

The impact of context on discourse marker use in two conversational genres



DARINKA VERDONIK AND ANDREJ ŽGANK

UNIVERSITY OF MARIBOR, SLOVENIA

AGNES PISANSKI PETERLIN

UNIVERSITY OF LJUBLJANA, SLOVENIA

Discourse Studies
Copyright © 2008
SAGE Publications
(Los Angeles, London,
New Delhi and Singapore)
www.sagepublications.com
Vol 10(6): 759–775
10.1177/1461445608096572

ABSTRACT The relationships between text or talk and the context are among the basic fields of pragmatic research and an insight into their nature may contribute to a better understanding of language use. In this article, we use the results of an analysis of discourse marker use in two different conversational genres (telephone conversation and television interviews) in an attempt to examine the impact of context on the use of discourse markers, generalized for each analysed genre. In the first stage of the analysis, we observe important differences between the two genres: discourse markers are far more frequently used in telephone conversations than in television interviews. In the second stage of the analysis, we identify several contextual factors which contribute to the differences in the use of discourse markers. In this way, we obtain insight into this particular aspect of genre context-talk relationships, and identify some of the characteristics of the genres in question.

KEY WORDS: *context-talk relationships, conversational genre, discourse markers, language use, pragmatic expression, sociocognitive context model*

Introduction

Recently, there has been growing interest in a set of pragmatic expressions often referred to as discourse markers (e.g. Archakis, 2000; Blakemore, 2002; Dedaiæ, 2005; Fox Tree, 2006; Fraser, 1999; Fukushima, 2005; Gonzalez, 2005; Jucker and Ziv, 1998; Kroon, 1998; Montes, 1999; Redeker, 1990; Schiffrin, 1987; Schourup, 1999; Tchizmarova, 2005; Wang et al., 2007; Waring, 2003, to name just a few). There have been a number of attempts to provide a definition of this pragmatic category; for example, Schiffrin (1987) uses an interactional sociolinguistic approach in which markers serve to connect utterances on the multiple planes of participation framework; Fraser (1999) defines his own approach as a grammatical-pragmatic perspective; Blakemore (2002) discusses discourse markers within the relevance theory framework.

In this study we will not question the definition of discourse markers. The theoretical framework for our interpretation of discourse markers is based on the work of Redeker (1990) and Schiffrin (1987). Schiffrin (1987) argues that discourse markers are used on five different planes of talk: exchange structure (turns, adjacency pairs), action structure (speech acts), ideational structure (semantic units: propositions or ideas), participation framework (social relations between the speaker and the hearer (e.g. teacher–student), also influenced by the relations of the speaker/hearer to the talk and ideas, presented in the talk), information state (cognitive capacities of the speaker/hearer – organization and management of knowledge and meta-knowledge). We keep the distinction between the ideational structure and all the other planes of talk. A similar distinction is observed by Redeker (1990), who distinguishes between the markers of ideational structure and the markers of pragmatic structure. Some of the conclusions presented by Schiffrin (1987) support the idea that there may be a broader difference between the discourse markers functioning primarily on the ideational plane, and the discourse markers functioning primarily on all the other planes of talk. In this article, we focus on the discourse markers which have mainly a pragmatic function (in Redeker's, 1990, terms, markers of pragmatic structure) and contribute least to the ideational structure, that is, we consider discourse markers to be primarily pragmatic elements conveying no or minimal propositional content. We use this theoretical framework, described in more detail in our previous work (Verdonik et al., 2007a), to define and annotate discourse markers in our analysis.

Studies on discourse markers have focused on various aspects of this pragmatic category in many languages: English (Fox Tree, 2006; Fraser, 1996; Redeker, 1990; Schiffrin, 1978; Waring, 2003), Latin (Kroon, 1998), French (Vlemings, 2003), Spanish (Montes, 1999), Catalan (Gonzalez, 2005), Greek (Archakis, 2001), Japanese (Fukushima, 2005), Taiwan Mandarin (Wang and Tsai, 2007; Wang et al., 2007), Bulgarian (Tchizmarova, 2005), Croatian (Dedaia, 2005), etc. For Slovenian, which is the language used in this analysis, there have been relatively few studies exploring their use: Pisanski Peterlin (2005), Schlamberger Brezar (2007), Smolej (2004), Verdonik (2006) and Verdonik et al. (2007a).

Most studies of discourse markers present a) particular discourse marker(s) use and its (their) meaning and functions from different perspectives, for example, in contrast with the uses of other discourse markers (Fox Tree, 2006), in search of the best method to define the meaning and functions of discourse markers (Schourup, 2001), discussing the information encoded by a particular marker (Dedaia, 2005; Matsui, 2001), examining coherence relations which discourse markers establish by contrasting two languages (Gonzalez, 2005), comparing discourse marker use in different discourse modes (Wang et al., 2007), etc. Other approaches dealing with discourse marker use are less common, for example, Montes (1999) focuses her attention on the development of discourse markers in child speech, the same topic is addressed by Andersen et al. (1999) whose approach also entails contrasting different languages. As far as we are aware, none of the previous studies has explicitly focused on the relationships

between discourse marker use and the context, even though this is one of the basic aspects of pragmatics and may contribute to a better understanding of discourse marker use.

In our understanding of context, we follow Van Dijk's (1997, 2001, 2007) sociocognitive approach. Van Dijk (2007) points out the relevance of cognitive approach for an account of context-text relationships. He neglects the traditionally assumed 'more or less direct relationships between situational, societal, political or cultural aspects of the "environment" of text and talk, on the one hand, and the structures of discourse itself, on the other' (Van Dijk, 2007: 161), proposing instead a sociocognitive account of context, where 'contexts are not "objective" or "deterministic" constraints of society or culture at all, but subjective participant interpretations, constructions or definitions of such aspects of the social environment' (Van Dijk, 2007: 163). In order to understand the nature of context and how it influences and controls text and talk, Van Dijk uses the theoretical notion of mental models, provided by cognitive psychology. He believes that 'contexts are subjective definitions of events or situations, but in this case not of the situation we talk *about* but the situation *in which* we now participate when we engage in talk or text. That is, contexts are the participants' mental models of communicative events' (Van Dijk, 2007: 170). Furthermore, he claims that 'their initial design provide the *plans* that precede all discourse' (Van Dijk, 2007: 170) and that 'context models are the goal of discourse understanding and interaction' (Van Dijk, 2007: 170–1), and most crucially, they 'ongoingly control discourse production and understanding' (Van Dijk, 2007: 171).

In this article,¹ we use the results of an analysis of discourse marker use in two different conversational genres in an attempt to examine the contextual factors which impact the use of discourse markers in each genre. By 'contextual factors' we mean the elements of context models, such as assigned relevance and importance; the roles, positions and social relations of participants; goals, intentions, purposes; group membership and social affiliation; personal association and identification with social groups; social attitudes and ideologies; setting; and others (see Van Dijk, 1997). We assume that most of the differences in discourse marker use in the two genres can be explained by these contextual factors. We are aware of the fact that there may be other contextual factors which may have an impact on discourse marker use, but in this analysis, we limit our attention to those factors which can be identified by contrastive analysis. By defining contextual factors which contribute to differences in the use of discourse markers in the two conversational genres analysed here, we obtain an insight into this particular aspect of context-talk relationships and identify some of the characteristics of the genres in question.

For this purpose, we use an analysis of two conversational genres; we take the concept of genre as it was developed within the framework of the ethnographic tradition (Hymes, 1977; Saville-Troicke, 1982), and especially genre analysis, above all the work of John Swales (1990). Swales (1990: 58) defines genres as 'a class of communicative events, the members of which share some set of communicative purposes'. In our analysis, we focus on two conversational

genres, telephone conversation and television interviews. The genres differ in the channel (telephone versus personal contact), degree of spontaneity (more versus less spontaneous) and privacy (private versus public conversations). The first genre involves interaction between strangers in economic transaction, the second genre is an informative and entertaining TV show.

Data and analytical procedure

The data used for the analysis include natural conversations in Slovenian language of approximately 30,000 tokens in length for each genre.

For the analysis, telephone conversations between a customer and an agent in a travel agency, tourist office and hotel reception were selected from the Turdis corpus (Verdonik and Rojc, 2006). The selection was named Turdis-2; it includes 65 conversations of 214.49 minutes in length (32,547 tokens – 19,936 in conversations with a travel agency, 6350 in conversations with a tourist office and 6261 in conversations with a hotel reception). The number of speakers is 68; 27 are male and 41 are female; 31 are travel agents and 37 are callers.

The television interviews from late night broadcast news shows used in the analysis were selected from the BNSI Broadcast News database (Žgank et al., 2004) and named BNSIint. The selection includes 30 conversations of 198.35 minutes (31,236 tokens) in length. The number of speakers is 47; 36 are male and 11 are female; six are interviewers and 41 are interviewees.

In both corpora used in the analysis, we have a similar average number of tokens per conversation for both genders, but important differences for the discursive roles: the interviewee has approximately twice as many tokens per interview (552 on average) as the interviewer (250 on average), and the travel agent approximately twice as many tokens (650 on average) as the caller (335 on average).

The analytical procedure is based on corpus analysis, combined with qualitative analysis. It consists of four steps. First, one annotator (a co-author of Verdonik et al., 2007a) manually annotated the discourse markers in both corpora, following the theoretical framework, based on the work of Redeker (1990) and Schiffrin (1987), as presented in the introductory section, and practical guidelines for annotation as presented in Verdonik et al. (2007a). In short, these guidelines specify that discourse markers are

expressions in conversation that contribute least to the ideational plane/conceptual meaning/propositional content, and mostly have pragmatic functions:

- they help signal connections to the propositional content,
- they help build a relationship between the participants in a conversation,
- they help express the speaker's attitude to the content of the conversation,
- they help organize the course of a conversation. (Verdonik et al., 2007a)

Second, since the annotation of discourse markers is based on the subjective interpretation of the annotator, an attempt was made to evaluate the validity of corpus annotation by tagging validation. We chose approximately

10 percent of the data from each corpus and had two independent annotators annotate this data again, following the annotation scheme presented in Verdonik et al. (2007a). The annotators had instructions not to discuss any open issues. The results of the tagging validation are presented in the next section. The comparison was limited to the number of occurrences for each discourse marker, as noted by each annotator, while the issue of how frequently the annotators agreed on whether a specific token or phrase was a discourse marker or not was disregarded. This is because subsequent analysis was restricted to the total frequency of discourse markers in the two corpora in the following steps, while particular instances of discourse markers were not analysed.

The third step was corpus analysis: it involved making a list of expressions functioning as discourse markers and establishing their frequency when used as a discourse marker for each of the genres. The differences between the genres as reflected by the results are processed in the last step of the analytical procedure.

In the last and the most important step, we interpret corpus analysis results by defining the context-discourse marker use relationships, typical of each genre. As Biber et al. (1998) point out, an important benefit of corpus analysis is that it enables generalizable findings that hold across texts. However, this is done at the price of disregarding local discourse features. As our analysis of context-discourse marker use relationships builds on corpus results, it is done on the macro-level, that is, we use generalizations and focus on what is common to all discourses of a genre, disregarding local context details. We are aware that in this way we overlook some micro-context patterns that may occur across the genres, however, in this we follow Van Dijk's line of thinking that language users typically access and use contexts primarily at the overall macro-level and disregard micro-contexts:

Language users therefore will typically abstract (micro)details of the context and infer overall, macro-concepts describing the ongoing (and ongoingly relevant) aspects of the context, such as the macro speech act now being engaged in, the ongoing (sub-)genre, overall setting information (approximate time and place), main actors, and the overall goal of conversation. This will allow them to disregard local context details during the rest of the discourse, and focus to what is important. (Van Dijk, 1997: 194–5)

We therefore consider macro-concepts describing the ongoing aspects of the context (i.e. macro-contexts) and micro-details of the context (i.e. micro-contexts) to be distinct levels of context models. While macro-contexts should be explored from a broad and distant perspective, we need a very close and detailed perspective in order to explore micro-contexts. Consequently, we need different methods (and different studies) for exploring micro-contexts than for exploring macro-contexts.

The last step of analysis is structured as following. First, the most typical meanings and functions of the discourse markers analysed here, as defined in Verdonik et al. (2007a) and Verdonik (2007), are summed up. Taking into account the meanings and functions, as well as the frequency of use of particular

discourse markers in each corpus, determined in the previous step of the analysis, we draw conclusions about some of the characteristics of the two genres (see Verdonik et al., 2007b). Finally, we search for contextual factors which contribute to these characteristics. We use Van Dijk's (1997) elements of context models (assigned relevance and importance; the roles, positions and social relations of participants; goals, intentions, purposes; group membership and social affiliation; personal association and identification with social groups; social attitudes and ideologies; setting; and others) as guidelines for describing all the macro-context factors of the two genres. After specifying the macro-contexts of each genre, we define only those macro-context factors that contribute to the characteristics of the two genres specified above. These are the factors which may influence discourse marker (non-)use.

We present the results of each step of the analytical procedure in the following sections. The main focus of our attention is, of course, step four of the analytical procedure, but first we briefly present the results of the second and third steps of the analytical procedure.

Results of tagging validation

The data for tagging validation consist of eight telephone conversations from the Turdis-2 Corpus of 22.63 minutes or 3181 tokens in length, and of three television interviews from the BNSIint Corpus of 9.39 minutes or 3157 tokens in length.

The tagging validation showed that the three annotators agreed that the expressions *aha* 'oh',² *aja* 'I see', *eee* 'um' and *no* 'well' (see the following section, for a detailed list of discourse markers and their presentation in this article) always function as discourse markers and that further evaluation of annotation was not necessary. For the expression *dobro* . . . 'right', the annotation was also the same in all three versions. Differences in annotation of less than 5.0 percent were observed for the expressions *ja* 'yes' and *glejte* 'look'. More pronounced differences appeared for the expression *ne?* 'right?' in the BNSI interviews – even over ± 20.0 percent, while they only amounted to ± 2.3 percent in the Turdis-2 Corpus. However, this is partly due to the low frequency of *ne?* 'right?' in the BNSIint test corpus (fewer than 20 instances). The situation is similar for the expression *zdaj* 'now' in the Turdis-2 Corpus – differences in annotation amounted to more than ± 15.0 percent, while the expression was used in fewer than 20 instances in the test corpus. All in all, we can conclude that both corpora were reliably annotated, with some deviations for the discourse marker *zdaj* 'now' in the Turdis-2 Corpus and *ne?* 'right?' in the BNSIint Corpus.

Test annotation also showed that when it comes to new expressions, not yet defined as discourse markers in Slovenian, there was least agreement among the three annotators. For example, in the Turdis-2 test corpus, the annotator of the entire corpus tagged 1 example from the 'residue' category (see the 'Results of corpus analysis' section), test annotator 1 tagged 14 examples of eight different expressions and test annotator 2 tagged 23 examples of eight different

expressions. This further confirms the difficulties to uniformly define discourse markers and opens an interesting question which might be a starting point for further research on Slovenian discourse markers.

Results of corpus analysis

In the third step of the analytical procedure, we compare the frequency of discourse marker use in the two corpora. The list of expressions functioning as discourse markers and their frequency are presented in Table 1. Since the number of tokens in each corpus was not completely the same, we normalized the frequency to 10,000 tokens.

Different expressions or different forms of a single expression (the Slovenian language is highly inflectional) can have the same or very similar meanings and functions. In this article we build on a previous detailed study of the functions of discourse markers, where we combined a quantitative and a qualitative method, so that all the uses of a particular discourse marker were annotated, each concordance and its co-text was examined and similar types of usage were grouped together (see Verdonik, 2006, 2007; Verdonik et al., 2007a). In Verdonik et al. (2007a), we showed that in some cases even very diverse expressions (e.g. different phonetic realizations, morphological variants of a single stem and even different stems) may have very similar meanings and functions. Of course, there are differences as well, for example, one form is more formal, another more informal, but since we lack the space to present each discourse marker in detail, we discuss the results for these groups of discourse markers together. Thus, in this article, the following sets of discourse markers are discussed together:

- *dobro . . .* ‘right . . .’³ stands for *dobro, v redu, okej, prav,*
- *eee . . .* ‘um . . .’ stands for *eee, eeem, eeen, nnn, mmm, eeeh* and similar fillers,
- *glejte* ‘look’ stands for *glejte, poglejte, lejte, glej, lej,*
- *ne?* ‘right?’ stands for *ne?, a ne?, ali ne?*
- *veste* ‘you know’ stands for *veste, a veste, morate vedeti,*
- background (also backchannel) signals include *aha, aja* ‘oh, I see’, *dobro* ‘okay’, *ja* ‘I see’, *mhm* ‘mhm’, *okej* ‘okay’, *seveda* ‘of course’, *tako* ‘that’s right’, *tudi* ‘that too’.
- ‘interjections’ stands for expressions such as *uuu, aaa, fff, ah, eh, hm, hja, evo, jo, joj, ma* and others;
- ‘residue’ stands for some types of usage of certain connectives, such as *seveda* ‘of course’, *torej* ‘so’, *in* ‘and’, pronouns, such as *kaj* ‘somewhat’, *kako* ‘how’, particles, such as *kajne* ‘right?’, *recimo* ‘let’s say’, adverbs, such as *pol* ‘then’, which appear quite frequently, but very seldom as discourse markers.

The categories of discourse markers which we labelled ‘interjections’ and ‘residue’ were not part of subsequent analysis since the discourse markers in these categories were used very infrequently (mostly once or twice) and it was therefore impossible to draw any conclusive information about their discourse functions.

TABLE 1. *The list of expressions potentially used as discourse markers and their frequency when used as a discourse marker, calculated per 10,000 tokens*

<i>Discourse marker</i>	<i>Turdis-2</i>	<i>BNSInt</i>
<i>aha</i> 'oh'	72	0
<i>mhm</i> 'mhm'	30	2
<i>ja</i> 'yes'	185	22
background signals	321	11
<i>ne?</i> 'right?'	186	16
<i>Dobro . . .</i> 'right . . .'	69	14
<i>no</i> 'well'	28	35
<i>glejte</i> 'look'	24	15
<i>veste</i> 'you know'	8	2
<i>zdaj</i> 'now'	64	1
<i>Eee . . .</i> 'um . . .'	387	413
<i>mislím</i> 'I mean'	7	1
Total	1381	532
'interjections'	9	2
'residue'	7	5

A comparison of the frequency of discourse marker use presented in Table 1 reveals important differences between the two genres: discourse markers in general are far more frequently used in telephone conversations than in television interviews, with two exceptions: the discourse markers *eee . . .* 'um . . .' and *no* 'well'. This is the starting point for our main analytical step which will lead us to the answer to the central question of this article: which contextual factors contribute to these differences in the two genres.

Analysis of context to discourse marker use relationships

We analysed the context to discourse marker use relationships for each discourse marker separately, following the procedure defined in step four of the 'Data and analytical procedure' section. The results of the analysis show that the contact between the participants in conversation is less close and intimate, that there are fewer expressions of agreement and understanding, fewer signals for turn-change, fewer expressions of emotions, and fewer signals when to end a conversation in television interviews than in telephone conversations.

Below we present the analysis and its results in more details. First, the results of the corpus analysis are briefly summed up. Second, the meanings and pragmatic functions of each discourse marker, as defined in Verdonik (2007) and Verdonik et al. (2007a), are presented briefly. For each discourse marker, a brief overview of the main functions identified by previous research is provided. On the basis of these functions, the characteristics of each genre which manifest themselves through the frequency of use of particular discourse markers in each corpus are identified. When these results are similar for a set of discourse markers, we discuss that set together in order to avoid repetition. Therefore, we

discuss *aha* 'oh', *mhm* 'mhm', *ja* 'yes', background signals and *ne?* . . . 'right? . . .' together: they have the same or very similar pragmatic functions (see the 'Data and analytical procedure' section for the definition of these functions) and the differences in the frequency of use are also quite similar. Furthermore, we discuss *glejte* 'look' and *veste* 'you know' together as they are the only two verbs on our list. We also discuss *eee* . . . 'um . . .' and *mislim* 'I mean' together because of their similar pragmatic functions. In the second part of this section, we outline the contextual factors which we believe contribute to the characteristics of each genre we have identified and therefore also contribute to the use or non-use of particular discourse marker in each genre.

AHA 'OH', MHM 'MHM', JA 'YES', BACKGROUND SIGNALS, NE? . . . 'RIGHT?'

The discourse markers *aha* 'oh', *mhm* 'mhm', *ja* 'yes', background signals and *ne?* 'right?' are all used much more frequently in the Turdis-2 Corpus (from eight times to 29 times more frequently) than in the BNSInt Corpus. According to Verdonik (2007) and Verdonik et al. (2007a), the most significant characteristics of these discourse markers include the following: *ja* 'yes' and *mhm* 'mhm' express agreement and understanding, *aha* 'oh' expresses understanding, and in some cases also the speaker's emotions, especially his or her attitude towards the content of the conversation. All of these expressions are also used as background signals with more or less the same characteristics. *Ne?* 'right?' functions in the opposite direction; it is used to check whether the hearers understand or agree with what has been said. All the discourse markers in this group are connective elements. They also help to create an intimate, close contact between the participants in conversation. Moreover, they function as signals indicating when and where to change a turn: *ne?* 'right?' often signals that the speaker is ready to give the turn over to another participant, *ja* 'yes', *mhm* 'mhm', *aha* 'oh' are usually used at the beginning of a new turn. *Aha* 'oh' sometimes expresses emotions towards the content of the conversation (e.g. surprise).

The above findings suggest that there is much less need to express agreement and understanding, build close, intimate contact between the participants in conversation, signal the turns or express emotions in television interviews than in telephone conversations.

DOBRO . . . 'RIGHT . . .'

The discourse marker *dobro* . . . 'right . . .' is used more frequently in the Turdis-2 Corpus, but the difference in frequency is less pronounced (there are five times as many instances in the Turdis-2 Corpus as in the BNSInt Corpus) than for the previous set of discourse markers. In general, according to Verdonik (2007) and Verdonik et al. (2007a), the most significant pragmatic functions, as well as the contextual factors shaping the use of *dobro* . . . 'right . . .', as we will show below, are similar as for *ja* 'yes', *mhm* 'mhm', *aha* 'oh'. However, there is an important difference in the use of *dobro* . . . 'right . . .' between the two corpora: in the Turdis-2 Corpus, *dobro* . . . 'right . . .' is often used as the speaker's signal that he or she wants to end the conversation, while in the BNSInt Corpus, it generally signals topic change (and not the end of the conversation).

This suggests that there is less need to reach an agreement on when to end the conversation in television interviews than in telephone conversations.

NO 'WELL'

No 'well' is used more often in television interviews than in telephone conversations. According to Verdonik (2007) and Verdonik et al. (2007a), it is a connective element (it connects different topic slots); in many types usage, it expresses the speaker's attitude towards the topic of conversation; when used at the end of an utterance, it stresses the content of the utterance. The uses of no 'well' in the Turdis-2 Corpus are diverse: we find all the above mentioned functions. It is relatively infrequent compared to other discourse markers in the Turdis-2 Corpus. On the other hand, no 'well' is the second most frequent discourse marker in the BNSInt Corpus. Its use is more homogeneous than in the Turdis-2 Corpus: no 'well' is used chiefly at the beginning of a new turn when the speaker does not agree completely or does not corroborate the statement made by the previous speaker. It is also used as a connective element when the speaker starts a new topic.

The use of no 'well' in our data suggests that in television interviews the participants in a conversation tend not to agree with each other quite as much and as frequently as in telephone conversations.

GLEJTE 'LOOK' AND VESTE 'YOU KNOW'

Glejte 'look' and *veste* 'you know' are not very frequently used as discourse markers, *veste* 'you know' is especially infrequent in our data and we still lack the necessary knowledge about its pragmatic functions, therefore we focus only on *glejte* 'look'. According to Verdonik (2007) and Verdonik et al. (2007a), *glejte* 'look' is used approximately twice as frequently in telephone conversations as in television interviews. It is used to attract the hearers' attention, it signals that the speaker is about to explain something. It functions as a connective element, contributing to a closer, more informal relationship between the participants in a conversation.

The last function can probably account for most of the differences in the frequency of use. As we have already observed in connection with the use of the discourse markers *aha* 'oh', *mhm* 'mhm', *ja* 'yes', *ne?* 'right?' in our data, the contact between the participants is closer and more intense in telephone conversations than in television interviews.

EEE . . . 'UM . . .' AND MISLIM 'I MEAN'

Eee . . . 'um . . .' is the second discourse marker used slightly more frequently in television interviews than in telephone conversations (the first is no 'well'). In both corpora, *eee . . .* 'um . . .' is by far the most frequent discourse marker. According to Verdonik (2007) and Verdonik et al. (2007a), *eee . . .* 'um . . .' signals that a speaker has not finished his or her turn, but needs some time in the production process; when used to interrupt the speaker's turn, it signals that the participant uttering this discourse marker wishes to take over the turn, and it is often used at unexpected events in utterance structure (e.g. self-repair,

inserted utterance). On the other hand, *mislim* 'I mean' is one of the most infrequent discourse markers, therefore it is hard to come to definite conclusions about its usage, but its pragmatic functions are closer to those of *eee . . .* 'um . . .' than to the functions of the discourse markers discussed above.

The small difference in the frequency of use confirms to a certain extent the above observation that television interviews are less personal, private and more formal, since *eee . . .* 'um . . .' has no interpersonal or attitudinal functions. As far as the organization of discourse is concerned, we can conclude that the need for organizing the course of a conversation, that is, to gain time for the production processes, to keep the turn or to signal the need to take over the turn, arises only slightly more frequently in television interviews than in telephone conversations.

ZDAJ 'NOW'

The discourse marker *zdaj* 'now' was almost never used in television interviews, while it was fairly common in telephone conversations. According to Verdonik (2007) and Verdonik et al. (2007a), *zdaj* 'now' connects different topic slots; it indicates that the speaker is about to say something or to give the answer, but hesitates a bit thus gaining time for production. *Zdaj* 'now' warns other participants that there is something to follow and that the speaker has not finished his or her turn.

Surprisingly, in spite of the partly similar functions of *zdaj* 'now' and *eee . . .* 'um . . .', the results of the corpus analysis are very different. We believe that a further analysis of the discourse marker *zdaj* 'now' in Slovenian language would be necessary for a conclusive explanation. It seems that *zdaj* 'now' as discourse marker may be too informal and therefore less appropriate for use in formal television interviews.

To sum up, the use of the discourse markers analysed here shows the following:

- there are fewer expression of agreement and understanding;
- the contact between the participants in conversation is less close and intimate;
- there are fewer signals for turn-change;
- there are fewer expressions of emotions; and
- there are fewer signals when to end a conversation in television interviews than in telephone conversations.

Following the procedure described in the penultimate paragraph of the 'Data and analytical procedure' section, we identify different contextual factors which have contributed to the differences between the two genres we have observed: social relationships between the participants in conversation, participants in conversation (above all, the number of participants), discursive roles and relationships, topic of conversation, opinion on the topic, attitude towards the topic, general goal of conversation, goal of participants in conversation, and channel of communication. Consequently, these are also factors which contribute to differences in discourse marker (non)-use in each genre.

Bellow we describe the relationships between discourse marker use and the contextual factors specified above.

The interviewer's relationship towards the interviewee is generally provocative, and the interviewees, when there are two, usually present opposing points of view. They compete for the affection of the audience and each try to prove his or her own right. This does not encourage the use of expressions of agreement and mutual understanding in the BNSInt Corpus. In telephone conversations, the opposite is true: the agent (and, to some extent, the costumer as well) attempts to build a positive, friendly relationship with the hearer, and this, to some degree, encourages the use of expressions of close, intimate contact, mutual understanding and emotions. This brings us to the conclusion that *social relationships between the participants in conversation* influence discourse marker use.

In television interviews one of the participants is the audience, an outside observer of the conversation. This contributes to the more formal mode of conversation and prevents a closer, more intimate contact between the participants. In telephone conversations, on the other hand, there are only two participants, both actively exchanging turns. This brings us to the conclusion that *the number of participants in conversation and their discursive roles* influence discourse marker use.

In television interviews, the interviewer has a privileged role: he or she is responsible for manage the course of the conversation; he or she can decide who gets the next turn, what the topic will be and when the conversation will end without asking for the interviewee's approval. This results in the use of fewer signals for turn-change and fewer signals for ending a conversation. In telephone conversations, on the other hand, the discursive roles are more equal in status; there is more need to signal and negotiate turn-changing, topic-changing or conversation ending. This brings us to the conclusion that *discursive roles and relationships* influence discourse marker use.

In television interviews, the topics are daily political, economic and other events in society: this makes the conversations more formal and discourages the use of expressions which signal close, intimate relationships or the expression of emotions. In telephone conversations, on the other hand, the topics are connected with the caller's free time and are consequently more intimate. This brings us to the conclusion that *the topic of conversation* influences discourse marker use.

A particular characteristic of the television interviews in our data is that the interviewees, when there are two, have opposing views on the topic of conversation, as a consequence they tend not to affirm the statements made by the other participant. In telephone conversations, the situation is very different: the speakers strive to achieve mutual agreement with a similar opinion on the topic. This brings us to the conclusion that *the opinion on the topic* influences discourse marker use.

The interviewer in television interviews is expected to be objective towards the topic, therefore he or she rarely reveals whether or not he or she agrees with the interviewee's statements, or how he or she feels about the topic. In

telephone conversations, speakers are expected to have subjective attitudes towards the topic, for example, they may be enthusiastic about a place, but may dislike another location, etc., and it is common for them to express their attitudes or emotions. This brings us to the conclusion that *the attitude towards the topic* influences discourse marker use.

In television interviews, the general goal is to animate the audience by presenting a conflict and confronting two opponents, which creates an atmosphere of disagreement. The goal of an interviewee in a television interview is not only to provide information; he or she may also wish to gain the affection of the audience, and may therefore wish to have turns for him or herself, rather than signalling to the other interviewee to take over the turn. In telephone conversations, the general goal is to convince a customer to buy a product or to obtain the necessary information, therefore the agent strives to build a positive, friendly atmosphere of agreement. This brings us to the conclusion that *the general goal of conversation and the goal of participants in conversation* influence discourse marker use.

The telephone as a channel for communication in telephone conversations disables visual contact, therefore the only way to express attention and understanding is by voice (and not, for example, by head nodding or other mimics or motions). In television interviews, on the other hand, the active participants have personal contact. This brings us to the conclusion that *the channel of communication* influences discourse marker use.

Conclusion

In the introductory section of this article, we expressed the hypothesis that most of the differences in discourse marker use in television interviews and telephone conversations, the two genres analysed here, can be explained by the differences between the two genres which we have label contextual factors. The first stage of the analysis revealed important differences between the two genres: discourse markers are far more likely to be used in telephone conversations than in television interviews. In the second stage of the analysis, the following contextual factors, which contribute to the differences in the use of discourse markers, were identified:

- 1) Social relationships between the participants in conversation
- 2) Participants in conversation (above all, the number of participants)
- 3) Discursive roles and relationships
- 4) Topic of conversation
- 5) Opinion on the topic
- 6) Attitude towards the topic
- 7) General goal of conversation, goal of participants in conversation
- 8) Channel of communication

Since we tried to identify only those contextual factors which contribute to differences in discourse marker use in the genres analysed here, we must be aware of the fact that there may be other contextual factors which influence

discourse marker use, but have not emerged in this particular contrastive analysis. Because our analytical procedure builds on generalizations and common features of the genre, we also disregarded the micro-level environments, where some patterns may occur across the genre. A detailed analysis of the micro-level environment seems to be the logical next step in the analysis: while such an analysis is beyond the scope of the present study, it would shed more light on the micro-level contexts that influence discourse markers use as well.

The contextual factors we have identified are of different types: some can be classified as social (participants in conversation, their social relationships and discursive roles and relationships), others as psychological (opinion on the topic, attitude towards the topic, goals of conversation), or even as semantic (topic of conversation) and physical (channel of communication). The classification of contextual factors certainly needs further study.

In Van Dijk's (1997, 2007) understanding of contexts as a type of cognitive models, which we chose to follow here, context models are subjective interpretations and experiences. By analysing the characteristics and contexts of genres, we had to override the individual nature of context models (and conversations) and search for what is common, generalized and most typical in many conversations of the same genre. This holds for all levels of our analysis: for the description of the most typical functions of discourse markers, for the characteristics of the genres which manifest themselves through the use of discourse markers, and for the contextual factors we have identified which cause differences in the use of discourse markers. Genre analysis builds on the concept of the prototype, reasoning that 'examples of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience. If all high probability expectations are realized, the exemplar will be viewed as prototypical by the parent discourse community' (Swales, 1990: 58). Similarly, the contextual factors we have identified are prototypical of the genres used in the analysis. The question which arises here is whether it is possible to speak about prototype contexts within the framework of genre analysis. We believe it is.

In this article, we presented a study of context-to-talk relationships in two conversational genres. The study focuses on various aspects of language use, such as discourse markers (above all their use in Slovenian), genres, contexts, etc. However, the analysis presented here opens some important questions; here, we will mention only some of the most obvious: 1) Are the contextual factors identified here which contribute to differences in the use of discourse markers inter-connected and, if so, how and why? 2) Naturally, we are interested in other contextual factors which may influence discourse marker use, but were not identified in the present study. 3) Throughout the analysis, it was possible to detect certain characteristics of the genres analysed here merely through discourse marker use. A detailed study of these genres would reveal to what degree it is possible to describe genre characteristics on the basis of discourse marker use. 4) Last but not least, the idea that there are prototype contexts needs further analysis and discussion.

NOTES

1. This work was partially funded by Slovenian Research Agency, under contract number: J2-9742-0796-06.
2. The expressions in quotation marks are approximate English translations of the Slovenian discourse markers.
3. It is impossible to provide exact English equivalents for the Slovenian discourse markers examined in this article, because there are no one-to-one equivalents. A single discourse marker may have several functions in each of the two languages, but the functions in the two languages do not necessarily overlap. The translations provided here are therefore only informative, giving the general meaning of each discourse marker.

REFERENCES

- Andersen, E.S., Brizuela, M., DuPuy, B. and Gonnermas, L. (1999) 'Cross-linguistic Evidence for the Early Acquisition of Discourse Markers as Register Variables', *Journal of Pragmatics* 31: 1339–51.
- Archakis, A. (2001) 'On Discourse Markers: Evidence from Modern Greek', *Journal of Pragmatics* 33: 1235–61.
- Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Blakemore, D. (2002) *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Dedaia, M.N. (2005) 'Ironic Denial: *tabože* in Croatian Political Discourse', *Journal of Pragmatics* 37: 667–83.
- Fox Tree, J.E. (2006) 'Placing *Like* in Telling Stories', *Discourse Studies* 8(6): 723–43.
- Fraser, B. (1996) 'Pragmatic Markers', *Pragmatics* 6(2): 167–90.
- Fraser, B. (1999) 'What are Discourse Markers?', *Journal of Pragmatics* 31: 931–52.
- Fukushima, T. (2005) 'Japanese Continuative Conjunction *ga* as a Semantic Boundary Marker', *Journal of Pragmatics* 25: 81–106.
- Gonzalez, M. (2005) 'Pragmatic Markers and Discourse Coherence in English and Catalan Oral Narrative', *Discourse Studies* 7(1): 53–86.
- Hymes, D. (1977) *Foundations in Sociolinguistics: An Ethnographic Approach*. London: Tavistock Publications.
- Jucker, A.H. and Ziv, Y. (eds) (1998) *Discourse Markers: Descriptions and Theory*. Amsterdam: John Benjamins.
- Kroon, C. (1998) 'A Framework for the Description of Latin Discourse Markers', *Journal of Pragmatics* 30: 205–23.
- Matsui, T. (2001) 'Semantics and Pragmatics of a Japanese Discourse Marker *dakara* (so/in other words): A Unitary Account', *Journal of Pragmatics* 34: 867–91.
- Montes, R.G. (1999) 'The Development of Discourse Markers in Spanish: Intejctions', *Journal of Pragmatics* 31: 1289–319.
- Pisanski Peterlin, A. (2005) 'Text-organising Metatext in Research Articles: An English-Slovene Contrastive Analysis', *English for Specific Purposes* 25: 307–19.
- Redeker, G. (1990) 'Ideational and Pragmatic Markers of Discourse Structure', *Journal of Pragmatics* 14: 367–81.
- Saville-Troike, M. (1982) *The Ethnography of Communication: An Introduction*. Oxford: Blackwell.
- Schiffrin, D. (1987) *Discourse Markers*. Cambridge: Cambridge University Press.

- Schlamberger Brezar, M. (2007) 'Vloga povezovalcev v govornem diskurzu', *Jezik in slovnost* 52(3–4): 21–32.
- Schourup, L. (1999) 'Discourse Markers', *Lingua* 107: 227–65.
- Schourup, L. (2001) 'Rethinking Well', *Journal of Pragmatics* 33: 1025–60.
- Smolej, M. (2004) 'Členki kot besedilni povezovalci', *Jezik in slovnost* 49(5): 45–57.
- Swales, J.M. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Tchizmarova, I.K. (2005) 'Hedging Functions of the Bulgarian Discourse Marker *xajde*', *Journal of Pragmatics* 37: 1143–63.
- Van Dijk, T.A. (1997) 'Cognitive Context Models and Discourse', in M. Stamenov (ed.) *Language Structure, Discourse and the Access to Consciousness*, pp. 189–226. Amsterdam: Benjamin.
- Van Dijk, T.A. (2001) 'Discourse, Ideology and Context', *Folia Linguistica* 30(1–2): 11–40.
- Van Dijk, T.A. (2007) 'Discourse, Context and Cognition', *Discourse Studies* 8(1): 159–77.
- Verdonik, D. (2006) 'Mhm, ja, no, dobro, glejte, eee . . . : diskurzni označevalci v telefonskih pogovorih', *Jezik in slovnost* 51(2): 19–36.
- Verdonik, D. (2007) *Jezikovni elementi spontanosti v pogovoru: Diskurzni označevalci in popravljanja*. Maribor: Slavistično društvo Maribor.
- Verdonik, D. and Rojc, M. (2006) 'Are You Ready for a Call? – Spontaneous Conversations in Tourism for Speech-to-speech Translation Systems', in *Proceedings of 5th International Conference on Language Resources and Evaluation, Genoa, Italy*. Paris: European Language Resources Association.
- Verdonik, D., Pisanski Peterlin, A. and Žgank, A. (2007b) 'Diskurzni označevalci v dveh pogovornih žanrih', *Jezik in slovnost* 52(6): 19–33.
- Verdonik, D., Rojc, M. and Stabej, M. (2007a) 'Annotating Discourse Markers in Spontaneous Speech Corpora on an Example for the Slovenian Language', *Language Resources and Evaluation Journal* 41(2): 147–80.
- Vlemings, J. (2003) 'The Discourse Use of French *donc* in Imperative Sentences', *Journal of Pragmatics* 35: 1095–112.
- Wang, Y.F. and Tsai, P.H. (2007) 'Textual and Contextual Contrast Connection: A Study of Chinese Contrastive Markers across Different Text Types', *Journal of Pragmatics* 39: 1775–815.
- Wang, Y.F., Tsai, P.H. and Ling, M.Y. (2007) 'From Informational to Emotive Use: *meiyou* ("No") as a Discourse Marker in Taiwan Mandarin Conversation', *Discourse Studies* 9(5): 677–701.
- Waring, H.Z. (2003) '"Also" as a Discourse Marker: Its Use in Disjunctive and Disaffiliative Environments', *Discourse Studies* 5(3): 415–36.
- Žgank, A., Rotovnik, T., Sepesy Maučec, M., Verdonik, D., Kitak, J., Vlaj, D., Hozjan, V., Kačič, Z. and Horvat, B. (2004) 'Acquisition and Annotation of Slovenian Broadcast News Database', in *Proceedings of 4th International Conference on Language Resources and Evaluation, Lisbon, Portugal*, pp. 2103–2106. Paris: European Language Resources Association.



DARINKA VERDONIK is researcher at the Centre for Language Technologies at the Faculty of Electrical Engineering and Computer Science, University of Maribor. Her research interests include language resources and discourse analysis. She received her

PhD in 2006 from the Faculty of Arts, University of Ljubljana; the topic of her PhD thesis was pragmatic annotation of corpora for the needs of speech-to-speech translation technologies. ADDRESS: Faculty of Electrical Engineering and Computer Science, Smetanova 17, SI-2000 Maribor, Slovenia. [email: darinka.verdonik@uni-mb.si]

ANDREJ ŽGANK received his BSc degree from the Faculty of Electrical Engineering and Computer Science at the University of Maribor in 1999 and his PhD degree in 2003 from the same institution. Since 2003 he has been a Senior Researcher at the University of Maribor. His research interests include multilingual and cross-lingual speech recognition, acoustic modelling for large vocabulary continuous speech recognition, and the creation of language resources. ADDRESS: Faculty of Electrical Engineering and Computer Science, Smetanova 17, SI-2000 Maribor, Slovenia. [email: andrej.zgank@uni-mb.si]

AGNES PISANSKI PETERLIN works at the Department of Translation, Faculty of Arts, University of Ljubljana. In 2006, she received her PhD in linguistics from the University of Ljubljana. Her research interests include contrastive rhetoric, translation of texts for specific purposes and discourse analysis. ADDRESS: Department of Translation, Faculty of Arts, University of Ljubljana, Askerceva 2, SI-1000 Ljubljana, Slovenia. [email: agnes.pisanski@guest.arnes.si]

