

## Combinatory potential of contrastive discourse markers in English and Lithuanian: a semantic functional analysis

Maslauskienė, Greta

Combinatory potential of contrastive discourse markers in English and Lithuanian: a semantic functional analysis

Lietuvių kalba, núm. 14, 2020

Vilniaus Universitetas, Lituania

Disponible en: <https://www.redalyc.org/articulo.oa?id=692574286008>



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.

# Combinatory potential of contrastive discourse markers in English and Lithuanian: a semantic functional analysis

Greta Maslauskienė greta.maslauskiene@flf.vu.lt  
Vilnius University, Lithuania

## 1 Introduction

Discourse markers (henceforth, DMs) have been under linguistic scrutiny for more than three decades. They have been extensively studied in English (Fraser 1990; 2006; 2009; 2013; Haselow 2012; Aijmer & Elgemark 2013; Traugott 2014; *inter alia*), Italian (Waltereit 2002), French (Fagard 2010), Japanese (Shinzato 2017), as well as across languages, such as English and Dutch (Van Olmen 2011; 2013), English, Swedish and German (Aijmer 2007), English and Hebrew (Maschler 1994), Italian and Latin (Fedriani & Ghezzi 2014). Despite the abundance of synchronic and diachronic studies on DMs, there is still a lack of consensus as to how these linguistic devices should be defined, categorized and described in terms of their syntactic, semantic and pragmatic properties.

A plethora of terms is used to refer to DMs in the studies written in English, namely, *cue phrases* (Knot & Sanders 1998), *discourse connectives* (Blakemore 1987; Hall 2007), *discourse operators* (Redeker 1991), *discourse particles* (Schourup 1999). The term *discourse markers* has become widely used (see, for example, Schiffrin 1987; Fraser 1999; 2006; 2013 and Blakemore 2002) and frequently regarded as a theory-neutral term (Furko 2005, 19). For the purposes of this study, the aforementioned term *discourse markers* is adopted.

The syntactic, semantic and pragmatic properties of DMs have been extensively investigated in English from the synchronic and diachronic perspectives. Diachronic studies have shed light on the development of DMs, tracing back their sources (e.g. adverbs, imperatives). Some scholars account for the evolution of DMs in terms of *grammaticalization* (Brinton 1996; Traugott 2014; Lenker 2000), while others link their development to *pragmaticalization* (Aijmer 1997; Defour et al. 2010). The comparison of cross-linguistic diachronic data has led to the conclusion that the paths of development of DMs may differ across languages (see Defour et al. 2010). Synchronic studies have focused on the semantic-pragmatic profile of DMs and their syntactic behaviour. Areas of inquiry into DMs include but are not limited to discourse structuring,

Lietuvių kalba, núm. 14, 2020

Vilniaus Universitetas, Lithuania

Recepción: 02 Marzo 2020

Aprobación: 10 Abril 2020

Redalyc: <https://www.redalyc.org/articulo.oa?id=692574286008>

pragmatic functions and their (inter)subjective meanings (see Brinton 2008). Recent studies have shown that DMs display a variety of textual and interpersonal functions determined by the syntactic position of DMs (see, for example, Traugott 2014).

The potential of DMs to form DM combinations has led to new perspectives in the analysis of their semantic-pragmatic properties and grammaticalization processes involved in their development. As some studies suggest, not only syntactic position, but also their combinability potential may have the impact on the semantic-pragmatic functioning of DMs. For example, Aijmer (2016) convincingly demonstrates that *anyway* displays functional differences when it occurs as a stand-alone DM and in combinations. Fraser (2013) has identified possible combinations of English DMs that signal contrast and has categorized them according to their functional profile.

Researchers provide different accounts of DMs as far as their categorial status is concerned. Focusing on spoken discourse, Schiffrin (1987) includes interjections such as *oh* and *now*, as well as non-verbal expressions into the category of DMs, which make up a diverse group. Fraser (2009, 297), on the other hand, regards DMs as a purely linguistic phenomenon, as the scholar defines DMs as lexical expressions, such as *but*, *so* and *in addition* that signal specific semantic relationship between discourse segments. These extra-clausal linguistic units contribute to the interpretation of the whole utterance, for they guide the addressee as to how to interpret the utterance next to a DM (Fraser 1999). Similarly, Aijmer (2007, 33) refers to DMs in terms of “linguistically encoded clues which signal the speaker’s potential communicative intentions.” In *Longman Grammar of Spoken and Written English* (Biber et al. 1999, 875), the linguistic items described as DMs in Fraser (1996; 2009) are analysed under the category of *linking adverbials*. Moreover, the term *discourse marker* is restricted to a small number of inserts (*well*, *right*, *now*, *I mean*, *you know*, *you see*, *mind you* and *now then*) that signal either transition in a conversation, or an interactive relationship between the speaker, the hearer, and the message.

Terminological variety is also observed in the Lithuanian literature on the linguistic items defined as DMs in this study. DMs are referred to as *diskurso žymikliai* ‘discourse markers’ (Masaitienė 2003; Buitkienė 2011; Šimčikaitė 2012), *diskurso žymekliai* ‘discourse markers’ (Pugačiauskaitė 2012), *diskurso markeriai* ‘discourse markers’ (Usonienė 2016), *diskurso jungtukai* ‘conjunctive discourse markers’ (Verikaitė 2005), *diskurso jungtukai ir jungiamieji žodžiai* ‘discourse connectives’ (Bielinskienė 2009). The category of DMs is not mentioned in *Lithuanian Grammar* (Ambrazas et al. 1997) at all, and items that fall under the category of DMs are classified in this grammar as *coordinating conjunctions* (Ambrazas et al. 1997, 427). Mention of semantic bleaching and the emergence of additional pragmatic functions of some expressions that would qualify as DMs is made with reference to *interpolation* (Ambrazas et al. 1997, 685-689).

The investigation of the syntactic, semantic and pragmatic properties of Lithuanian DMs is sparse, even though some research on individual DMs (Vaičiūnienė & Mažeikienė 2016) or their classes (Buitkienė 2011; Jasionytė-Mikučionienė 2016; Valančė 2017) has been emerging. Contrastive studies have focused on the textual and interactional functions of individual English and Lithuanian DMs in spoken discourse (Masaitienė 2003) and in fiction (Buitkienė 2011). Other linguistic analyses have explored the semantic-pragmatic profiles of DMs in relation to their syntactic position. For instance, drawing on the Relevance theory, Bielinškienė (2009) has analysed the multifunctionality of certain DMs in the left periphery of a sentence and, in particular, has revealed the multifunctionality of *ir* ‘and’, *ar’or* ‘and’ and *o* ‘but’. Focusing on the occurrences of DMs in the left and right peripheries of a clause as well as in medial position, Jasionytė-Mikučionienė (2016) has successfully demonstrated that the (inter)subjective meanings of the DMs *žiūrėk* ‘look’ and *klausyk* ‘listen’ are dependent on their syntactic position and the communicative type of the host sentence.

Although a number of studies have examined the semantic-pragmatic profile of DMs and the correlation between their functions and syntactic status in both English and Lithuanian, the combinatory potential of DMs has been scarcely explored, especially from a cross-linguistic perspective. This paper aims at investigating the combinatory potential of contrastive DMs (henceforth, CDMs) in English and Lithuanian, paying special attention to their semantic-pragmatic profile across discourse types. The study largely draws on the definition of DMs proposed by Fraser (1996; 1999; 2009), considered as one of the most comprehensive accounts of DMs, later adopted by Aijmer (2007) and Heine (2013).

## 2 Data and methods

This contrastive study based on the data obtained from English and Lithuanian corpora explores semantic-functional features of CDMs and their combinations in English and Lithuanian academic discourse, spoken discourse and fiction. The study draws on the classification of English CDMs provided in Fraser (2013) (*but, however, yet, still, nevertheless, on the other hand, instead, conversely, in contrast, on the contrary, alternatively*) and their most common Lithuanian translation correspondences (. ‘but’, *bet* ‘but’, *tačiau* ‘however’, *vis dėlto* ‘nevertheless’, *vis tiek/vis tik* ‘still’, *priešingai* ‘on the contrary’, *užtat* ‘instead’, *kita vertus/antra vertus* ‘on the other hand’, *nepaisant to* ‘nevertheless’, *iš kitos pusės* ‘on the other hand’, *bet kokiu (kuriuo) atveju* ‘nevertheless’, *vietoj(e) to* ‘instead’).

Following Fraser (2013, 322), CDMs are classified into General Contrastive (henceforth, GC) and Specific Contrastive (henceforth, SC) DMs. The most frequent English CDMs (*but, however, yet, still, nevertheless*) are considered as GCs as they display the most general semantic-pragmatic profile. On the contrary, the other CDMs (*on the other hand, instead, conversely, in contrast, rather, on the contrary,*

*alternatively*) are defined as SCs because they signal very specific types of contrast and their semantic-pragmatic profile is much more restricted. In Lithuanian translation, English GCs mainly correspond to *bet* 'but', *tačiau* 'however', . 'but', *vis dėlto* 'nevertheless', and *vis tiek/ vis tik* 'still', which also occur among the most frequent Lithuanian CDMs and display diverse semantic- functional properties. Conversely, such markers as *priešingai* 'on the contrary' and *kita vertus* 'on the other hand', which dominate among the translation correspondences of SCs, such as *on the other hand, instead, conversely, in contrast, rather, on the contrary, alternatively*, have a more restricted distribution and patterns of use (Belogurova 2018). Hence in this study *bet* 'but', *tačiau* 'however', *o* 'but', *vis dėlto* 'nevertheless', and *vis tiek/ vis tik* 'still' are classified within the group of GCs, whereas *priešingai* 'on the contrary', *užtat* 'instead', *kita vertus/ antra vertus* 'on the other hand', *nepaisant to* 'nevertheless', *iš kitos pusės* 'on the other hand', *bet kokiū (kuriuo) atveju* 'nevertheless', *vietoj(e) to* 'instead' are categorized as SCs.

A combination of DMs is defined as a pair of DMs occurring together and signalling the relation between the information in a discourse segment that precedes that DM combination (henceforth, S1) and the discourse segment that follows the combination (henceforth, S2), for example:

- (1) <...> I don't much like the distinction between good literature and pulp literature but, on the other hand, I do accept that it has a function. (BNC-spok)

I chose to consider concordances with CDMs used in separation, if alterations in word order did not change the meaning of a sentence, as in (2a-2b):

- (2) a. It wasn't er very widely used, but it nevertheless er had some important applications. (BNC- spok)  
 b. It wasn't er very widely used, [but nevertheless] it er had some important applications.

The example above shows that the movement of the subject in (2b) does not affect the meaning of the sentence, thus both instances (2a-2b) contain CDM combinations.

Following Fraser (2013), I treated a sequence of CDMs as a combination only if each of the CDMs in that combination could occur as a stand-alone marker in the same sentence, for example:

- (3) a. Some people try to introduce toys that do not reinforce gender stereotypes but, nevertheless, girls are much more likely to get dolls to dress up in the latest fashion, boys more likely to get toys for military-type games. (BNC-acad)  
 b. Some people try to introduce toys that do not reinforce gender stereotypes but/nevertheless, girls are much more likely to get dolls to dress up <...>.

As can be seen, both *but* and *nevertheless* could be used in the same sentence as stand-alone markers without causing ungrammaticality.

Three corpora were used in the study. Written Lithuanian data was selected from the sub-corpus of fiction in the Corpus of the Contemporary Lithuanian Language (CCLL-fict). The specialized

synchronic corpus of written academic Lithuanian CorALit (Corpus of Academic Lithuanian) served as a source for written academic language. Spoken Lithuanian was extracted from the sub-corpus of spoken discourse in the CCLL (CCLL-spok). Data for written and spoken English was retrieved from the British National Corpus (BNC), namely the sub-corpora of fiction, academic and spoken discourse. The size of each data source employed for the identification of the combinatory potential of CDMs in English and Lithuanian is provided in Table 1.

**Table 1**

Size of the subcorpora used for the investigation of the combinatory potential of English and Lithuanian CDMs

<b>English</b>	<b>Lithuanian</b>
BNC-acad (15,331,668)	CorALit-acad (8,670,613)
BNC-fict (15,909,312)	CCLL-fict (18,461,597)
BNC-spok (9,963,663)	CCLL-spok (557,822)

The frequencies of CDM combinations and their categories were normalised per 1,000,000 words. The search for CDM combinations in the data sources was conducted manually. For the qualitative analysis, each concordance was analysed in terms of semantic-functional properties of the CDM combinations.

### 3 CDM Combinations across discourse types in English and Lithuanian

The quantitative analysis reveals 1,928 Lithuanian and 2,178 English occurrences of CDM combinations (4,106 in total) in the corpora.

The normalized frequencies of CDM combinations (see Table 2 below) show their predominant use in the Lithuanian language:



Table 2

Normalized frequency f1000000 and distribution of CDM combinations across types of discourse in English and Lithuanian

English		BNC-acad	BNC-fict	BNC-spok	Total
	GC-GC	19.31	33.45	28.7	81.46
	GC-SC	44.16	16.15	11.23	71.54
	SC-GC	1.05	0.12	0	1.17
	SC-SC	0	0	0	0
	Total	64.52	49.72	39.93	154.17
Lithuanian		CorALit	CCLL-fict	CCLL-spok	Total
	GC-GC	1.3	73.47	145.24	220.01
	GC-SC	3.38	18.85	55.61	77.84
	SC-GC	0	0	0	0
	SC-SC	0	0.05	0	0.05
	Total	4.68	92.37	200.85	297.90

The overall normalized frequency of the Lithuanian CDM combinations in academic discourse, spoken discourse and in fiction is approximately two times higher than the frequency of English CDM combinations. Lithuanian CDMs predominantly occur in spoken discourse and the least amount of CDM combinations is detected in academic discourse. On the contrary, the opposite situation is observed in English, as CDM combinations mostly occur in academic discourse, whereas spoken discourse contains the smallest amount of CDM combinations.

The two languages under scrutiny differ mostly in the use of GC-GC and GC-SC combinations. In English, the predominant use of GC-GC is observed in fiction and GC-SC combinations are mostly used in academic discourse. On the other hand, Lithuanian GC-GC and GC-SC combinations are mainly detected in spoken discourse. Furthermore, SC-GC combinations have only been found in the English corpora. While one SC-SC combination has been identified in Lithuanian fiction, there are no English SC-SC combinations detected in the English data.

Following Fraser (2013), the CDM combinations found in the English and Lithuanian data were organized in a form of matrices, which illustrate the co-occurrences of each English and Lithuanian CDM under scrutiny with another CDM (see Tables 3-4). The vertical column on the left refers to the CDMs that go first in a combination, while the horizontal row on the top represents items that go second in a CDM combination. The dark grey cells indicate that there is a significant amount (more than 5/1,000,000) of instances of the combination use, whereas the light grey cells mark infrequent cases (less than 5/1,000,000). The white cells show that there are no combinations detected in the corpora. The black cells indicate combinations consisting of identical DMs (e.g. *but-but*), which were not considered in the investigation. Finally, the first five CDMs are set apart from the following 6-12 CDMs by heavy black lines to distinguish the contrast between GCs and SCs.

**Table 3**  
The matrix of CDM combinations in English

GC						SC					
1	2	3	4	5	6	7	8	9	10	11	12
GC	but	however	yet	still	nevertheless	on the other hand	instead	conversely	in contrast	rather on the contrary	alternatively
1	but										
2	however										
3	yet										
4	still										
5	nevertheless										
6	on the other hand										
7	instead										
8	conversely										
9	in contrast										
10	rather										
11	on the contrary										
12	alternatively										

**Table 3**  
The matrix of CDM combinations in English

GC						SC					
1	2	3	4	5	6	7	8	9	10	11	12
GC	Importar importar o 'but'	bet 'but'	tačiau 'however'	vis dėlto 'nevertheless'	vis tik/vis 'still'	priešingai 'on the contrary'	užtat 'instead'	kita vertus/antra 'on the other hand'	nepaisant to 'nevertheless'	iš kitos pusės 'on the other hand'	bet koki (kuriuo) atveju 'nevertheless' to 'instead'
1	o 'but'										
2	bet 'but'										
3	tačiau 'however'										
4	vis dėlto 'nevertheless'										
5	vis tik/vis 'still'										
6	priešingai 'on the contrary'										
7	užtat 'instead'										
8	kita vertus/antra vertus 'on the other hand'										
9	nepaisant to 'nevertheless'										
10	iš kitos pusės 'on the other hand'										
11	bet koki (kuriuo) atveju 'nevertheless'										
12	vietoj(e) to 'instead'										

The first feature emerging from Tables 3-4 is the ranging frequency and variety of CDM combinations in the English and Lithuanian data. The most frequent CDM combinations in Lithuanian are *bet-vis tiek/vis tik* 'but-still', *bet-vis dėlto* 'but-nevertheless', *bet-užtat* 'but-instead', *o-vis dėlto* 'but-nevertheless' and *bet-o* 'but-still'; and the most frequent CDM combinations in English include *but-still*, *but-rather*, *but-nevertheless*, *but-instead*, *but-on the other hand* (Table 5).



Table 5

Most frequent CDM combinations in the English and Lithuanian written and spoken corpora f1000000

English		Lithuanian	
CDM combination	Normalised frequency	CDM combination	Normalised frequency
<i>but-still</i>	48.65	<i>bet- vis tiek/ vis tik</i> 'but-still'	126.24
<i>but-rather</i>	41.24	<i>bet-vis dėlto</i> 'but-nevertheless'	36.41
<i>but-nevertheless</i>	20.52	<i>bet-užtat/užtai</i> 'but-nevertheless'	33.39
<i>but-instead</i>	15.97	<i>o-vis dėlto</i> 'but-nevertheless'	17.44
<i>but-on the other hand</i>	10.82	<i>bet-o</i> 'but-but'	12.55

*Bet* 'but'-CDM and *but*-CDM are not only the most frequent combinations (see Table 6), but also the most diverse, as *but* and *bet* combine with all the other CDMs.

Table 6. Most frequent groups of CDM combinations in the English and Lithuanian written and spoken corpora (f/1,000,000)

Table 6

Most frequent groups of CDM combinations in the English and Lithuanian written and spoken corpora f1000000

English		Lithuanian	
CDM combination	Normalised frequency	CDM combination	Normalised frequency
<i>but</i> -CDM	141.92	<i>bet</i> -CDM	234.37
<i>yet</i> -CDM	6.76	<i>o</i> -CDM	43.82
<i>however</i> -CDM	3.41	<i>tačiau</i> -CDM	15.31

Fraser (2013, 328-329) maintains that, unlike other CDMs, *but* never goes second in the combination, which might contribute to the abundance of *but*-CDM combinations in the English data. Alongside *but*-CDM combinations, this study provides illustration of *yet/however/still-but* combinations (see section 3.1). However, these findings are not conclusive as instances of *yet/however/still-but* combinations were not numerous (see Table 7).

### 3.1 General Contrastive-General Contrastive Combinations in English

Overall, thirteen GC-GC combinations have been found in the English data (see Table 7).

Table 7

Normalized frequency f1000000 and distribution of English GCGC combinations across discourse types

Combination	BNC-acad	BNC-fict	BNC-spok	Total
<i>but-still</i>	9.39	25.71	13.55	<b>48.65</b>
<i>but-nevertheless</i>	4.7	3.27	12.55	<b>20.52</b>
<i>yet-still</i>	1.3	3.33	0.5	<b>5.13</b>
<i>but-yet</i>	0.78	0.13	1.3	<b>2.21</b>
<i>however-still</i>	1.83	0.13	0	<b>1.96</b>
<i>however-but</i>	0.72	0.57	0.1	<b>1.39</b>
<i>nevertheless-still</i>	0.39	0.13	0	<b>0.52</b>
<i>but-however</i>	0	0	0.5	<b>0.5</b>
<i>still-nevertheless</i>	0.13	0	0.2	<b>0.33</b>
<i>yet-but</i>	0.07	0	0	<b>0.07</b>
<i>still-but</i>	0	0.06	0	<b>0.06</b>
<i>yet-nevertheless</i>	0	0.06	0	<b>0.06</b>
<b>Total</b>	<b>19.31</b>	<b>33.45</b>	<b>28.7</b>	<b>81.40</b>

The most frequent GC-GC combination in English is *but-still*, which is also the most frequent CDM combination in general. Another common GC-GC combination in the English data is *but- nevertheless*.

The distribution of GC-GC combinations across different types of discourse in English does not show as much variation as in Lithuanian (see section 3.2). English GC-GC combinations occur most frequently in fiction, whereas in spoken discourse they are slightly less frequent. Academic discourse, compared to fiction and spoken discourse, contains the lowest overall frequencies of English GC- GC combinations.

Another feature emerging from Table 7 is the abundant use of *but*-CDM combinations. It should be noted that *but* co-occurring with another CDM loses its specific functions and becomes semantically “bleached”. *But* signals a contrast between S1 and S2 but it does not specify the type of contrast, which is marked by the second CDM in the combination (Fraser 2013, 330). In my view, *however* and *yet* undergo such semantic weakening when they co-occur with *still*:

(4) But when discussing such an instance, one policeman expressed his dislike for blacks but still found the film 'brilliant'. (BNC-acad)

(5) She'd used him and then abandoned him, and yet still he'd continued to protect her. (BNC-fict)

(6) Agreement on economies to be proposed by Parliament was reached, the crisis was apparently over, and the members dispersed to their weekend retreats. The economies, however, still amounted to only 56 million. (BNC-acad)

As examples (4-6) show, the meaning of such combinations (in this case, concession) is determined by the meaning of *still*, as initial *but*, *yet* and *however* signal a contrast between S1 and S2 without further specifying the type of the contrast.

A close analysis of the concordance of *but-still*, *but-yet* and *yet-still* suggests (see examples (4-5)) that these combinations might signal that the message in S2 is a part of shared knowledge. In some cases, such combinations occur with emphatic *do* (see (7-8)) to strengthen the contrast.

(7) And erm there's lots and lots of erm poverty but yet the people do seem fairly content with their lot. (BNC-spok)

(8) Throughout Latin America, except for Cuba, abortion is illegal but, nevertheless, a considerable number of women do resort to abortion. (BNC-acad)

A frequent co-occurrence of emphatic *do* with a DM to strengthen concessive meaning has also been demonstrated in Szczyrbak (2017) with the case of *indeed*.

### 3.2 General Contrastive-General Contrastive Combinations in Lithuanian

In Lithuanian, eleven GC-GC combinations have been found, as illustrated in Table 8.

**Table 8**  
Normalized frequency f1000000 and distribution of the  
Lithuanian GCGC combinations across discourse types

Combination	CorALit	CCLL-fict	CCLL-spok	Total
<i>bet-vis tiek/ vis tik</i> 'but-still'	0.12	32.9	93.22	<b>126.24</b>
<i>bet-vis dėlto</i> 'but-nevertheless'	0	14.9	21.51	<b>36.41</b>
<i>o-vis dėlto</i> 'but-nevertheless'	0.36	11.7	5.38	<b>17.44</b>
<i>bet-o</i> 'but-but'	0	0	12.55	<b>12.55</b>
<i>o-vis tiek/ vis tik</i> 'but-still'	0.12	4	3.59	<b>7.71</b>
<i>tačiau-vis tiek/ vis tik</i> 'however-still'	0	6.7	0	<b>6.7</b>
<i>tačiau-vis dėlto</i> 'however-nevertheless'	0.7	3	1.8	<b>5.5</b>
<i>vis dėlto-vis tiek/ vis tik</i> 'nevertheless-still'	0	0.05	3.59	<b>3.64</b>
<i>bet-tačiau</i> 'but-however'	0	0	1.8	<b>1.8</b>
<i>o-bet</i> 'but-but'	0	0	1.8	<b>1.8</b>
<i>o-tačiau</i> 'but-however'	0	0.22	0	<b>0.22</b>
<b>Total</b>	<b>1.3</b>	<b>73.47</b>	<b>145.24</b>	<b>220.01</b>

As the normalized frequencies in Table 8 show, the most frequent GC-GC combinations in the Lithuanian corpora are *bet-vis tiek/ vis tik* 'but-still', *bet-vis dėlto* 'but-nevertheless' and *o-vis dėlto* 'but-nevertheless', which are among most frequent Lithuanian CDM combinations in general. Scarcely used in academic discourse, such combinations in fiction usually mark concession:

(9) Jos plaukai buvo dar šlapi, bet vis tiek labai raudoni. (CCLL-fict) 'Her hair was wet, but still very red.'

In the instance above, *bet-vis tiek* signals that S2 is true despite S1. GC-GC combinations, referred to as *complex adversative conjunctions* in *Lithuanian Grammar* (Ambrazas et al. 1997), are very typical of the Lithuanian language.

The predominant use of these and other Lithuanian GC-GC combinations is observed in spoken discourse, where they display a

wide variety of functions. For example, *bet-vis tiek* ‘but-still’ is used in argumentative contexts to signal disagreement with the message in S2.

- (10) -Tu, aš turiu ir medalį dar pirmos vietos!! <...> -Tai ten duodavo už dyką net <...> -Nu ir kas,  
bet vis tiek šokdavom gražiai. (CCLL-spok)  
‘-You, I also have got a medal for the first place!! <...> -They used to give those away for free <...> -So what, but still we used to dance beautifully.

In (10) the speaker marks disagreement with the idea expressed by the interlocutor in S1 (that the medal has low value), and questions its validity by pointing to the idea in S2 (that being able to dance beautifully was already fulfilling on its own). The disagreement is further intensified by the co-occurrence of a DM *nu ir kas* ‘so what’, used to signal contradiction with the interlocutor’s previous claims.

*Bet-o* ‘but-but’ and *o-bet* ‘but-but’, frequently occur in spoken discourse when the speaker wants to draw the attention of the addressee to the discrepancy between S1 and S2. For example, in (11) the idea in S1 is regarded as insufficient or unclear:

- (11) Ką tu studijuoji? -Elektroniką ir automatiką <...> -Gerai, o bet tai kokia tavo specialybė?  
(CCLL-spok)  
‘-What do you study? -Electronics and automation <...> -Okay, but what is your major?’

The example above shows that the addressee’s reply does not suffice and the speaker attempts to clarify and reformulate his/her initial question by using the combination. Example (11) illustrates the semantic bleaching of the DMs . and *bet* that combine and the emergence of new intersubjective meanings of the combination. *O-bet* ‘but-but’ shows the intersubjective meaning dimension as the speaker makes it explicit that the addressee’s reply did not match the speaker’s expectations. Bearing in mind the overlapping functions and translation correspondences of the two DMs in the combination, it is thus possible to render *o-betas* a stand-alone marker ‘but’.

In combinations *bet-tačiau* ‘but-however’, *o-tačiau* ‘but-however’, *vis dėlto-vis tiek* ‘nevertheless- still’ initial CDMs frequently function as markers of hesitation, or as turn-holding signals, providing the speaker some time to reformulate his/her ideas, as in (12):

- (12) Vieniems atrodo tai ne patys geriausi [pakeitimai], kitiems yra argumentų, dėl ko jie turėtų būti, tie pakeitimai. Tai čia, manau, vis dėlto dar vis tiek bus ir komitete svarstoma <...> (CCLL- spok)  
‘Some think these are not the best changes, others have arguments why there should be changes. This is, I think, however, still going to be considered by the committee.’

‘Some think these are not the best changes, others have arguments why there should be changes. This is, I think, **however**, **still** going to be considered by the committee.’

In the example above, initial *vis dėlto* does not signal that S2 is true despite the message in S1; instead, the CDM marks hesitation or turn

holding, thus illustrating the semantic bleaching of its primary meaning. Due to this semantic weakening *vis dėlto* in this case corresponds mostly and is rendered to 'however'. In actual speech hesitation might be reinforced by prosodic breaks, pauses and restarts (see Fischer & Alm 2013, 73).

### 3.3 General Contrastive-Specific Contrastive Combinations in English

In English, nine GC-SC combinations have been found in fiction, academic discourse and spoken discourse (see Table 9).

Table 9

Normalized frequency f1000000 and distribution of English GCSC combinations across discourse types

Combination	BNC-acad	BNC-fict	BNC-spok	Total
<i>but-rather</i>	33.53	4.9	2.81	41.24
<i>but-instead</i>	5.48	8.99	1.5	15.97
<i>but-on the other hand</i>	2.28	1.82	6.72	10.82
<i>yet-on the other hand</i>	1.04	0.19	0.2	1.43
<i>but-on the contrary</i>	0.91	0.25	0	1.16
<i>but-conversely</i>	0.52	0	0	0.52
<i>but-alternatively</i>	0.26	0	0	0.26
<i>but-in contrast</i>	0.07	0	0	0.07
<i>yet-in contrast</i>	0.07	0	0	0.07
<b>Total</b>	<b>44.16</b>	<b>16.15</b>	<b>11.23</b>	<b>71.54</b>

While GC-GC combinations have been found to dominate in spoken discourse, the frequency and distribution of GC-SC combinations shows the opposite, as spoken discourse contains the least amount of GC-SC combinations. Moreover, the highest number of GC-SC combinations is detected in academic discourse, due to frequent *but-rather* combinations.

GC-SC combinations in the English data are mostly headed by a CDM *but*. The most frequent GC-SC combinations in English are *but-rather/instead/on the other hand*; the least frequent combinations are *but/yet-in contrast*.

One of the most common functions of GC-SC combinations is to mark that a message in S2 provides a correct interpretation of S1, usually expressed as a negative:

- (13) What we have is not one, ongoing Parliament, but rather a series of parliaments, each now limited in its life to five years. (BNC-acad)

In (13), *but-rather* marks that the information contained in S2 corrects or reformulates the idea expressed in S1. This function is typically found in contexts where difficult concepts are clarified, which may explain why *but-rather* primarily occurs in academic discourse. *But-instead* also displays this function, yet this combination occurs mostly in fiction, probably due to its acceptability in less formal contexts.

Although not very frequent, *but-on the contrary* is similar to *but-instead* and *but-rather*, yet this combination marks the message in S2 as strikingly different from the message in S1:

(14) Your chairman was kind enough <...> to say that my attitude, for an historian, was very unhistorical and I took it that that was meant as praise. Some of my former colleagues would agree that my recent work is unhistorical but on the contrary condemn it for this <...> (BNC- fict)

As shown in (14), *but-on the contrary* emphasizes the juxtaposition of two conflicting opinions in S2 and S1, i.e. praise and condemnation.

*But-alternatively/on the other hand/conversely/in contrast* and *yet-on the other hand/in contrast* mark that two features of the same phenomenon are contrasted (15-17); in some cases, the symmetrical form *on the one hand... on the other hand* is used (17):

(15) I don't think he's really interested; but, on the other hand, he wouldn't refuse me anything. (BNC-fict)

(16) <...> that might mean victory for the state in the second debate. But, alternatively, it might mean that there is a further debate to conduct. (BNC-acad)

(17) On the one hand the effect of human activity upon earth environment has been very apparent and increasing, and yet on the other hand until the 1950s or 1960s the significance of human activity did not attract much attention by physical geographers (BNC-acad)

In most cases, *but-on the other hand* combinations are used in contexts where the speaker expresses his/her disagreement (typically stated in S2) with the opinion in S1 which is part of the shared knowledge (see instances 18-19).

(18) <...> and those sorts of things will obviously apply whatever group you're dealing with. But on the other hand I think the erm social work is now going (pause) back again to a situation where it's really rather more specialized. (BNC-spok)

(19) You know fair enough, some of the older men can't move as fast as a eighteen year old, but on the other hand it's what they got in their head that counts. (BNC-spok)

In the examples above, a variety of additional markers, such as the grammaticalized use of the epistemic marker *obviously* in (18) or *you know fair enough* in (19), are used in S1 to strengthen intersubjective agreement-seeking. In some cases, *I think* highlights the speaker's attitude (epistemic modality) (18) or performs even more grammaticalized functions, such as marking hesitation or word-search. The clustering of modal and interpersonal meanings with *I think* is also attested in Aijmer (1997).

### 3.4 General Contrastive-Specific Contrastive Combinations in Lithuanian

Seventeen GC-SC combinations have been identified in the Lithuanian data (see Table 10):

**Table 10**  
Normalized frequency f1000000 and distribution of the  
Lithuanian GCSC combinations across discourse types

Combination	CorAL it	CCLL- fict	CCLL- spok	Total
<i>bet-užtat</i> 'but-instead/still'	0.12	8.17	25.1	<b>33.39</b>
<i>bet- kita vertus/antra vertus</i> 'but-on the other hand'	0.2	2.7	7.17	<b>10.07</b>
<i>bet-iš kitos pusės</i> 'but-on the other hand'	0.12	0.32	8.96	<b>9.4</b>
<i>o-iš kitos pusės</i> 'but-on the other hand'	0.12	0	9	<b>9.12</b>
<i>o-priešingai</i> 'but-on the contrary'	0.35	3.4	0	<b>3.75</b>
<i>o-užtat</i> 'but-instead'	0	0	3.59	<b>3.59</b>
<i>bet-priešingai</i> 'but-on the contrary'	0.5	1.51	0	<b>2.01</b>
<i>bet-bet kokių (kuriuo) atveju</i> 'but-nevertheless'	0	0	1.79	<b>1.79</b>
<i>tačiau-kita vertus/antra vertus</i> 'however-on the other hand'	0.92	0.65	0	<b>1.57</b>
<i>o-kita vertus/antra vertus</i> 'but-on the other hand'	0.58	0.32	0	<b>0.9</b>
<i>tačiau-bet kokių (kuriuo) atveju</i> 'however-nevertheless'	0.35	0.38	0	<b>0.73</b>
<i>bet-nepaisant to</i> 'but-nevertheless'	0.12	0.54	0	<b>0.66</b>
<i>tačiau-priešingai</i> 'however-on the contrary'	0	0.27	0	<b>0.27</b>
<i>tačiau-nepaisant to</i> 'however-nevertheless'	0	0.27	0	<b>0.27</b>
<i>tačiau-vietoj(e) to</i> 'however-instead'	0	0.22	0	<b>0.22</b>
<i>tačiau-užtat</i> 'however-instead'	0	0.05	0	<b>0.05</b>
<i>vis tiek- bet kokių (kuriuo) atveju</i> 'still-nevertheless'	0	0.05	0	<b>0.05</b>
<b>Total</b>	<b>3.38</b>	<b>18.85</b>	<b>55.61</b>	<b>77.84</b>

Lithuanian GC-SC combinations are mainly detected in spoken discourse. On the other hand, fiction contains the widest variety of different GC-SC combinations.

As Table 10 illustrates, the most prevailing GC-SC combinations are headed by a CDM *bet* 'but'. The most frequent GC-SC combinations in the Lithuanian corpora are *bet-užtat* 'but-instead', *bet- kita vertus/antra vertus* 'but-on the other hand' and *bet-iš kitos pusės* 'but-on the other hand'.

GC-SG combinations mostly mark that S2 is true in spite of what is indicated in S1. However, in some cases other meanings emerge; for instance, commonly used in spoken discourse, *bet-užtat* 'but-instead/still' signals that what is indicated in S2 compensates for what is mentioned in S1:

(20) Teko daug darbo nudirbti, bet užtat rezultatai džiugino. (CCLL-spok) 'I had to work a lot, but still results brought me joy.'

As (20) shows, *bet-užtat* 'but-still' signals that the feeling of accomplishment is seen as rewarding enough to make up for the hardship experienced while working a lot.



Frequent *o-priešingai* ‘but-on the contrary’, *bet-priešingai* ‘but-on the contrary’ and *tačiau-priešingai* ‘however-on the contrary’ combinations stress the contradiction or denial of a message in S1, as in (21):

- (21) Dūzgimo tikslas - ne pabrėžti šokį, o priešingai, jį pajvairinti. (CCLL-fict)  
 ‘The goal of buzzing is not to accentuate the dance, but on the contrary, to add variety to it.’

In addition, these combinations tend to occur in contexts where the message in S1 is formulated with a negative, as illustrated in (21). Thus, S2 usually provides a correct interpretation of the message in S1.

Although *tačiau-vietoj(e) to* ‘however-instead’ and *bet-vietoj(e) to* ‘but-instead’ combinations display low frequencies, they are important means of marking contradiction between the expectation or past intention indicated in S1 and the real outcome specified in S2, as in (22):

- (22) Už kelių žingsnių turėjau pradėti kilti į kalną, tačiau vietoje to įžengiau į jo vidų. (CCLL-fict) ‘In a few steps I was supposed to start ascending the mountain, however instead I walked inside of it.’

(22) shows that this combination marks that the prior plan (ascending the mountain) was not realised.

### 3.5 Specific Contrastive-General Contrastive Combinations in English

Five possible SC-GC combinations have been found in the English data. The frequency and distribution of each SC-GC combination are provided in Table 11:

Table 11

Normalized frequency f1000000 and distribution of English SCGC combinations across discourse types

Combinations	BNC-acad	BNC-fict	BNC-spok	Total
<i>in contrast-however</i>	0.52	0	0	0.52
<i>on the other hand-however</i>	0.2	0.06	0	0.26
<i>instead-however</i>	0.07	0.06	0	0.13
<i>conversely-however</i>	0.13	0	0	0.13
<i>alternatively-however</i>	0.13	0	0	0.13
<b>Total</b>	<b>1.05</b>	<b>0.12</b>	<b>0</b>	<b>1.17</b>

The results obtained are in line with the findings in Fraser (2013), who claims that the only SC-GC combinations are the ones with *however* in the final position of a combination. *However* combines with all SCs, except for *rather* and *on the contrary*. SC-GC combinations make up the bulk of all combinations in academic discourse. Few instances of SC-GC combinations have been found in fiction and no instances of these combinations have been attested in spoken discourse.

In SC-GC combinations, the initial CDM retains its core meaning of contrast and *however* is used as a focus marker (see Fraser 2013) which signals that S2 contains information relevant to the interpretation of S1, for instance:

(23) <...> Norfolk priests were very much more affluent. In contrast, however, the omission of information for the Westwood nunnery leaves the Worcestershire aggregate defective, while data relating to secular clergy in the Aylesbury hundreds of Buckinghamshire are probably inadequate. (BNC-acad)

In (23), *in contrast* introduces comparison of the accumulated wealth by the clergy across counties in the UK, such as Norfolk, Worcestershire and Buckinghamshire. A CDM *however* signals that S2 contains information that might challenge the validity of the comparison shown in S1.

### 3.6 Specific Contrastive-Specific Contrastive Combinations in Lithuanian

The combination *užtat-priešingai* ‘but-on the contrary’ is the only SC-SC combination identified in the Lithuanian data. Its frequency amounts to (0.05/1,000,000).

(24) Šalia jų sėdėjo tipingas svajotojas <...>. Užtat priekyje, šonuose, priešingai, matyti įsitempę žmonės <...> (CCLL-fict)

‘There sat a typical dreamer next to them <...>. But in front of them, on the sides, on the contrary, one could see tense people <...>’

In (24) the combination marks the juxtaposition of two entities (a typical dreamer and tense people). *Užtat* (which in general mostly corresponds to ‘instead’) is rendered in this case to ‘but’ as it indicates a contrastive relation between S1 and S2 and draws attention to the upcoming opposition, signalled by *priešingai* ‘on the contrary’.

## 4 Conclusions

The data shows that the frequency and distribution of CDM combinations are language and discourse dependent. CDM combinations are more common in Lithuanian than in English. Lithuanian CDMs have been found to dominate in spoken discourse. The least amount of Lithuanian CDM combinations has been observed in academic discourse. On the contrary, English CDMs commonly occur in academic discourse, as opposed to fiction and spoken discourse. The overall dominance of Lithuanian CDM combinations stems from the fact that the majority of CDM combinations are of GC-GC type. Such combinations have proved to be most multifunctional and predominantly found in spoken discourse revealing a variety of pragmatic functions as attracting attention, marking disagreement and hesitation.

Lithuanian *bet*-CDM ‘but-CDM’ and English *but*-CDM combinations display the highest frequencies and functional variation. Both *bet* ‘but’ and *but* combine with all the other CDMs under investigation. *Bet*-CDM ‘but-CDM’ combinations are most common in spoken discourse, while *but*-CDM combinations prevail in academic discourse. The most frequent CDM combinations in Lithuanian are *bet-vis tiek/ vis tik* ‘but-still’, *bet-vis dėlto* ‘but-nevertheless’, *bet-užtat* ‘but-instead’, *o-vis dėlto* ‘but-nevertheless’ and *bet-o* ‘but-but’, whereas *but-still*,

*but-rather, but-nevertheless, but- instead and but-on the other hand* prevail in English.

The study raises diverse implications for further research. Future studies could be carried out by using corpora comprising other types of discourse with a focus on other types of DMs, not only those signalling contrast. In addition, a more complex investigation semantic-pragmatic profile of DM combinations should be done with the focus on the development of their (inter)subjective meanings in spontaneous speech.

## Literatūra

- Aijmer Karin 1997, I think- an English modal particle, *Modality in Germanic languages: Historical and comparative perspectives*, 99 (1), 1-47, [https://www.researchgate.net/publication/246407952\\_I\\_think-a\\_n\\_English\\_modal\\_particle](https://www.researchgate.net/publication/246407952_I_think-a_n_English_modal_particle)
- Aijmer Karin 2004, Pragmatic markers in spoken interlanguage, *Nordic Journal of English, Studies* 3 (1), 173–190, <http://ojs.ub.gu.se/ojs/index.php/njes/article/view/241/238>
- Aijmer Karin 2007, The meaning and functions of the Swedish discourse marker *alltså*—Evidence from translation corpora, *Catalan Journal of Linguistics*, 6 (1), 31-59, <https://revistes.uab.cat/catJL/article/view/v6-aijmer/135>
- Aijmer Karin 2016, Pragmatic markers as constructions, The case of *anyway*, *Outside the clause: Form and function of extra-clausal constituents*, 29-58, [https://www.researchgate.net/publication/309858381\\_Pragmatic\\_markers\\_as\\_constructions\\_The\\_case\\_of\\_anyway](https://www.researchgate.net/publication/309858381_Pragmatic_markers_as_constructions_The_case_of_anyway)
- Aijmer Karin, Elgemark Anna 2013, The pragmatic markers *look* and *listen* in a cross-linguistic perspective, *Of butterflies and birds, of dialects and genres: Essays in honour of Philip Shaw*, Nils-Lennart Johannesson, Melchers Gunnel, Björkman Beyza, eds., Stockholm: Stockholm University Press.
- Ambrazas Vytautas, Geniušienė Emma, Girdenis Aleksas, Sližienė Nijolė, Tekorienė Dalija, Valeckienė Adelė, Valiulytė Elena 1997, *Lithuanian Grammar*, Vilnius: Baltos lankos.
- Belogurova Greta 2018, *Combinatory potential of contrastive discourse markers in a cross-linguistic perspective: semantic functional analysis* (Unpublished master's thesis), Vilnius: Vilnius university.
- Biber Douglas, Johansson Stig, Leech Geoffrey, Susan Conrad, Edward Finegan, Randolph Quirk 1999, *Longman grammar of spoken and written English*, 2, London: Longman.
- Bielinskienė Agnė 2009, Relevancijos teorija ir diskurso jungtųjų bei jungiamųjų žodžių pragmatika, *Kalbų studijos* 15, 53-62, [https://www.kalbos.lt/zurnalai/15\\_numeris/09.pdf](https://www.kalbos.lt/zurnalai/15_numeris/09.pdf)
- Blakemore Diane 1987, *Semantic constraints on relevance*, Oxford: Blackwell.
- Blakemore Diane 2002, *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*, 99, Cambridge: Cambridge University Press.
- Brinton Laurel J. 1996, *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin/New York: Mouton de Gruyter.
- Brinton Laurel J. 2008, *The comment clause in English: Syntactic origins and pragmatic development*. Cambridge: Cambridge University Press.

- Buitkienė Janina 2011, Discourse markers in English and Lithuanian texts, *Kalba ir kontekstai*, 4, 99–107, <http://etalpykla.lituanistikadb.lt/fedora/objects/LT-LDB-0001:J.04~2011~1407327889119/datastreams/DS.002.1.01.ARTIC/content>
- Defour Tine, D'Hondt Ulrique, Simon-Vandenberghe Anne-Marie, Willems Dominique 2010, Degrees of pragmaticalization: The divergent histories of *actually* and *actuellement*, *Languages in Contrast* 10 (2), 166–193, [https://www.researchgate.net/publication/233628593\\_Degrees\\_of\\_pragmaticalization\\_The\\_divergent\\_histories\\_of\\_'actually'\\_and\\_actuellement](https://www.researchgate.net/publication/233628593_Degrees_of_pragmaticalization_The_divergent_histories_of_'actually'_and_actuellement)
- Fagard Benjamin 2010, *É vida, olha...*: Imperatives as discourse markers and grammaticalization paths in Romance: A diachronic corpus study, *Languages in Contrast* 10 (2), 245–267.
- Fedriani Chiara, Ghezzi Chiara 2014. The pragmaticalization of verbs of movement and exchange in Latin and Italian: Paths of development from lexicon to pragmatics. *Studia linguistica et philologica in honorem prof. univ. dr. Michaela Livescu*. Badescu Ilona, Popescu Mihaela, eds., Craiova: Editura Universitaria, 116–139.
- Fischer Kerstin, Alm Maria 2013, A radical construction grammar perspective on the modal particle-discourse particle distinction, *Discourse markers and modal particles: Categorization and description*, 47–87.
- Fraser Bruce 1990, An approach to discourse markers, *Journal of pragmatics*, 14 (3), 383–398. <https://www.sciencedirect.com/science/article/abs/pii/037821669090096V>
- Fraser Bruce 1996, Pragmatic markers, *Pragmatics* 6 (2), 167–190, <https://www.jbe-platform.com/content/journals/10.1075/prag.6.2.03fra>
- Fraser Bruce 1999, What are discourse markers?, *Journal of pragmatics*, 31 (7), 931–952, <https://www.sciencedirect.com/science/article/abs/pii/S0378216698001015>
- Fraser Bruce 2006, Towards a theory of discourse markers, *Approaches to discourse particles*, Fischer Kerstin, ed., Amsterdam: Elsevier, 189–204.
- Fraser Bruce 2009, An account of discourse markers, *International review of pragmatics*, 1 (2), 293–320, [https://www.researchgate.net/publication/233594552\\_An\\_Account\\_of\\_Discourse\\_Markers](https://www.researchgate.net/publication/233594552_An_Account_of_Discourse_Markers)
- Fraser Bruce 2013, Combinations of contrastive discourse markers in English, *International Review of Pragmatics*, 5 (2), 318–340, [https://www.researchgate.net/publication/274985833\\_Combinations\\_of\\_Contrastive\\_Discourse\\_Markers\\_in\\_English](https://www.researchgate.net/publication/274985833_Combinations_of_Contrastive_Discourse_Markers_in_English)
- Peter 2005, The pragmatic marker-discourse marker dichotomy reconsidered – the case of well and of course, Ph.D. dissertation, Debreceni Egyetem: Bölcsészettudományi Kar, Debrecen, [https://www.academia.edu/2537042/The\\_Pragmatic\\_Marker-discourse\\_Marker\\_Dichotomy\\_Reconsidered\\_the\\_Case\\_of\\_Well\\_and\\_Of\\_Course](https://www.academia.edu/2537042/The_Pragmatic_Marker-discourse_Marker_Dichotomy_Reconsidered_the_Case_of_Well_and_Of_Course)
- Hall Alison 2007, Do discourse connectives encode concepts or procedures?, *Lingua*, 117 (1), 149–174, [https://www.researchgate.net/publication/22759965\\_Do\\_discourse\\_connectives\\_encode\\_concepts\\_or\\_procedures](https://www.researchgate.net/publication/22759965_Do_discourse_connectives_encode_concepts_or_procedures)
- Haselow Alexander 2012, Discourse organization and the rise of final then in the history of English. *English Historical Linguistics 2010: Selected Papers from the Sixteenth International Conference on English Historical Linguistics*

- (ICEHL 16), Hagedus Iren, Fodor Alexandra, eds., Amsterdam: John Benjamins, 153–175.
- Heine Bernd 2013, On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, 51 (6), 1205–1247, [https://www.researchgate.net/publication/274500591\\_On\\_discourse\\_markers\\_Grammaticalization\\_pragmaticalization\\_or\\_something\\_else](https://www.researchgate.net/publication/274500591_On_discourse_markers_Grammaticalization_pragmaticalization_or_something_else)
- Jasionytė-Mikučionienė Erika 2016, Imperatyvinės kilmės diskurso markeriai lietuvių kalboje: *klausyk ir žiūrėk* atvejis, *Kalbotyra*, 68, 23–41, <https://www.zurnalai.vu.lt/kalbotyra/article/view/10317>
- Knott Alistair, Sanders Ted 1998, The classification of coherence relations and their linguistic markers: An exploration of two languages, *Journal of pragmatics*, 30 (2), 135–175, <https://www.sciencedirect.com/science/article/abs/pii/S037821669800023X>
- Lenker Ursula 2000, *Soplice and witodlice*: Discourse markers in Old English. *Pathways of change: Grammaticalization in English*, Fischer Olga, Rosenbach Anette, Stein Dieter, eds., Amsterdam: John Benjamins, 229–249.
- Masaitienė Dalia 2003, Discourse Markers in English and Lithuanian, *Kalbotyra* (53), 64–70. Maschler Yael 1994, Metalinguaging and discourse markers in bilingual conversation, *Language in Society* 23 (3), 325–366, <https://www.jstor.org/stable/4168534?seq=1>
- Pugačiauskaitė Rasa 2012, *Discourse markers in English and Lithuanian texts*, Master's thesis, Vilnius: Lithuanian University of Educational Sciences, <https://vb.vdu.lt/object/elaba:2044691/2044691.pdf>
- Redeker Gisela 1991, Linguistic markers of discourse structure, *Linguistics*, 29 (6), 1139–1172, [https://www.researchgate.net/publication/243767147\\_Linguistic\\_markers\\_of\\_discourse\\_structure](https://www.researchgate.net/publication/243767147_Linguistic_markers_of_discourse_structure)
- Schiffrin Deborah 1987, *Discourse markers*, Cambridge: Cambridge University Press.
- Schourup Lawrence 1985, *Common discourse particles in English conversation*, New York: Garland Publishing.
- Shinzato Rumiko 2017, Grammaticalization of PMs/DMs/MMs in Japanese, *Pragmatic markers, discourse markers and modal particles: New perspectives*, Fedriani Chiara, Sansò Andrea, eds., Amsterdam: John Benjamins, 305–334.
- Šimčikaitė Alė 2012, Spoken discourse markers in learner academic writing, *Kalbų studijos/Studies about Languages*, 20, 27–34, <http://kalbos.ktu.lt/index.php/KStud/article/view/1196>
- Szczyrbak Magdalena 2017, Modal Adverbs of Certainty in EU Legal Discourse: A Parallel Corpus Approach, *Contrastive Analysis of Discourse-pragmatic Aspects of Linguistic Genres* 5, Aijmer Karin, Lewis Diana, eds., Springer.
- Traugott Elizabeth Closs 2014, Intersubjectification and clause periphery, *Intersubjectivity and intersubjectification in grammar and discourse: Theoretical and descriptive advances*, 65, Brems Lieselotte, Ghesquière Lobke, Van de Velde Freek, eds., Amsterdam/Philadelphia: John Benjamins, 7–28.
- Usonienė Aurelija 2016, *Reikšmės pasaulis: Tekstynais paremti semantiniai kalbų tyrimai*, Vilnius: Akademinė leidyba.

- Vaičiūnienė Vilhelmina, Mažeikienė Viktorija 2016, Cross-linguistic variation of the discourse marker *however* from English into Lithuanian, *Kalba ir kontekstai* 7, 142–152.
- Valančė Danguolė 2017, *Pabrėžiamųjų dalelių funkcijos ir vartoseną dabartinėje rašytinėje lietuvių kalboje*, Ph.D. dissertation, Kaunas: Vytautas Magnus University.
- Van Olmen Daniël 2011, *The imperative in English and Dutch: A functional analysis in comparable and parallel corpora*, Ph.D. dissertation, Antwerp: University of Antwerp.
- Van Olmen Daniël 2013, The imperative of *say* as a pragmatic marker in English and Dutch, *Journal of Germanic Linguistics* 23 (3), 247–287, [https://www.researchgate.net/publication/259435172\\_The\\_Imperative\\_of\\_Say\\_as\\_a\\_Pragmatic\\_Marker\\_in\\_English\\_and\\_Dutch](https://www.researchgate.net/publication/259435172_The_Imperative_of_Say_as_a_Pragmatic_Marker_in_English_and_Dutch)
- Verikaitė Daiva 2005, Variation of conjunctive discourse markers across different genres, *Žmogus ir žodis*, 7 (3), 68–75, [https://virtualbiblioteka.vu.lt/permalink/f/1ferss/TN\\_eeol\\_s187309](https://virtualbiblioteka.vu.lt/permalink/f/1ferss/TN_eeol_s187309)
- Waltereit Richard 2002, Imperatives, interruption in conversion, and the rise of discourse markers: A study of Italian ‘guarda’, *Linguistics*, 40 (5), 987–1010.

## Data sources

- BNC- British National Corpus (<http://corpus.byu.edu/bnc/>)
- CCLL- Corpus of Contemporary Lithuanian Language (<http://tekstynas.vdu.lt/>)
- CorALit- Corpus Academicum Lithuanicum (<http://coralit.lt/>)

## List of abbreviations

- CDM(s): contrastive discourse marker(s)
- DM(s): discourse markers(s)
- GC(s): general contrastive discourse marker(s)
- S1: discourse segment that precedes a DM
- S2: discourse segment that follows a DM
- SC(s): specific contrastive discourse marker(s)