SPEECH REPAIR CLUES USED FOR DETECTING SPEECH REPAIRS IN *MICASE*

Linda Fraňová

Abstract

This paper analyses speech repair clues in spontaneous speech in the MICASE corpus. An algorithm for detecting speech repairs without using prosodic information or a syntactic parser is proposed. Implementation of this algorithm into spontaneous speech is presented. Two types of speech repairs were analysed: modification and abridged repairs. The analysis shows the frequency of use of the individual speech repair clues such as editing term, word fragment and word correspondence. Modification speech repairs are shown to be the most common speech repair used in the analysed transcripts. Out of three presented speech repair clues the word correspondence is the most prevailing one used by speakers. A test for statistical significance is used to show the significance level in using speech repair clues by male and female speakers.

Key words

speech repairs, discourse markers, reparandum, individual speech repair clues, alteration, modification repairs, abridged repairs

1 Introduction

Schiffrin (1988) claims that some discourse devices such as *oh, well, I mean, I see* fulfil information management tasks in spoken discourse and they are also used to mark a speech repair (ibid.: 81). Therefore, it is assumed that discourse markers are very helpful in detecting speech repairs. To start with, one may ask a question why study disfluencies and why so many linguists have focused their work on spontaneous speech, and disfluencies and speech errors analysis. Firstly, they constitute a problem for parsing theories, which handle only well-formed utterances (Fox Tree 1995). Secondly, showing how speech planning can be broken down, departures from fluent and grammatical speech provide useful data about the architecture of the speech production system (Levelt 1989, MacKay 1970, 1972). Thirdly, disfluencies can display metalinguistic information to listeners about planning difficulties of the speaker (Brennan & Schober 2001), or they can serve as devices for coordinating conversational interaction (Shriberg 1996). Finally, they pose problems for speech recognition systems (Hindle 1983, Nakatani & Hirschberg 1994, Shriberg 1996). Disfluent speech poses also a problem for listeners because they must consider the utterance as if disfluency in the utterance would not occur (Levelt 1989). Consider the following sentence:

Discourse and Interaction 7/2/2014, pp. 5-19
ISSN 1802-9930
DOI: 10.5817/DI2014-2-5

Example 1:

This is the food I want to ***eat-*** *I mean,* ***make****.*

In order to provide a fluent version, the problem with the utterance must be recognized. After recognition of the problem, the listener or hearer must determine what the problem is and how it should be corrected. Sometimes, it is enough to go back only one word and to replace it, and sometimes the replacement of more words is required in order to correct the information. Many linguists have discussed several different strategies for detecting and correcting speech repairs (e.g. Levelt 1989, Hindle 1983, Bear et al. 1992, Nakatani & Hirschberg 1994, Heeman & Allen 1994, Stolcke & Shriberg 1996). By analysing disfluencies in task-oriented conversations Oviatt (1995) found that long utterances have higher disfluency rates than short ones. This finding is supported by Shriberg's (1996) study of disfluencies in three different task-oriented conversational corpora, in which she found that the longer the sentence the less likely it is to be fluent. The aim of this work is to analyse speech repair clues including discourse markers and to show their significance in marking speech repairs in *The Michigan Corpus of Academic Spoken English* (MICASE) (Simpson, Briggs, Ovens & Swales 2002). MICASE consists of 1.8 million words of spoken English used in academic settings. All recordings have been done at the University of Michigan in Ann Arbor. The corpus contains data from a wide range of speech events such as small and large lectures, discussion sections, student presentations, seminars, undergraduate lab sessions, lab group, office hours, advising consultations, dissertation defences, study groups, interviews, etc. There are several features of MICASE that make it very suitable for the current research. Each speech event is categorized according to various contextual attributes. The speaker attributes include gender, age, academic position/role, native speaker status and first language. For the present research two transcripts, namely Intro Biology and Intro to American Politics, have been chosen based on interactivity rating and word count frequency of speech repair indicators. This paper proposes an algorithm which can detect speech repairs by identifying word fragments, editing terms, and word correspondences without using the higher levels of syntactic or semantic knowledge. Such a method can be used to detect speech repairs automatically.

2 Speech repairs

All of us produce occasional interruptions in our speech. When we speak we communicate new thoughts and new feelings about which we have never talked before. Therefore, it is very common that the speaker produces unwanted

utterances which they must repair immediately in order not to confuse the hearer. Fortunately, speech repairs tend to have standard forms, so the listener can identify them easily. Following Shriberg (1994), Nakatani and Hirschberg (1994) divide speech repairs into four intervals: the reparandum, the interruption point (IP), the editing term and the alteration. Let us consider the following example:

Example 2:

- Reparandum is the stretch of speech which is being removed – the wrong part of the utterance. It might end in the middle of the word, resulting in a word fragment.
- The end of reparandum is called the interruption point. This is the point where the disfluency is realized.
- The editing term is a word or phrase with a predictable meaning. It is used to fill the pause and to provide time to plan what will be said. Editing terms include discourse markers such as *I mean, well, let's see, uh,* etc.
- The last part called the alteration is the replacement for the reparandum.

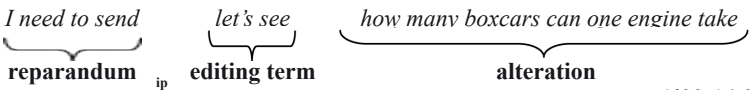
For the hearer to understand the intended utterance, he/she must detect the repair and determine the extent of the reparandum and the editing term.

2.1 Types of speech repairs

Following Hindle's (1983) division, Heeman and Allen (1999: 529) divided repairs into the following categories: fresh starts, modification repairs, and abridged repairs.

Fresh starts occur when the speaker abandons what he/she has just said and starts again. In this case there is no relation between *reparandum* and *alteration*. Consider the following example from Train corpus analysed by Heeman and Allen:

Example 3:



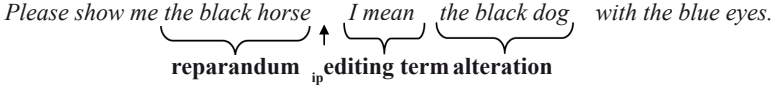
(d93-14.3 utt2)

Example 3 illustrates a fresh start where the speaker abandons the partial utterance *I need to send* and replaces it by *how many boxcars can one engine take?*

When analysing speech repairs a decision whether to include fresh starts into the analysis or not has to be made. Some authors include this type of speech repair in their research (Levelt 1983, Heeman & Allen 1994) and some do not (Johnson 1997). The difference from other types of speech repairs is the need to keep some part of the removed text and, therefore, they require a different method of analysis. As a result fresh starts are not considered in this research mainly because of their different nature.

Modification repairs also known as repairs with non-empty reparandum have strong word correspondences between the *reparandum* and *alteration*.

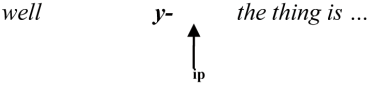
Example 4:



As Heeman and Allen (1999: 530) state, "modification repairs can in fact consist solely of the reparandum being repeated by the alteration". Moreover, the in-word repair when the speaker interrupts within a word and corrects a part of it will be considered as modification repair in this research.

Abridged repairs consist only of a word fragment. Hesitations (filled pauses) will not be characterized as speech repairs. Consider the following example taken from MICASE corpus from the speech event Office hours:

Example 5: (OFC150MU042)



To decide if filled pauses and cue phrases should be treated as signals of abridged repairs can only be done by taking into account if they are in the middle of an utterance or not. Moreover, words that can be used as editing terms can be ambiguous as to whether they are being used in a repair or not. Heeman and Allan (1994) regard as editing terms also filled pauses and hesitation sounds. Nevertheless, in our research we have considered abridged repairs as *possible-repairs*, since there is no direct evidence to suggest that a repair has been made (Levelt 1983). To make this research more objective only abridged repairs consisting of word fragment will be taken into account.

3 Discourse markers

Brinton (1996) defines discourse makers as short words or phrases such as *well, so, oh, you know* which occur with high frequency. In general, the language is always communicative and discourse markers enable spoken utterances to be clear and comprehensive. By implementing discourse items into a spoken language, a speaker connects the ideas together and signals how an upcoming spoken utterance relates to the current unit of speech (Aijmer 1996). These connective speech particles are a valuable source of information for understanding the discourse segment that they introduce. They might introduce transition, addition, contrast, result, consequence, enumeration, apology, repair, clarification, etc. The work deals with distribution of all discourse markers indicating errors in communication. Such markers are *I mean, you know, okey, well, oh, I see, etc.*

4 Repair indicators

The current work focuses only on lexical clues, and thus prosodic information is not taken into account. Here we show results on detection and correction of repairs by combining pattern matching with syntactic and semantic analysis. First, a particular speech repair needs to be identified using several identification cues. One of the most common indication cues is the presence of syntactic anomaly at the interruption point (Bear et al. 1992). Speech repairs are also usually accompanied by word correspondences between reparandum and alteration.

In summary, for this research the following clues will be used to predict speech repairs:

– editing terms;
– word fragments;
– word matching;
– syntactic anomalies/syntactic and semantic knowledge.

Even if a repair is indicated by the above clues, the extent of the repair still needs to be detected. To accomplish this and to make the research more objective ten rules introduced by Heeman and Allen (1994) have been incorporated in our list of thirteen rules, as shown below:

1. editing terms must be adjacent;
2. editing terms must immediately follow the interruption point;
3. a fragment, if present, must immediately precede the interruption point;
4. word correspondences must straddle the interruption point and cannot be marked on a word labelled as an editing term or fragment;

5. word correspondences must be cross-serial; a word correspondence cannot be embedded inside another correspondence (e.g. *how would that – how long that would take*);
6. if there are no other word correspondences, there can only be five intervening words, excluding fragments and editing terms between the first part and the second part of the correspondence;
7. in the removed text, two adjacent matches can have at most four intervening words;
8. in the resumed text, two adjacent matches can have at most four intervening words;
9. for two adjacent matches, the number of intervening words in the removed text can be at most one more that the number of intervening words in the resumed text;
10. a word replacement (except those added by the detection clues) must either only have fragments and editing terms between the two words that it marks, or there must be a word correspondence in which there are no intervening words in either the removed text or the resumed text;
11. pauses might co-occur with the interruption point of speech repairs as well as at the end of the editing term;
12. word fragment must not be at the end of an utterance;
13. if a word fragment includes only corresponding letters with the next word, then it is considered as word repetition (e.g. *wha- what*). If the first letter is corresponding with the first one in the next word and the other not, then the repair is marked as word replacement (e.g. *ste- spectrum*).

The last three rules mentioned above were introduced based on our analysis. Fragments at the end of utterance are not considered as repairs since they are usually considered as not a signal of a repair but rather unfinished word or sentence or a place where a speaker is interrupted by another speaker who takes over the turn. Moreover, the word correspondence might be a word repetition or a word replacement, but with the same part-of-speech tag. So far, mostly indicators of modification repairs were discussed.

4.1 Annotating speech repairs

It can be said that the main aspect of annotation is to mark the repair and its extent. Following Bear et al. (1992) and Heeman and Allen (1994), the following labels have been used: *m* for word matching, *r* for word replacements, other words in reparandum and alteration are annotated ***x***, the editing terms are labelled with *et* and the interruption point with *ip*.

Word matching

Most of the repairs include word repetitions. The letter *m* expresses the notion that the words on one side are exactly same as those on the other side as in the following example from the Dissertation defence speech event category.

Example 6: (DEF500SF016):

now to test this claim i	*i compared Koreans*
m	m

Word replacement

To indicate that the second word was intended by the speaker to replace the first we use *r*. However, the words must be of the same word class.

Example 7:

you	*can*	*see*	*the,*	*they*	*can*	*see*	*the*
r_1	m_1	m_2	m_3	r_1	m_1	m_2	m_3

Word deletion or insertion

Words which are not marked by *m* or *r* are marked by ***x*** if they occur between the repair site and a word marked *m* or *r*. Consider the following example from the Dissertation defence speech event category

Example 8: (DEF305MX131):

environments to be, similar,	*of*	*similar*	*morphologic*
m_1		x	m_1

Word fragments

Word fragments are frequently found as a repair marker. The fragments are marked in bold. If the labeller thinks that the word is meant to replace the fragment with another word, the label *r* should be used. However, if the labeller thinks that the word being replaced is the same, the label *m* should be used. This situation is also explained in Rule 13 in the previous section. Consider the following two examples from the Dissertation defence speech event category:

Example 9: (DEF305MX131):

on ***thos-***	*those plants*
m_1	m_1

Example 10: (DEF305MX131):

much ***overpredic-***	*overestimated*
r_1	r_1

5 Analysis of speech repairs in discussion sections

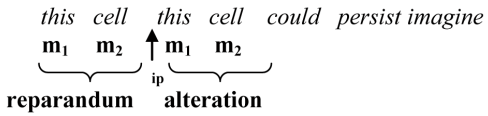
The Michigan Corpus of Academic Spoken English (MICASE) consists of 1.8 million words of spoken English used in academic settings. The analysed speech event category includes two selected transcripts, namely: Intro Biology and Intro to American Politics. Intro Biology is a highly interactive transcript with 6,899 words including 22 native speakers. All examples from this transcript are marked as DIS175JU081.

Types of repairs								
Total			Identification clues					
			with fragment		with editing term		word correspondence	
Modification repairs	60	100%	18	30%	2	3.3%	40	66.7%
Abridged repairs	11	100%	11	100%	0	0%	0	0%
Total	71	100%	29	40.8%	2	2.8%	40	56.4%

Table 1: Types of speech repairs in Intro Biology Discussion Section

According to the above table, the most common type of speech repair used in the analysed transcript are modification repairs, which are mostly signalled by word correspondence, as in the following example:

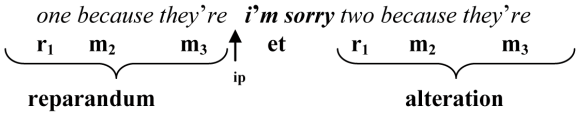
Example 11: (DIS175JU081):



The above example shows a strong word correspondence between words *this cell* in reparandum and *this cell* in alteration. This word matching is marked by letter *m* plus a numerical index. The word marked with m_1 corresponds with word also marked with m_1 and the word labelled m_2 corresponds with the word also labelled with m_2.

The analysis revealed that only two cases of speech repairs were marked by editing term.

Example 12: (DIS175JU081):



The above speech repair includes reparandum *one because they're*, editing term *i'm sorry* and alteration *two because they're.* When we found the identification clue *i'm sorry* we activated detection rule which marked it as a possible editing term in the speech repair. The editing term immediately follows the interruption point. The first word after the editing term is *two*, for which we find a word replacement in the reparandum. The next word is *because*, for which a word correspondence is found and Rule 6 is followed. The last word correspondence is *they're.* The word correspondences are added into the repair pattern because all of them are cross-serial and are not embedded inside another correspondence and therefore adhere to Rule 5.

The modification repairs were analysed according to division into three groups, as shown in Table 2. To make the analysis more relevant only 'clean' cases of word repetitions and word replacements are considered as word repetitions and replacements. It means that the repair, i.e. either repetition or replacement, does not include combination of repetition, replacement or deletion. Therefore, repair patterns such as mx.mx, mr.mr, or rm.rxm are considered as other speech repairs. 'Clean' word repetitions such as m.m, mm.mm or mmm.mmm are more frequently used than word replacements and most of them are marked with word fragment.

	word repetitions		word replacement		other speech repairs		TOTAL	
with fragment	12	20%	2	3.3%	4	6.7%	18	30%
with editing term	0	0%	0	0%	2	3.3%	2	3.3%
TOTAL	12	20%	2	3.3%	6	10%	**20**	**33.3%**

Table 2: Modification repairs marked by word fragment or editing term in Intro Biology Discussion Section

The number 33.3 per cent represents modification speech repairs marked by word fragment or editing term. It is proved again that if no word correspondence would be taken into consideration as a speech repair clue more than 66 per cent of

modification speech repairs would be missed. Editing terms were used only with other speech repairs and word fragment was used mostly with word repetitions:

Example 13: (DIS175JU081):

*...**<M>~~cont- contracts~~:M: (m.m)** pneumonia and dies...* **m . m**

Therefore, without using word correspondence as one of the detection clue we would miss many speech repairs, such as:

Example 14: (DIS175JU081):

*...genetic information **<M>as the as the:M:** parent.* **mm . mm**

	Speech repairs: Female and male speakers			
	fragment	**editing term**	**correspondence**	
Female	**17** *16.8* (0)	**2** *1.9* (0.01)	**37** *37.3* (0)	56
Male	**1** *1.2* (0.03)	**0** *0.1* (0.13)	**3** *2.7* (0.04)	**4**
	18	**2**	**40**	**60**

χ^2 = 0.223, df = 2, P(χ^2 >0.223) = 0.894

Table 3: Speech repairs in Intro Biology Discussion Section: Male and female contrasted

We used a Chi-square test in order to determine whether there is a significant difference between the expected and observed numbers. In other words, is the difference between observed and expected numbers due to sampling error, or is it a real difference. Our results show that the difference in using speech repair clues by male and female speakers is not significant and that the differences are simply a result of the random sample that was selected.

The second analysed text was Intro to American Politics, which includes 18 speakers out of whom six are male and twelve are female native speakers. The transcript, which is highly interactive, has 7,220 words. All the examples from this transcript are marked as DIS495JU119.

Types of repairs								
Total			Identification clues					
			with fragment		with editing term		word correspondence	
Modification repairs	125	100%	30	24%	16	12.8%	79	63.2%
Abridged repairs	11	100%	11	100%	0	0%	0	0%
Total	136	100%	41	30.1%	16	11.8%	79	58.1%

Table 4: Types of speech repairs in Intro to American Politics Discussion Section

As the above table shows, out of 136 repairs only eleven were classified as abridged repairs. Abridged repairs do not have reparandum and consist of word fragment only.

Example 15: (DIS495JU119):

...we think, ***s-*** *we need some sort...*

When we come to the word fragment *s-* in the above sentence, we look for word correspondences and the extent of reparandum and alteration. However, no word correspondence is found and the speech repair consists only of word fragment which classifies this repair as an abridged repair.

The analysed transcript includes 125 modification speech repairs. These are marked mostly by word correspondence, then word fragment and lastly by editing term.

Example 16 (DIS495JU119):

...it's our attitudes, ***<M>****or our* ***um*** *and our* ***:M: (rm.rm)*** *involvement* ***<M>****in, in* ***:M: (m.m)*** *politics and,* ***<M>****how we how we* ***:M: (mm.mm)*** *perceive this political system…*

The above analysed sentence includes three modification clues. The first speech repair is marked by editing term and the other two are marked by word correspondence. The first repair is identified by editing term *um.* A word correspondence for the next word *and* is found. It is a word replacement for word *or* in the reparandum. The next word encountered is *our* which is the last word for which we can find a word correspondence.

or our ***um*** *and our involvement*
$\mathbf{r_1\ m_1}$ $\mathbf{_{ip}}$ **et** $\mathbf{r_1\ m_1}$

The next repair is detected by word correspondence, namely word repetition of the word *in:*

involvement in, in politics
$\mathbf{m_1\ _{ip}\ m_1}$

This speech repair is acceptable since Rule 4 is obeyed and the word correspondence is cross-serially same as the last repair in the sentence:

how we how we perceive
$\mathbf{m_1\ m_2\ _{ip}\ m_1\ m_2}$

The last speech repair has two word correspondences *how* and *we.* Here the repair is complete because there are no other words between the first and the last accounted words. The result of correction of the analysed sentence is:

> *...it's our attitudes, ~~or our **um**~~ and our involvement ~~in,~~ in politics and, ~~how we~~ how we perceive this political system…*

	word repetitions		word replacement		other speech repairs		TOTAL	
with fragment	19	15.2%	2	1.6%	9	7.2%	30	24%
with editing term	10	8%	1	0.8%	5	4%	16	12.8%
TOTAL	29	23.2%	3	2.4%	14	11.2%	**46**	**36.8%**

Table 5: Modification repairs marked by word fragment or editing term in Intro to American Politics Discussion Section

Once again, we would miss more than half of modification speech repairs if we did not consider word correspondence as a speech repair clue. Modification repairs marked with word fragment or editing term make only 36.8 per cent of modification speech repairs. Speakers mostly repeated words or replaced one word with another or combination of both.

	Speech repairs: Female and male speakers			
	fragment	**editing term**	**correspondence**	
Female	**20** *14.9* (1.76)	**4** *7.9* (1.95)	**38** *39.2* (0.04)	**62**
Male	**10** *15.1* (1.73)	**12** *8.1* (1.92)	**41** *39.8* (0.04)	**63**
	30	**16**	**79**	**125**

χ^2 = 7.44, df = 2, P(χ^2 >7.44) = 0.0242

Table 6: Speech repairs in Intro to American Politics Discussion Section: Male and female contrasted

Numbers in bold represent the observed numbers for each category. In order to calculate the Chi-square we selected two variables to divide the data: gender and speech repair clue. The test shows how gender and speech repair clues are completely independent of one another. In other words, we test if the differences are due to chance alone or not. To test this hypothesis we need to calculate how the data should be distributed if our hypothesis of independence is correct. These numbers are marked in the above table in italics. The results show that the differences are significant at level 0.05. Therefore, we can state we deal with dependent variables and speech repair clues are dependent on gender.

6 Conclusion

The results of the analysis of these transcripts show that by using our algorithm we can clearly identify modification and abridged repairs. The present paper shows that using the proposed algorithm we can detect most speech repairs without using higher levels of syntactic or semantic knowledge. Abridged repairs were marked as possible speech repairs and only those which consisted of word fragment were taken into account. We have come to two quite important findings. First, word fragments and editing terms mark about 30 per cent of all modification repairs. The results show that the most common speech repair clue is word correspondence. Speakers do not use editing term as a speech repair indicator frequently, which is seen in the results. Second, the presence of word fragment or editing term does not give conclusive evidence as to whether the repair is a modification of abridged repair. Detection cues taken from Heeman and Allen (1994) were extended by further three rules to improve our work. This paper provides a clear distinction between word repetitions and word

replacements marked by word fragment which may be crucial when analysing speech repairs. In order to make the application of this algorithm more objective more work needs to be done. It would be essential to show that our model works also on other transcripts from different corpora. Although the algorithm shows to be promising, it would be also very interesting to analyse transcripts by two independent speech repair analysts and to show the percentage of precision with which the algorithm can detect a particular type of speech repair.

References

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

Bear, J., Dowding, J. and Shriberg, E. E. (1992) 'Integrating multiple knowledge sources for detection and correction of repairs in human-computer dialogue.' In: Thomson, H. S. (ed.) *Proceedings of the 30th Annual Meeting on Association for Computational Linguistics*. Stroudsburg: Association for Computational Linguistics. 56-63.

Brennan, S. E. and Schober, M. F. (2001) 'How listeners compensate for disfluencies in spontaneous speech.' *Journal of Memory and Language 44/2*, 274-296.

Brinton, L. (1996) *Pragmatic Markers in English. Grammaticalization and Discourse Functions.* Berlin: Mouton de Gruyter.

Fox Tree, J. E. (1995) 'The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech.' *Journal of Memory and Language 34/6*, 709-738.

Heeman, P. A. and Allen, J. F. (1994) 'Detecting and correcting speech repairs.' In: Pustejovsky, J. (ed.) *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics.* New Mexico: Association for Computational Linguistics. 295-302.

Heeman, P. A. (1999) 'POS tags and decision trees for language modelling.' In: Fung, P. and Zhou, J. (eds) *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*. Maryland: Association for Computational Linguistics. 129-137.

Hindle, D. (1983) 'Deterministic parsing of syntactic non-fluencies.' In: Marcus, M. (ed.) *Proceeding of the 21st Annual Meeting.* Stroudsburg: Association for Computational Linguistics. 123-128.

Johnson, K. (1997) 'Identification and correction of speech repairs in the context of an automatic speech recognition system.' In: *Durham Theses.* Durham University, 1997. Online document. 13 May 2012 <http://etheses.dur.ac.uk/5306>.

Levelt, W. J. M. (1983) 'Monitoring and self-repair in speech.' *Cognition 14/1,* 41-104.

Levelt, W. J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge, Massachusetts: The MIT Press.

MacKay, D. G. (1970) 'Spoonerisms: The structure of errors in the serial order of speech.' *Neuropsychologia 8/3,* 323-350.

MacKay, D. G. (1972) 'The structure of words and syllables: Evidence from errors is speech.' *Cognitive Psychology 3/2,* 210-227.

Nakatani, C. H. and Hirschberg, J. (1994) 'A corpus-based study of repair cues in spontaneous speech. *'The Journal of the Acoustical Society of America 95/3*, 1603-1616.

Schiffrin, D. (1988) *Discourse Markers.* Cambridge: Cambridge University Press.
Shriberg, E. E. (1996) 'Disfluencies in SWITCHBOARD.' In: Bunnell, H. T. and Idsardi, W. (eds) *Proceeding of the International Conference on Spoken Language Processing*, *Addendum*. Philadelphia: IEEE Service Center. 11-14.
Shriberg, E. E. (1994) *Preliminaries to a Theory of Speech Disfluencies.* Doctoral thesis, University of California at Berkeley.
Stolcke, A. and Shriberg, E. (1996) 'Statistical language modelling for speech disfluencies.' In: Reeves, S. J. (ed.) *Proceedings of the International Conference on Audio, Speech and Signal Processing.* Atlanta: Institute of Electrical and Electronics Engineers. 405-408.
Oviatt, S. (1995) 'Predicting spoken disfluencies during human-computer interaction.' *Computer Speech and Language 9/1,* 19-35.

Sources

Simpson, R. C., Briggs, S. L., Ovens, J. and Swales, J. M. (2002) '*The Michigan Corpus of Academic Spoken English*.' Ann Arbor, MI: The Regents of the University of Michigan.