

Pragmatically annotated corpora in speech-to-speech translation

Darinka Verdonik

Faculty of Electrical Engineering and Computer Science, University of Maribor
Smetanova ul. 17, 2000 Maribor, Slovenia
darinka.verdonik@uni-mb.si

Abstract

The aim of this paper is to discuss and specify some pragmatic language categories that could be used as attributes in spontaneous speech corpora, especially the corpora used for developing speech-to-speech translation systems components. When developing the speech-to-speech translation, researchers have to deal with spontaneous (conversational) speech phenomena like hesitations, turn-taking behaviors, self-repairs, false starts, filled pauses... This makes speech-to-speech translation a very hard task, with much space for improvement. Language technologies use linguistically annotated corpora and lexica (morphologic, syntactic, semantic...) to achieve better performance. In this paper I suggest to include pragmatic attributes of annotation to deal with some of the above mentioned phenomena of spontaneous speech.

Pragmatically annotated corpora in speech-to-speech translation

Namen tega prispevka je definirati nekatere pragmatične jezikovne kategorije, ki jih lahko uporabimo kot atribute v pragmatično označenih govornih korpusih, zlasti tistih, ki se uporabljajo pri razvoju sistemov strojnega simultanelega prevajanja govora. Raziskovalci, ki delajo na področju tehnologije strojnega simultanelega prevajanja govora, opozarjajo, da je v pogovoru polno elementov, kot so obotavljanja, menjavanje vlog, samopopravljanja, napačni začetki, premori... Te značilnosti so problematične za strojno simultano prevajanje govora in zahtevajo ustrezne rešitve. Pri razvoju jezikovnih tehnologij se uporabljajo jezikoslovno označeni korpusi in slovarji (oblikoslovni, skladenski, semantični...), saj pripomorejo k večji uspešnosti tehnologije. V tem prispevku predlagam vključevanje pragmatičnih atributov za označevanje govornih korpusov, da bi na tak način premoščali težave pri razvoju strojnega simultanelega prevajanja govora, ki jih navajam zgoraj.

1. Introduction

Many projects developing speech-to-speech translation systems (eg. Verbmobil – <http://verbmobil.dfki.de/>, Janus – <http://www.is.cs.cmu.edu/mie/janus.html>, EuTrans – <http://www.cordis.lu/esprit/src/30268.htm>, Nespole! – <http://nespole.itc.it/>) had to face the reality of spontaneous (conversational) speech. It is usually observed that spontaneous speech includes »disfluencies, hesitations (um, hmm, etc.), repetitions« (Waibel, 1996), »pauses, hesitations, turn-taking behaviors, etc.« (Kuremtasu et al., 2000), »self-interruptions and self-repairs« (Tillmann, Tischer, 1995), disfluencies such as »a-grammatical phrases (repetitions, corrections, false starts), empty pauses, filled pauses, incomprensible utterances, technical interruptions, and turn-takes« (Costantini et al., 2002). Such characteristics can cause many problems for automatic speech recognition and speech centered translation, which are part of a speech-to-speech translation system.

In linguistics (I refer to linguistics not only as a study of language system, but also as a study of language use) most of the above mentioned characteristics are considered as pragmatic, and are the subject of interest in some fields of discourse analysis or pragmatics. In this paper I will try to specify some basic pragmatic attributes that cover some of these spontaneous speech characteristics and that could be easily annotated in spontaneous speech corpora. There have been few tries to annotate some pragmatic elements in speech corpora for use in developing speech technologies or natural language processing (eg. Heeman et al., 1998; Heeman, Allen, 1999; Miltsakaki et al., 2002), however pragmatics as level of annotation in language resources is far from being broadly discussed or accepted. The processing problems

when dealing with spontaneous speech encourage us to try it and discuss it, and to encourage some further discussion on pragmatically annotated corpora is one of the aims of this paper.

The research presented in this paper is based on a corpus in the Slovenian language, therefore the attributes for annotation are defined for the Slovenian, but the presented concepts themselves are general. More details on all aspects of the research which is a basis for this discussion can be found in (Verdonik, 2006).

For the Slovenian language, speech-to-speech translation system recently became an interesting issue. (Žganec et al., 2005) present a design concept of the Voice TRAN, speech-to-speech translation system that would be able to translate simple domain-specific sentences in the Slovenian-English language pair. The other concept for the speech-to-speech translation system including the Slovenian language is named Babilon, and it is presented on the http://www.dsplab.uni-mb.si/Dsplab/Slo/Projects_slo_demo.php.

The structure of this article is the following: first I describe the corpus (Turdis-1) that was used to track, analyze and specify the pragmatic attributes for annotation. Chapters 3, 4 and 5 bring specification of the three levels of pragmatic annotation in spontaneous speech corpora: conversation structure (sections, turns, utterances), discourse markers and repairs. In chapter 6 some conclusions are drawn.

2. Data for the analysis – the Turdis-1

For the analysis I used a speech corpus of telephone conversations in tourism. Tourist domain seems to be one of the most promising and popular for speech-to-speech translation systems (it was the main or one of the main domains for speech-to-speech translation projects like

Verbmobil, Janus, Nespole!, EuTrans...). Since the tourist domain in general is too broad as a domain of interest for typical speech-to-speech translation applications, it was further restricted to the following sub-domains:

- telephone conversations in tourist agency
- telephone conversations in tourist office
- telephone conversations in hotel reception

Conversations with professional tourist agents and real tourist organizations were recorded. The callers were contacted personally; they were mostly employees and students of the University of Maribor. The tourist organizations which participated in recording were: two local hotels, local tourist office and four local tourist agencies. All conversations were in the Slovenian language which was also the mother tongue of all the callers. Recorded material was transcribed using the Transcriber tool (<http://trans.sourceforge.net/en/presentation.php>). We considered some of the EAGLES recommendations (<http://www.lc.cnr.it/EAGLES96/spokentx/>) and principles of transcribing BNSI Broadcast News database (Žgank et al., 2004) when transcribing. More details about recording and transcribing can be found in (Verdonik, Rojc, 2006).

From the recorded material 30 conversations were selected for the present study. This selection is named Turdis-1. The total length of the recordings in the Turdis-1 is 106 minutes, the average length of a conversation 3,5 minutes, the number of tokens is 15,717, number of word forms 2735, number of utterances 2171. The table 1 shows more details about number and length of conversations, and the table 2 about number and gender of speakers.

Table 1: Number and total length of conversations in the Turdis-1 database.

	No. of conv.	Total length
Tourist agency	14	53,33 min.
Tourist office	8	28,1 min.
Hotel reception	8	24,38 min.
Total	30	106,2 min.

Table 2: Gender of the speakers (callers and tourist agents) in the Turdis-1 database.

	Male	Female
Tourist agents	3	17
Callers	14	10
Total	17	27

3. Conversation structure

When processing natural speech, we need to find the most appropriate segments for processing first. This is especially important when talk of one speaker is longer than what is usually understood as a segment (in speech technologies) or an utterance (in discourse analysis). So the basic units of transcribing conversations are usually turns and segments/utterances. Both need some further clarifications.

3.1. Turns

Turn is understood as the talk of one speaker before the next speaker starts to talk. But in natural conversation it often happens that at the exchange point talk of both speakers overlap (so called overlapping speech). When transcribing, different solutions are possible for overlapping speech. The one I suggest here is that we segment overlapping speech as a new, overlapping turn, but include special tags for tracking connections between the text in overlapping speech and the text in the previous or the following segments. This is because when we tag the overlapping speech as a special segment, we have probably put some borders to the text which are not consistent with prosodic, syntactic and semantic borders (i.e. utterances), therefore also the previous or/and the following segment may be syntactically, semantically and prosodically incomplete.

Another issue of discussion is how to transcribe backchannel signals (short expressions that hearer pronounces in order to confirm to the speaker that he is listening, that he understands, that he is interested...). I suggest not to annotate them as overlapping speech, but as special speech events.

3.2. Segments/utterances

Segments/utterances are usually the basic units for processing speech. In written text corresponding units could be sentences. It is quite clear what counts as a sentence in the written text, but there seems to be less agreement on what counts as an utterance in the spontaneous speech. For use in developing speech technologies, I believe syntactic, semantic and also prosodic features (especially intonation and pauses) must be considered when segmenting speech to utterances.

3.3. Sections

Sections can be as well an interesting attribute for annotating conversation structure. Here, I will consider only opening and closing sections in a conversation, which are very important for pragmatically successful conversation. It is open for a discussion, whether other topic shifts during the course of a conversation are to be annotated.

In opening and closing sections in the analyzed telephone conversations I find more or less standard pragmatic acts and standard phrases used. This can make speech-to-speech translation task easier.

In an opening section a caller starts communication by telephone ring. First talk in conversation is agent's, always introducing himself and/or organization he works at, very often also greeting. Next turn is caller's, he is always greeting, very often introducing himself, and after this explaining a reason for the call.

Closing sections are very delicate, because none of the participants in a conversation should feel forced to end the conversation. Analysis shows that discourse markers *dobro/v redu/okej/prav* (Eng. *good, alright, right, okay, well, just*) can be used as signals for closing the conversation. Next act is usually thanking, which is also a signal for closing the conversation. The last act of every conversation are greetings.

4. Discourse markers

Discourse markers are expressions like *oh, well, now, y'know, and...* In conversation, they are most often used the way that they do not contribute much to the propositional content, but have more or less pragmatic, communicative functions. As such I find them an interesting attribute for annotation.

Studies of discourse markers were increasing in the last decades, not only for English but for many languages worldwide (see for example special issues of *Discourse Processes* (1997, 24/1) and *Journal of Pragmatics* (1999, 31/10), workshops like *Workshop on Discourse Markers* (Egmond aan Zee, Netherlands, January 1995) or *COLING-ACL Workshop on Discourse Relations and Discourse Markers* (Montreal, Canada, August 1998), books like (Schiffrin, 1987; Jucker, Ziv, 1998; Blakemore, 2002) etc.).

There are basically three different approaches to discourse markers: coherence-based (most known is Schiffrin's research (1987)), relevance theory approach (very known is work of Blakemore (1992; 2002)) and grammatical-pragmatic approach (Fraser, 1990; 1996; 1999).

For the Slovenian language there are only few researches of what I here name discourse markers, some more some less close to the discursive perspective: (Gorjanc, 1998), (Schlamberger Brezar, 1998), (Smolej, 2004a). (Pisanski, 2002; 2005) represents broader research on text-organizing metatext in research articles.

4.1. Guidelines for annotating discourse markers

When overviewing the researches on discourse markers, we find out that there is still no agreement on what counts as a discourse marker. But what we find common is acknowledgement that there are two basically different kinds of meaning, communicated by utterances: Schiffrin (1987) distinguishes ideational plane on the one hand, and exchange structure, action structure, participation framework and information state on the other hand; Blakemore (2002) distinguishes conceptual vs. procedural meaning; Fraser (1996) distinguishes propositional content and pragmatic information; researches on metadiscourse (eg. Pisanski, 2002; 2005) distinguish metadiscourse and propositional content. Even though these distinctions are not completely parallel, they have a lot in common. Discourse markers in these distinctions are expressions that function primarily pragmatically and contribute the least to the ideational/propositional/conceptual domain.

As one of the most extensive, detailed and also most often cited studies of discourse markers, based on recorded material of natural conversations, I take work of Schiffrin (1987) as the example. I keep the distinction between ideational structure and all the other planes of talk. Similar distinction is set by Redeker (1990), who distinguishes markers of ideational structure and markers of pragmatic structure. Since we are interested in expressions that function primarily pragmatically and contribute the least to the ideational/propositional/conceptual domain, the aim was to annotate discourse markers that function primarily as pragmatic markers.

According to this basic theoretical framework I annotate discourse markers in the Turdis-1 corpus and make a detailed analysis of annotated expressions in order to define their pragmatic functions in a conversation, to confirm or reject the chosen expressions, and to point to problematic points in annotating discourse markers.

4.2. Expressions functioning as discourse markers

According to the framework for annotating, defined in previous chapter, I annotated the expressions that contribute the least to the propositional content of an utterance in the Turdis-1 corpus. Such expressions were: *ja* (Eng. *yes, yeah, yea, well, I see* – please notice that the English expressions are only approximate description to help readers who do not speak the Slovenian language; it is based on the author's knowledge of English, Slovenian-English dictionary and British National Corpus (<http://www.natcorp.ox.ac.uk/>); usage of discourse markers is culturally specific and we would need a comparative study to be able to specify the English equivalents more exactly), *mhm* (Eng. *mhm*), *aha* (Eng. *I see, oh*), *aja* (Eng. *I see, oh*), *ne?/a ne?/ali ne?/jel?* (no close equivalent in English, a bit similar to *right?, isn't it?* etc.), *no* (Eng. *well*), *eee/mmm/eeem...* (Eng. *um, uh, uhm*), *dobro/v redu/okej/prav* (Eng. *good, alright, right, okay, well, just*), *glejte/poglejte* (Eng. *look*), *veste/a veste* (Eng. *y'know*), *mislim* (Eng. *I mean*), *zdaj* (Eng. *now*), and backchannel signals: *mhm* (Eng. *mhm*), *aha* (Eng. *I see, oh*), *ja* (Eng. *yes, yeah, yea, I see*), *aja* (Eng. *I see, oh*), *dobro* (Eng. *okay, alright, right*), *okej* (Eng. *okay, alright, right*), *tako* (Eng. *thus*), *tudi* (Eng. *also*), *seveda* (Eng. *of course*). I use the term backchannel signals for isolated uses of discourse markers when hearer does not take over the turn and also does not show intention to do so, but merely expresses his attention, agreement, confirmation, understanding etc. of what speaker is saying.

The results of the analysis showed that some of these expressions always function as discourse markers: such are *mhm* (Eng. *mhm*), *aha* (Eng. *I see, oh*), *aja* (Eng. *I see, oh*), *no* (Eng. *well*), *eee/mmm/eeem...* (Eng. *um, uh, uhm*).

Others (eg. *(a/ali) ne?* (in Eng. similar *right?, isn't it?* etc.), *dobro/v redu/okej/prav* (Eng. *good, alright, right, okay, well, just*), *glejte/poglejte* (Eng. *look*), *veste/a veste* (Eng. *y'know*), *mislim* (Eng. *I mean*)), can function either as a discourse marker, for example *dobro* as discourse marker:

K25: *dobro* gospa najlepša hvala da ste se tako potrudli ne? / *okay* madam thank you so much for your efforts

or as an important element of propositional content, for example *dobro* in a proposition:

K39: *ker[+SOGOVRNIK_je] jim nikol nič ni dobr in vedno etc. / because[+OVERLAP_yes] nothing is ever good enough for them and they always etc.*

but differences between both usages are easy to recognize for a human annotator. For automatic detection it may be helpful, that (according to the analysis of the Turdis-1 corpus) the analyzed expressions in the function of discourse marker are usually positioned at the borders between utterances.

But for some of the analyzed expressions, particularly *ja* (Eng. *yes, yeah, yea, well, I see*) and *zdaj* (Eng. *now*), the border between discourse marker and propositional function was blurred. There were usages where these expressions were functioning clearly pragmatically, other usages where they were functioning clearly as part of a proposition, but also usages where it was not clear which of these two basic functions was more important, for example:

K39: eem treh ali pa štirih Nemcev to zaenkrat še ne vem s() se pravi oni[+SOGOVORNIK_mhm] so pač iz Nemčije[+SOGOVORNIK_mhm] / um three or four German people this I do not know exactly s() so they[+OVERLAP_mhm] are from Germany[+OVERLAP_mhm]

K39: #nikol# še niso bli v Sloveniji / they have #never# been to Slovenia

K39: in zdej bi jih ze() pač za takšne štir pet dni počitnic ki jih bojo meli v Sloveniji bi jim pač seveda etc. / and now I would f() for some four five days of vacation they will have in Slovenia I would of course etc.

Such examples confirm that the border between pragmatic and semantic level is certainly not a clear cut, and annotating in corpora needs careful considerations on every step.

The above mentioned expressions are of course not all discourse markers of the Slovenian language. But the outlined considerations may be the starting point for further discussion about discourse markers. In the Turdis-I corpus, discourse markers were manually annotated, but the analysis showed that further annotation can be at least partially automatic.

4.3. Pragmatic functions of the analyzed discourse markers

Since the analyzed expressions do not contribute much to the content of a message, we can suppose that they have some pragmatic functions. This suggestion is supported by the fact that the analyzed discourse markers were used more than 2000 times in 15,000 tokens corpus, what corresponds to something more than 13% of all tokens, and that is quite a lot. I used the conversational analysis method (see Levinson, 1983, 286-287), and as the results of the analysis I specified the following pragmatic functions of discourse markers:

- signaling connections to propositional content (backward or forward)
- building relationships between participants in conversation (for example checking and confirming a hearer's presence, interest in conversation, understanding...)
- expressing speaker's attitude to the content of the conversation (eg. surprise, dissatisfaction...)
- organizing the course of conversation (signals in turn-taking system, signals for changing the topic and ending a conversation, signals of disturbances (eg. self-repairs) in utterance structure/production)

5. Self-repairs

As I pointed out in the introduction, spontaneous speech characteristics like disfluencies, self-interruptions and self-repairs, corrections, false starts etc. are problematic for spontaneous speech processing. In pragmatics most of these phenomena are treated as

disfluencies or as self-repairs. In the Slovenian language the phenomena did not draw special attention before this research, it was merely noticed for example in (Smolej, 2004b; Krajnc, 2004).

Some of the most cited and known researches on self-repairs were done by (Schegloff, Jefferson, Sacks, 1977; Schegloff, 1979), by (Levelt, 1983), also (Allwood et al., 1990) etc. Disfluencies were studied for example by (Lickley, 1994; Shriberg, 1994; Tseng, 1999). They consider the term more neutral, but it includes broader phenomena (for example for Shriberg (1994) disfluencies are *um*'s, repetitions and self-repairs, for Tseng (1999) restarts, repetitions, pauses, speech errors, speech repairs). Here, based on pragmatic researches of the phenomena, I discuss only self-repairs. I try to define them the way that we can use a definition of the self-repair to annotate the part of an utterance that needs to be eliminated in further processing because it is unfinished structure, replaced by another structure.

I suggested to annotate segment/utterance the way that it can be treated as a basic unit for processing, and I want to define the self-repair the way that it is a structure that needs to be eliminated, therefore I define the self-repair as a phenomenon on the level of a segment/utterance.

5.1. Defining self-repairs

(Blanche-Benveniste, 1991; Smolej, 2004b in the Slovenian linguistics) discuss two levels or axes of producing a text: syntagmatic (horizontal) and paradigmatic (vertical). In the eyes of this theory a self-repair is a structure, where speaker does not continue fluent speech, but stops and goes back to some previous point on syntagmatic level of text, for example:

*kolko pa potem stane nočitev pa recimo **da so** eee
da je poln penzijon /
and how much then costs one night for example **that we** um
that it is with
breakfast*

But when listing, explaining, inserting structures etc. speaker also goes back to some previous point on syntagmatic level of text, for example when explaining:

*študenti organiziramo en tak letni **sestanek** oziroma
srečanje /
the students we organize some sort of annual **meeting** or
gathering*

A typical self-repair as I want to define it here always begins by cut-off, therefore I do not define examples as the last one as a self-repair.

Next, I analyze pragmatic aspects of self-repairs. First I try to define reasons for cutting-off. I find that they may be circumstantial (bad telephone connection), social (especially turn-taking), or psychological (a speaker needs more time to prepare what he will say, a speaker changes his strategy how to say something, a speaker notices a mistake in what he told, a speaker has problems when pronouncing and re-pronounces some previous element(s)). It is only when a speaker changes his strategy, when he notices a mistake or has problems when pronouncing, that we can talk about self-repair. At the same time the first condition has to be fulfilled, i.e. a

speaker goes back to some previous point on syntagmatic level of text.

According to this definition I annotate self-repairs in the Turdis-1 corpus. They appear in 185 utterances, which is approx. in 8% of all the utterances.

5.2. Structure of self-repairs

I find four basic structure elements of self-repairs:

1. A part of a text that will be corrected, therefore it should be eliminated in automatic processing. In 90% of examples in the Turdis-1 corpus it is not longer than 3 words.
2. A cut-off.
3. Self-repair signals: metadiscursive element(s) can follow right after cut-off, for example discourse markers *eee* (Eng. *um*), *zdaj* (Eng. *now*), *mislim* (Eng. *I mean*) etc., pause, prolonged vowel etc. But these are used only in 55% of all self-repairs in the Turdis-1 corpus.
4. Repairing element/s, i.e. the new text that replaces the part of a text that was corrected. In 65% in the Turdis-1 corpus repaired elements include repetition of at least one token or some phonemes of the cut-off token from the part of a text that was corrected.

6. Conclusion

In this paper I have discussed the idea to include pragmatic tags to spontaneous speech corpora used for developing speech-to-speech translation components (and of course for other speech technologies, dealing with spontaneous speech, for example dialog systems). Based on the analysis of the corpus (Turdis-1) of telephone conversations I tried to define three basic levels of annotation.

Annotating basic conversation structure elements – segments/utterances, turns, sections – is usual in conversation corpora. In this paper I point to some problematic points of annotation: annotating overlapping speech and backchannel signals, defining utterances to achieve consistency of annotation, annotating opening and closing sections which include mostly standard pragmatic acts and phrases.

Next, I suggest annotating discourse markers. Discourse markers attracted much attention of linguists, but annotating discourse in speech corpora used for developing speech technologies is not broadly accepted yet, even though there are/were some tries. Overview of the researches of discourse markers in discourse analysis shows that there is no agreement on what counts as discourse marker. Therefore I try to specify a framework for annotation that would be the most useful for speech-to-speech translation purposes. As discourse markers, I specify the expressions that contribute the least to the propositional content of an utterance, but have mostly pragmatic functions. The analysis shows that most of them are used at the borders between utterances, so they can be used to help segmenting spoken text to segments/utterances. They are very frequently used in a conversation – more than 13% of all the words in the Turdis-1 corpus. This supports the idea that discourse markers are very important elements of natural conversation.

Last I try to define self-repairs the way that self-repair as attribute in speech corpora annotates a part of spoken text that needs to be eliminated in further processing – it is unfinished structure, replaced by some other structure. I conclude that self-repairs are an event where a speaker goes back to some previous point on syntagmatic level of text, in order to change a strategy, correct a mistake or repair problems when pronouncing. Self-repairs are present in approx. 8% of all the utterances in the Turdis-1 corpus.

Possibilities for further annotation of pragmatic elements in spontaneous speech corpora are many more, for example speech acts, adjacency pairs, other metatextual elements, repetitions etc. There is a wide area for researches, experiments and discussion.

7. Acknowledgements

I sincerely thank to all the tourist companies that participated in recording: the tourist agencies **Sonček**, **Kompas**, **Neckermann Reisen** and **Aritours**, to the **Terme Maribor**, especially the **Hotel Piramida** and the **Hotel Habakuk**, and to the **Mariborski zavod za turizem** with the tourist office **MATIC**. I also thank to all the tourist agents in these companies who participated in recording and to all the callers who were ready to use the Turdis system.

8. References

- Allwood, J., J. Nivre, E. Ahlsen. 1990. Speech management: On the non-written life of speech. *Nordic Journal of Linguistics*, 13/1.
- Blakemore, Diane. 1992. *Understanding utterances*. Oxford, Cambridge: Blackwell Publishers.
- Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Blanche-Benveniste, Claire. 1991. *Le français parle*. Etudes grammaticales. Paris: CNRS.
- Constantini, E., S. Burger, F. Pianesi. 2002. NESPOLE!'s multilingual and multimodal corpus. In proceedings of the 3rd International Conference on Language Resources and Evaluation 2002, LREC 2002, Las Palmas, Spain.
- Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics*, 14, 383-395.
- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics*, 6/2, 167-190.
- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics*, 31, 931-952.
- Gorjanc, V. 1998. Konektorji v slovničnem opisu znanstvenega besedila. *Slavistična revija*, XLVI/4, 367–388.
- Heeman, Peter, Donna Byron, James Allen. 1998. Identifying Discourse Markers in Spoken Dialogue. In *Working Notes of AAAI Spring Symposium on Applying Machine Learning to Discourse Processing*, Stanford, CA.
- Heeman, Peter, James Allen. 1999. Speech repairs, intonational phrases and discourse markers: modeling speakers' utterances in spoken dialog. *Computational Linguistics*, 25(4).
- Jucker, Andreas H., Yael Ziv (Eds.). 1998. *Discourse Markers: Descriptions and Theory*. Amsterdam: John Benjamins.

- Krajnc, M. 2004. Besediloskladenjske značilnosti javne govornje besede (na gradivu mariborščine). *Slavistična revija*, 52/4, 475-498.
- Kurematsu, A., Akegami, Y., Burger, S., Jekat, S., Lause, B., MacLaren, V., Oppermann, D., Schultz, T. 2000. Verbmobil Dialogues: Multifaced Analysis. In Proceedings of the International Conference of Spoken Language Processing.
- Levelt, W. J. M. 1983. Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Lickley, Robin J. 1994. *Detecting disfluency in spontaneous speech*. PHD thesis. University of Edinburgh.
- Miltsakaki, E., R. Prasad, A. Joshi, B. Webber. 2004. The Penn Discourse Treebank. In Proceedings of the Language Resources and Evaluation Conference'04, Lisbon, Portugal.
- Pisanski, Agnes. 2002. Analiza nekaterih metabesedilnih elementov v slovenskih znanstvenih člankih v dveh časovnih obdobjih. *Slavistična revija*, 50/2, 183-197.
- Pisanski Peterlin, Agnes. 2005. Text-organising metatext in research articles: an English-Slovene contrastive analysis. *Engl. specif. purp. (N.Y. N.Y.)*, 24/3, 307-319.
- Redeker, G. 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14, 367-381.
- Schegloff, E., G. Jefferson, H. Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language*, 53/2, 361-382.
- Schegloff, E. 1979. The relevance of repair to syntax-for-conversation. In Givon, T. (ed.). *Syntax and Semantics 12, Discourse and Syntax*. New York: Academic Press. 261-286.
- Schiffirin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Schlamberger Brezar, M. 1998. Vloga povezovalcev v diskurzu. In *Jezik za danes in jutri*. Ljubljana: Društvo za uporabno jezikoslovje Slovenije. 194-202.
- Shriberg, E. E. 1994. *Preliminaries to theory of speech disfluencies*. PHD thesis. University of California at Berkeley.
- Smolej, Mojca. 2004a. Členki kot besedilni povezovalci. *Jezik in slovstvo*, 49/5, 45-57.
- Smolej, Mojca. 2004b. Načini tvorjenja govornega diskurza – paradigmatska in sintagmatska os. In Erika Kržišnik (ed.). *Aktualizacija jezikovnozvrstne teorije na Slovenskem: členitev jezikovne resničnosti (Obdobja, Metode in zvrsti, 22)*. Ljubljana: Center za slovenščino kot drugi/tuji jezik.
- Tillmann, Hans G., Bernd Tischer. 1995. Collection and exploitation of spontaneous speech produced in negotiation dialogues. In proceedings of the ESCA Workshop on Spoken Language Systems, 217-220, Vigsø.
- Tseng, Shu-Chuan. 1999. Grammar, prosody and speech disfluencies in spoken dialogues. PHD thesis. University of Bielefeld.
- Verdonik, Darinka, Matej Rojc. 2006. Are you ready for a call? – Spontaneous conversations in tourism for speech-to-speech translation systems. In proceedings of the 5th International Conference on Language Resources and Evaluation, Genoa, Italy.
- Waibel, Alex. 1996. Interactive translation of conversational speech. *IEEE Computer*, 29/7, 41-48.
- Žganec Gros, J., F. Mihelič, T. Erjavec, Š. Vintar. 2005. The VoiceTRAN Speech-to-Speech Communicator. In Proc. of the 8th Intl. Conf. on Text, Speech and Dialogue, TDS 2005. Czech Republic, Karlovy Vary.
- Žgank, A., T. Rotovnik, M. Sepesy Maučec, D. Verdonik, J. Kitak, D. Vlaj, V. Hozjan, Z. Kačič, B. Horvat. Acquisition and Annotation of Slovenian Broadcast News Database. 2004. In Proceedings of the 4th International Conference on Language Resources and Evaluation. Lisbon, Portugal.