Using pause durations to discriminate between lexically ambiguous words and dialog acts in spontaneous speeech

Rajdip Dhillon

Dept. of Linguistics, Yale Univ., 370 Temple St., Rm. 204, New Haven, CT 06520, USA
rajdip.dhillon@yale.edu

Using multiparty conversations conducted in spontaneous Scottish English speech, this study examines whether pause durations can be used to discriminate between lexically ambiguous words pertaining to different dialog acts and, more generally, whether pause durations signal different types of dialog acts. In this study, the dialog acts examined are 'floor grabbers', which are used to gain attention or to interrupt a current speaker who has the floor, and 'acceptance/agreements', which express agreement to or acceptance of a previous speaker's utterance. Experiments 1 and 2 examine pause durations following the lexical items 'right' and 'good' respectively when used as 'floor grabbers' or 'acceptance/agreements'. Experiment 3 examines pause durations following 'floor grabbers' and 'acceptance/agreements', regardless of the lexical item. Results of all three experiments indicate that pauses following 'floor grabbers' are significantly longer than those following 'acceptance/agreements'. This study shows that pause durations are a reliable means of discriminating between lexically ambiguous words and distinguishing dialog acts. The study also sheds light on how stable pause usage is in spontaneous speech.

1   Introduction

The current study brings together two areas of research in the domain of prosody. The first is strictly in the domain of pauses [1, 3, 6, 8, 9, 16, 17, 19, 20, 21] and the second regards dialog acts [2, 5, 15], which describe the pragmatic functions of utterances. Within the context of spontaneous Scottish English speech, the current study examines whether pause durations can be used to discriminate between lexically ambiguous words which pertain to different dialog acts and, more generally, whether pause durations can be used to signal different types of dialog acts.

Pauses have a number of functions in speech. According to Suci [17], pauses are a form of speech fracturing in that they are used to think, plan ahead, and breathe and they represent an interruption in the flow of language. Butterworth [1] attributes three functions to pauses. Pauses represent delays incurred in lexical selection, delays incurred in formulating syntactic plans, and markings of clause endings. In addition, Butterworth suggests that some pauses may be communicatively obligatory as they have the potential to affect meaning by having a definite structural role. Butterworth [1] offers the following example in which pauses are necessary to resolve potential structural ambiguities.

| Old men | and women | (1) |
| Old | men and women | |

| Gentlemen | lift the seat (*instruction*) | (2) |
| Gentlemen lift the seat (*statement*) | | |

With respect to the location of pauses in speech, Suci [17] claims that pauses are much more likely to occur at phrase boundaries than at other points. In a study of pauses in English and French spontaneous speech, Grosjean and Deschamps [8] showed that about seventy percent of all pauses occurred at major constituent breaks. In their study, Grosjean and Deschamps [8] also found that pauses which occurred at major constituent breaks were significantly longer than those which occurred within constituents. In a study examining the length of segments and the placement of pauses, Grosjean, Grosjean, and Lane [9] found that speakers tend to place pauses between segments of equal length. They observed that pauses are more frequent and longer at the ends of sentences than within sentences. More generally, their study demonstrates that the surface structure of a sentence is a good predictor of pause distribution when sentences and constituents are balanced.

While syntactic structure is a good predictor of pause location, it is also of use in predicting pause duration. In a study examining pause durations in synchronous speech, Zvonik and Cummins [20] found that pause duration depends on either or both the length or the syntactic complexity of a preceding phrase. They also observed reduced variability in pause duration when speakers read a text in synchrony. In another study, Zvonik and Cummins [21] investigated whether pause duration is correlated with the length of sentences or phrases preceding and following a pause. Their results indicated that the probability of a pause being short (<300ms) rises greatly if both the preceding and following phrases are short (less than or equal to ten syllables). In another study looking at pause duration, Strangert [16] examined how pause behaviour in Swedish relies on syntactic structure. In Strangert's study, the complexity of NPs and VPs in a sentence and the length of words immediately preceding a boundary were varied and both of these factors were found to influence pause duration. Specifically, Strangert's results showed that pause duration tended to increase when longer words preceded the boundary. With respect to specific syntactic structures, as NP complexity increased, so did pause duration. Pause duration decreased when the NP had the simplest structure and as the VP increased in complexity. While Strangert's results are compelling, they must be interpreted with caution, as only one speaker was used in the study.

When considering how pauses affect the intelligibility of speech, Uchanski et al. [19] conducted a study in which they inserted pauses into conversational speech and deleted pauses from clear speech – which has a slower speech rate than that of conversational speech. The resulting speech was less intelligible in both cases, indicating that there is a relationship between pauses and the timing involved in speech.

As the current study examines pauses in Scottish English, a point worth considering is how the behavior of pauses varies across languages. Campione and Veronis [3] did just this in a study examining pause durations. Specifically, pauses were examined in read speech in English, French, German, Italian, and Spanish and in spontaneous speech in French. What Campione and Veronis discovered was that pause durations do in fact vary across languages – for instance, the average duration of a pause tended to be lower in Italian and higher in Spanish.

In examining other areas of speech in which variation across languages is exhibited, Cruttenden [4] found that the way in which people hesitate in speech may to some extent be language-specific. Cruttenden notes that, in Japanese,

there are very typical hesitation 'morphemes' such as *eto* that are unknown in Indo-European languages. Cruttenden also remarks that even closely related languages may differ in this respect. For instance, speakers of RP English typically have different filled pauses from those speaking Scottish English.

As there is evidence for variation in pause and filled pause behavior across languages, there is also evidence for variation in pause behavior across speakers. Pauses vary depending on many factors such as speech rate, speaking style, and discourse. Fletcher [6] reported that most speakers vary the number of pauses they employ, but some vary pause length to alter speech rate, especially when speeding up. In addition to pauses, speakers also differ in habitual lexical and prosodic patterns [10, 12]. In a study of disfluencies, Shriberg [13] observes a distinction between 'repeaters', speakers who tend to produce more repetitions than deletions or false starts, and 'deleters', speakers who show the opposite pattern. The distinction is such that 'deleters' are faster speakers than 'repeaters', which suggests that the groups may employ different speaking and prosodic strategies.

As the current study examines pause durations in spontaneous speech, issues involved in using spontaneous speech for such purposes should be pursued. A speaker's emotion or state of being has been shown to affect speech patterns [14]. Greenberg and Fosler-Lussier [7] detail various phonetic issues which make using spontaneous speech rather difficult and problematic. For instance, they point to issues which begin at the level of the phone. They note that phones are not very reliable sources of information in spontaneous speech as they are subject to non-canonical phonetic phenomena such as spurious frication, devoicing of normally voiced segments, and acoustic cue trading. Greenberg and Fosler-Lussier further show that there is very little lexical constancy at the level of the phone. In looking at fifty common words, they found that the most popular pronunciation of those words accounted for only ten to twenty-five percent of the variants. In looking at the syllable, Greenberg and Fosler-Lussier observed that the syllable exhibits some measures of stability with respect to pronunciation variation. In their study, about twenty-two percent of phonetic segments associated with a word in terms of its dictionary form were omitted in production and only one percent of syllables failed to be uttered. Syllable onsets were generally realized in canonical form about eighty-five to ninety percent of the time, in contrast to both the syllable nucleus and coda, which were canonically realized in only about sixty to sixty-five percent of the instances. In addition, Greenberg and Fosler-Lussier observe that, as speaking rate increases, so does the probability of non-canonical production.

The current study investigates pause durations in spontaneous Scottish English speech. The central purpose of the study is to determine whether pause durations can be used to discriminate between lexically ambiguous words pertaining to different dialog acts and between dialog acts in general. As has been mentioned, dialog acts describe the pragmatic functions of utterances [5]. Types of dialog acts include basic statements, yes/no questions, commands, suggestions, apologies, backchannels, types of positive responses, types of negative responses, and floor

mechanisms[1]. Detecting dialog acts requires attention to factors such as pragmatic function, syntax, prosody, surrounding context, and interaction between speakers. In a previous study, Bhagat et al. [2] investigated whether automatically extracted prosodic features can serve as cues to dialog acts in naturally-occurring meetings. Specifically, Bhagat et al. examined lexically ambiguous dialog acts occurring in the ICSI Meeting Recorder corpus. The dialog acts they examined were *acceptance/agreements*, *acknowledgements*, *backchannels*, and *floor grabbers* when conveyed by the lexical items 'yeah', 'right', 'okay', and 'uhhuh'. The prosodic features Bhagat et al. investigated in the detection of these dialog acts were duration, energy, pitch, and pauses. Their results indicate that automatically extracted prosodic features do provide cues that distinguish among the four types of lexically ambiguous dialog acts – and the best performance is achieved by combining the different types of prosodic features rather than using them separately.

As Bhagat et al. conducted their study using the ICSI Meeting Recorder corpus, which primarily contains American English and has some amount of British English and English speech from non-native speakers, the current study makes use of spontaneous speech solely in Scottish English. The first experiment examines pause durations following the lexical item 'right' when used as an *acceptance/agreement* or as a *floor grabber*. The second experiment examines pause durations following the lexical item 'good' also when used as an *acceptance/agreement* or as a *floor grabber*. The third experiment takes the findings of the first two experiments to more generally examine pause durations following the dialog acts of *acceptance/agreement* and *floor grabber*. The goal here it to see how reliable pause durations are when discriminating between lexically ambiguous words assigned to different dialog acts and, more generally, to see if pause durations can be used to signal different dialog acts. In addition, the results of this study will shed light on how stable pause usage is in spontaneous speech.

2   Experiment 1: 'Right'

This experiment seeks to determine whether pause durations following the lexical item 'right' depend on whether the lexical item is used as an *acceptance/agreement* or as a *floor grabber*. Dhillon et al. [5] describe *acceptance/agreements* as dialog acts pertaining to utterances which express agreement to or acceptance of a previous speaker's question proposal, or statement. Dhillon et al. describe *floor grabbers* as dialog acts which are typically used by speakers to gain attention or to interrupt the current speaker who has the floor. They tend to occur at the beginning of a speaker's turn and are generally louder than the surrounding speech.

---

[1] Floor mechanisms are types of dialog acts which are employed by a speaker to gain the 'floor' in order to commence speaking (a *floor grabber*), to maintain the floor and continue speaking without relinquishing the floor to another speaker (a *floor holder*), and to 'hold off' from making an utterance when given the floor and expected to speak (a *hold*). For a more detailed examination of dialog acts, see Dhillon et al. [5].

Method

Data. Samples of spontaneous Scottish English speech were obtained from the Scottish Corpus of Texts and Speech.[2] As the dialog acts detailed by Dhillon et al. [5] are constructed in reference to multiparty scenarios and some dialog acts, such as *floor grabbers*, operate more productively within multiparty scenarios rather than scenarios in which only two speakers are present, only audio files containing conversations with three or more speakers were considered. Only cases in which the word 'right' functioned as either a *floor grabber*, as seen in (3), or as an *acceptance/agreement*, as seen in (4), and occurred at the end of an Intonational Phrase (IP) and was immediately followed by an utterance made by the same speaker were extracted.

> Ehm. **Right**. First of all, let's start with how (3) you feel
>
> **Right**. Well, I don't think I ever call them anything except trousers when I was young.
>
> But Colin's **right**. It says a 'young person' (4)
>
> If ye're – Aye. That's **right**. I never thought o that.
>
> That's **right**. Yes.

Because the recordings available from the corpus were all made on a single channel, instances in which the word 'right' was located with the aforementioned conditions met yet were plagued by speech overlap from other speakers were eliminated from the data set. A total of 74 utterances made by 17 speakers (10 male, 7 female) were thus extracted successfully. Of these 74 utterances, 51 were cases in which 'right' was used as an *acceptance/agreement* and 23 were cases in which 'right' was used as a *floor grabber*.

Procedure. In the 74 utterances, pause durations between 'right' and the following word, which belonged to a new utterance, were measured in milliseconds using the Praat software. If a speaker released the final [t] in 'right', the stop burst was also included in the pause duration. The mean of pause durations following 'right' when used as an *acceptance/agreement* was computed and so was the mean of pause durations following 'right' when used as a *floor grabber*. A two-sample t-test was conducted on the individual results to examine whether the pause duration following 'right' differed depending on the dialog act.

In examining the instances in which 'right' appeared as a *floor grabber* or as an *acceptance/agreement*, the dialog act of the following utterance made by the same speaker was examined in order to evaluate whether the following dialog act also played a role in the pause duration following 'right'. Of the 51 cases in which 'right' was used as an *acceptance/agreement*, 34 were followed by *statements*, 16 were followed by another *acceptance/agreement*, and 1 was followed by a *rejection*. Of the 23 cases in which 'right' was used as a *floor grabber*, 13 were followed by *statements*, 7 were followed by *questions*, 1 was followed by an *acknowledgement*, 1 was followed by a *floor holder*, and 1 was followed by another *floor grabber*. Because there were substantial cases in which 'right' used as an *acceptance/agreement* was followed by a *statement* or by another *acceptance/agreement*, the means of the pause

durations occurring in these two types of cases were computed and a two-sample t-test comparing the individual results of both of these cases was conducted. As, to some extent, there were sufficient instances in which 'right' used as a *floor grabber* was followed by a *question* or a *statement*, the means of the pause durations occurring in these two types of cases were computed and a two-sample t-test comparing the individual results of both of these cases was also conducted.

Results

The mean pause duration following 'right' when used as an *acceptance/agreement* was 183ms. When used as a *floor grabber*, the mean pause duration was 345ms. Assuming $\alpha = 0.10$, a two-sample t-test comparing pause durations in all cases of 'right' used as an *acceptance/agreement* to all cases of 'right' used as a *floor grabber* proved to be statistically significant ($t(28) = -1.85$, $p < 0.10$).

In instances in which 'right' was used as an *acceptance/agreement* and was followed by another *acceptance/agreement*, the mean pause duration following 'right' was 198ms. In instances in which 'right' was used as an *acceptance/agreement* and was followed by a *statement*, the mean pause duration following 'right' was 178ms. A two-sample t-test comparing pause durations of 'right' followed by a *statement* to those cases of 'right' followed by another *acceptance/agreement* did not reveal statistically significant results ($t(35) = 0.30$, n.s.).

In instances in which 'right' was used as a *floor grabber* and was followed by a *statement*, the mean pause duration following 'right' was 372ms. In instances in which 'right' was used as a *floor grabber* and was followed by a *question*, the mean pause duration following 'right' was 261ms. A two-sample t-test comparing pause durations of 'right' followed by a *statement* to those cases of 'right' followed by a *question* did not reveal statistically significant results ($t(11) = -0.58$, n.s.).

3 Experiment 2: 'Good'

This experiment seeks to determine whether pause durations following the lexical item 'good' depend on whether the lexical item is used as an *acceptance/agreement* or as a *floor grabber*.

Method

Data. Samples of spontaneous Scottish English speech were obtained from the Scottish Corpus of Texts and Speech. As the dialog acts detailed by Dhillon et al. [5] are constructed in reference to multiparty scenarios and some dialog acts, such as *floor grabbers*, operate more productively within multiparty scenarios rather than scenarios in which only two speakers are present, only audio files containing conversations with three or more speakers were considered. Only cases in which the word 'good' functioned as either a *floor grabber*, as seen in (5), or as an *acceptance/agreement*, as seen in (6), and occurred at the end of an Intonational Phrase (IP) and was immediately followed by an utterance made by the same speaker were extracted.

> **Good**. Ehm so we've done 'hot', 'cold', (5) 'unwell'?

Good. Right. Well gosh that's us gotten round them very, very quickly indeed.

That's **good**. I like that story.   (6)

Good. Yeah. That's a nice one.

That's **good**. Isn't it?

Because the recordings available from the corpus were all made on a single channel, instances in which the word 'right' was located with the aforementioned conditions met yet were plagued by speech overlap from other speakers were eliminated from the data set. A total of 32 utterances made by 3 speakers (1 male, 2 female) were thus extracted successfully. Of these 32 utterances, 19 were cases in which 'good' was used as an *acceptance/agreement* and 13 were cases in which 'good' was used as a *floor grabber.*

Procedure. In the 32 utterances, pause durations between 'good' and the following word, which belonged to a new utterance, were measured in milliseconds using the Praat software. If a speaker released the final [d] in 'good', the stop burst was also included in the pause duration. The mean of pause durations following 'good' when used as an *acceptance/agreement* was computed and so was the mean of pause durations following 'good' when used as a *floor grabber.* A two-sample t-test was conducted on the individual results to examine whether the pause duration following 'good' differed depending on the dialog act.

While in Experiment 1, the dialog act following 'right' was examined to see if it played a role in the pause duration following 'right', the same was not done here in Experiment 2 as the data set was too small to produce reliable results. Specifically, in the 19 cases in which 'good' was used as an *acceptance/agreement*, it was followed by a *question* in 7 instances, by another an *acceptance/agreement* in 6 instances, and by a *statement* in 6 instances. In the 13 cases in which 'good' was used as a *floor grabber*, it was followed by a *statement* in 7 instances, by another *floor grabber* in 3 instances, by a *question* in 2 instances, and by an *acceptance/agreement* in 1 instance.

Results

The mean pause duration following 'good' when used as an *acceptance/agreement* was 168ms. When used as a *floor grabber*, the mean pause duration was 389ms. A two-sample t-test comparing pause durations in all cases of 'good' used as an *acceptance/agreement* to all cases of 'good' used as a *floor grabber* proved to be statistically significant (t(22) = -3.35, p < 0.005).

4   Experiment 3: *floor grabber* vs. *acceptance/agreement*

This experiment seeks to determine whether pause durations depend on the type of dialog act they follow. In this case, the goal is to verify whether pauses following *floor grabbers* differ in duration from those which follow *acceptance/agreements*.

Method

Data. Data from Experiments 1 and 2 was used in this experiment. A total of 106 utterances made by 18 speakers (10 male, 8 female) were used. Of these 106 utterances, 70 were *acceptance/agreements* (51 containing 'right', 19 containing 'good') and 36 were *floor grabbers* (23 containing 'right', 13 containing 'good').

Procedure. The mean of pause durations following *acceptance/agreements* was computed and so was the mean of pause durations following *floor grabbers.* A two-sample t-test was conducted on the individual results to examine whether the pause durations differed if the dialog act was a *floor grabber* or an *acceptance/agreement.*

Results

The mean pause duration following *acceptance/agreements* was 179ms. The mean pause duration following *floor grabbers* was 361ms. A two-sample t-test comparing pause durations of *acceptance/agreement* to *floor grabbers* proved to be statistically significant (t(49) = -3.00, p < 0.005).

5   Discussion and conclusion

The primary goal of the current study is to see how reliable pause durations are when discriminating between lexically ambiguous words pertaining to different dialog acts and to see if pause durations can be used to signal different types of dialog acts. The results of Experiment 1 showed that, in terms of the lexical item 'right', the pause following 'right' was significantly longer when 'right' was a *floor grabber* than when it was an *acceptance/agreement.* However, it was not the case that the dialog act following 'right' affected this pause duration. With respect to the lexical item 'good', in Experiment 2, it was determined that the pause following 'good' was significantly longer when 'good' was a *floor grabber* than when it was an *acceptance/agreement.* In making a general comparison of *floor grabbers* to *acceptance/agreements*, Experiment 3 demonstrated that *floor grabbers* are followed by significantly longer pauses than *acceptance/agreements*. Taken together, the results of this study suggest that pause durations can in fact be used in the discrimination of lexically ambiguous words belonging to different dialog acts and, more generally, that pause durations of particular lengths are capable of signaling different dialog acts.

The nature of *floor grabbers* is such that, upon being uttered, they require the speaker to assess whether he has in fact gained the 'floor' and, in addition, they do not relate to the content, if any, which follows. *Acceptance/agreements*, on the other hand, do not necessarily require such feedback form other participants within a conversation and can relate in content to what follows. Because *floor grabbers* require the speaker to take time to assess the success of the dialog act and potentially to plan for the content which follows, it is thus natural for pauses following *floor grabbers* to be longer than those following *acceptance/agreements.*

As Suci [17] showed that pauses occur at phrase boundaries, the results presented here further confirm Suci's findings, as dialog acts are aligned with phrasal boundaries. In terms of the functions of pauses, Butterworth [1] remarks that pauses represent delays incurred in lexical selection, delays incurred in formulating syntactic plans, and markings of clause endings. The current study adds to this body of work by demonstrating that pauses also represent delays incurred in pragmatic selection. Similarly, as Uchanski et al. [19] show that listeners more or less tend to expect pauses to occur or not

to occur in particular locations, it may be the case that listeners tend to expect pauses of certain durations to occur with particular dialog acts. With respect to duration, Zvonik and Cummins' [20] findings indicated that syntactic structure is of use in predicting pause duration, the results of current study reveal that dialog acts may be of use in predicting pause duration also.

While the findings of the current study do show that pause durations are a reliable means of discriminating between lexically ambiguous words and distinguishing dialog acts, this is not to say that pause durations are the *only* means of doing so. The Bhagat et al. [2] study showed that prosodic features such as duration, energy, pitch, and pauses are highly successful in disambiguating dialog acts when used in a combinatorial manner. What the current study does is show that pause duration is a *reliable* feature in the task of disambiguating words and dialog acts. The current study essentially successfully isolates pauses as being a discriminatory feature in such disambiguation tasks.

More broadly, to expand the work of Greenberg and Fosler-Lussier [7], the current study manages to illustrate that pause durations are used with a good amount of stability in spontaneous speech. Simply put, pause durations are a reliable source of information in spontaneous speech.

While Campione and Veronis [3] examined how pause durations vary cross-linguistically, the current study examines how pause durations behave in Scottish English. For future lines of research, it would be worth examining how pause durations vary across languages and other varieties of English, as examined in the context of dialog acts as done in the current study. In addition, while a lofty and intensive task, it would be worth while to examine pause durations in the full repertoire of dialog acts in a collection of languages to examine whether there is some sort of stability in relative pause durations for a given dialog act cross-linguistically or to see if languages simply vary in this respect.

References

[1] B. Butterworth, "Evidence from pauses in speech," in B. Butterworth (Ed.), *Speech and Talk*, 155-176. London: Academic Press (1980)

[2] S. Bhagat, H. Carvey, E. Shriberg, "Automatically generated prosodic cues to lexically ambiguous dialog acts in multiparty meetings," *Proceedings of the 15th ICPhS,* 2961-2964 (2003)

[3] E. Campione, J. Veronis, "A large-scale multilingual study of silent pause duration," *Proceedings of the Speech Prosody 2002 Conference,* 199-202 (2002)

[4] A. Cruttenden, *Intonation*. Cambridge: Cambridge University Press (1986)

[5] R. Dhillon, S. Bhagat, H. Carvey, E. Shriberg, "Meeting Recorder Project: dialog act labeling guide," *ICSI Technical Report TR-04-002* (2003)

[6] J. Fletcher, "Some micro and macro effects of temp change on timing in French," *Linguistics* 25, 951-967 (1987)

[7] S. Greenberg, E. Fosler-Lussier, "The uninvited guest: information's role in guiding the production of spontaneous speech," *Proceedings of the CREST Workshop on Models of Speech Production: Motor Planning and Articulatory Modeling*, Germany (2000)

[8] F. Grosjean, A. Deschamps, "Analyse dontrastive des variables temporelles de l'anglais et du francais: vitesse de parole et variables composantes, phenomenes d'hesitation," *Phonetica* 31, 144-184 (1975)

[9] F. Grosjean, L. Grosjean, H. Lane, "The patterns of silence: performance structures in sentence production," *Cognitive Psychology* 11, 58-81 (1979)

[10] S. S. Kajerekar, L. Ferrer, E. Shriberg, K. Sonmez, A. Stolcke, A. Venkataraman, J. Zheng, "SRI's 2004 NIST speaker recognition evaluation system," *Proceedings IEEE ICASSP 2005,* Philadelphia (2005)

[11] J. G. Martin, "On judging pauses in spontaneous speech," *Journal of Verbal Learning and Verbal Behavior* 9, 75-78 (1970)

[12] D. Reynolds, W. Andrews, J. Campbell, J. Navratil, B. Peskin, A. Adami, J. Qin, D. Klusacek, J. Abramson, R. Mihaescu, J. Godfrey, D. Jones, X. Bing, "The SuperSID project: exploiting high-level information for high-accuracy speaker recognition," *Proceedings of IEEE ICASSP 2003,* Hong Kong (2003)

[13] E. Shriberg, *Preliminaries to a theory of speech disfluencies*, Ph.D. thesis, University of California, Berkeley (1994)

[14] E. Shriberg, "Spontaneous speech: how people really talk and why engineers should care," *Proceedings of Eurospeech 2005*, 1781-1784 (2005)

[15] E. Shriberg, R. Bates, A. Stolcke, P. Taylor, D. Jurafsky, K. Ries, N. Coccaro, R. Martin, M. Meteer, C. Van Ess-Dykema, "Can prosody aid the automatic classification of dialog acts in conversational speech?" *Language and Speech* 3, 439-487 (1998)

[16] E. Strangert, "Relating prosody to syntax: boundary signaling in Swedish," *Proceedings off the 5th European Conference on Speech Communication and Technology*, 239-242 (1997)

[17] G. Suci, "The validity of pause as an index of units in language," *Journal of Verbal Learning and Verbal Behavior* 6, 26-32 (1967)

[18] M. Swerts, A. Wichmann, R-J Beun, "Filled pauses as markers of discourse structure," *Proceedings of ICSLP 1996*, 1033-1036 (1996)

[19] R. M. Uchanski, S. Choi, L. D. Braida, C. M. Reed, N. I. Durlach, "Speaking clearly for the hard of hearing IV: further studies of the role of speaking rate," *Journal of Speech and Hearing* 39, 494-509 (1996)

[20] E. Zvonik, F. Cummins, "Pause duration and variability in read texts," *Proceedings of ICSLP 2002*, 1109-1112 (2002)

[21] E. Zvonik, F. Cummins, "The effect of surrounding phrase lengths on pause duration," *Proceedings of Eurospeech 2003*, 777-780 (2003)