a significant potential for turn-management functions. A quantitative analysis of filled pause use supported some descriptive observations in the literature that filled pauses mostly function as pre-starts, floor-holders, and to some extent also as floor-yielders. In our corpus, filled pauses do not seem to function in competitive floor-grabbing

Turn-taking strategies also vary with the gender of the interlocutors and the role they perform in the communicative task, and may signal dominance in the conversations. The investigation of turn-latencies and distribution of turn-types showed that females and the speakers in the role of *Describers* tended to have shorter latencies than males and speakers in the role of *Placers*. Additionally, the difference between male and female latencies was much greater in the presence

of a female interlocutor than a male one. However, male *Placers* also tended to interrupt their interlocutors significantly more than other groups.

In future, we plan to test the potential of automatically extractable acoustic information from turns and chunks for identification and prediction of turn-type. Furthermore, given the observed effects of gender and task role, and thanks to the structure of our corpus, we plan to investigate how the change of the interlocutor affects the observed turn-taking behavior.

Note

This research has been done in part in collaboration with Julia Hirschberg and Agustin Gravano at Columbia University, New York. It was also supported in part by KEGA grant /3/6399/08, and COST 2102.

References

Barr, D. (2001) 'Paralinguistic correlates of discourse structure.' Poster presented at the 43rd annual meeting of the Psychonomic Society, Orlando, USA.

Beattie, G. (1982) 'Turn-taking and interruption in political interviews: Margaret Thatcher and Jim Callaghan compared and contrasted.' *Semiotica 39(1,2),* 93-114.

Benus, S., Gravano, A., Hirschberg, J. (2007) 'The prosody of backchannels in American English.' *Proceedings of 16th International Conference of Phonetic Sciences*. 1065-1068.

Benus, S., Enos, F., Hirschberg, J., Shriberg, E. (2006) 'Pauses and deceptive speech.' In: Hoffmann, R., Mixdorff, H. (eds) *Proceedings of 3rd International Conference on Speech Prosody*. Dresden: TUD Press.

Christenfeld, N., Creager, B. (1996) 'Anxiety, Alcohol, Aphasia, and Ums.' *Journal of Personality and Social Psychology 70(3),* 451-460.

Coates, J. (1997) *Language and Gender: A Reader.* Oxford: Blackwell Publishing.

Donahue, M. L. (1978) 'Form and function in mother-toddler conversational turn-taking', Occasional Paper No. 5, Michigan State University, East Lansing, Michigan.

Duncan, S. (1972) 'Some signals and rules for taking speaking turns in conversations.' *Journal of Personality and Social Psychology 23*, 283-292.

Ferreira, F., Lau, F., Bailey, K. (2004) 'Disfluencies, language comprehension, and Tree Adjoining Grammars.' *Cognitive Science 28(5),* 721-749.

Fox Tree, J. (2001) 'Listeners' uses of um and uh in speech comprehension.' *Memory and Cognition 29*, 320-326.

Fry, D. B. (1975) 'Simple reaction times to speech and non-speech stimuli.' *Cortex 11(4),* 355-60.

Gravano, A., Benus, S., Chavez, H., Hirshberg, J., Wilcox, L. (2007) 'On the role of context in the interpretation of 'okay'.' *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics.* 800–807.

Leeuw, E., de. (2007) 'Hesitation markers in English, German, and Dutch.' *Journal of Germanic Linguistics 19(2),* 85-114.

Local, J., Kelly, J. (1986) 'Projection and "silences": Notes on phonetic and conversational structure.' *Human Studies 9*, 185-204.

Sacks, H., Schegloff, E., Jefferson, G. (1974) 'A simplest systematics for the organization

of turn-taking in conversation.' *Language 50*, 696-735.
Schienberg, S., Holland, A. (1980) 'Conversational turn-taking in Wernike aphasia.' *Clinical Aphasiology Conference 1980*, BRK Publishers, 106-110.
Shriberg, E. (2001) 'To "Errrr" is Human: Ecology and Acoustics of Speech Disfluencies.' *Journal of the International Phonetic Association 31(1),* 153-169.
Shriberg, E., Lickley, R. (1993) 'Intonation of clause-internal filled pauses.' *Phonetica 50*, 172-179.
Smith, A. B., Locke, J. L., Farkas, L. (2008) 'Precursors of dyslexia in early conversational turn-exchange.' *Topics in Language Disorders 28(1),* 5-14.
Smith, V. L., Clark, H. H. (1993) 'On the course of answering questions.' *Journal of Memory and Language 32*, 25-38.
Stenström, A. (1990) 'Pauses in monologue and dialogue.' In: Svartvik, J. (ed.) *London-Lund Corpus of Spoken English: Description and Research.* Lund: Lund University Press. pp. 211-252.
Swerts, M. (1998) 'Filled pauses as markers of discourse structure.' *Journal of Pragmatics 30*, 485-496.
Taboada, M. (2006) 'Spontaneous and non-spontaneous turn-taking.' *Journal of Pragmatics 16(2-3),* 329-360.