

# ON SOME VARIATION IN THE USE OF DISCOURSE MARKERS BY CZECH AND GERMAN STUDENTS OF ENGLISH

*Renata Povolná*

## **Abstract**

Since recent studies on academic English have shown considerable cross-cultural variation in texts written by non-native speakers (Clyne 1987, Ventola & Mauranen 1991, Čmejrková & Daneš 1997, Duszak 1997, Chamonikolasová 2005, Stašková 2005, Mur-Dueñas 2008, Wagner 2011, Dontcheva-Navratilova 2012, Povolná 2012), the paper investigates a corpus of diploma theses written by Czech and German students of English with the aim of finding out how novice non-native writers from different discourse communities (Swales 2004) use causal and contrastive discourse markers (DMs) associated with hypotactic and paratactic relations in order to build coherence relations (Taboada 2006) in academic texts. In addition, the author attempts to find out whether there is any variation in the preferences of novice writers depending on the different fields of study, i.e. diploma theses written in the areas of linguistics and methodology, and whether the use of selected DMs by Czech and German students differs from the writing habits of native speakers of English.

## **Key words**

written academic discourse, diploma theses, discourse markers, causal and contrastive relations, hypotactic and paratactic relations

## **1 Introduction**

Semantic relations that may hold between adjacent or more distant segments of discourse tend to be expressed explicitly by some markers above all in written academic discourse where the expression of the author's argumentation becomes of crucial importance. Since causal and contrastive relations rank among the most complex of all semantic relations that may hold within the text (Kortmann 1991), the paper investigates ways in which students of English from two different discourse communities, i.e. different cultural backgrounds (Czech Republic and Germany), use discourse markers (DMs) to express these most complex relations when building coherence relations, i.e. "relations that hold together different parts of the discourse" (Taboada 2006: 567) in order to enable the reader(s) to perceive the text as being coherent. "The process of creating coherent texts involves an indication of relationships between the things one is 'on about'" (Halliday & Hasan 1989: 94), since an appropriate application of guiding signals indicating relationships between segments of discourse by

the writer can foster a deductive process of interpretation and understanding on the side of the reader(s). Moreover, a fundamental change in the understanding of academic discourse “from a formal text-based perspective to a functional perspective that concentrates on the writer and the writing process and, even more, on the reader and the cognitive construction of discourse in a community” (Schmied 2011: 1) has brought to the fore the need to focus on how meaning is conveyed between the writer and his/her prospective reader(s) in academic discourse communities.

By overtly signalling how the writer intends the discourse segment that follows to relate to the previous discourse segment(s), DMs, in particular those expressing cause and contrast, contribute to cohesion and enhance the establishment and maintenance of coherence in written academic discourse. Unlike coherence in spoken discourse, which can be permanently negotiated by all discourse participants (Povolná 2007, 2009), coherence in written discourse cannot be negotiated explicitly because the context is split (Fowler 1986) and there is no reciprocal management of the discourse (Seidlhofer & Widdowson 1999). Thus the writer, for example, the author of an academic text has to anticipate the “expectations of the reader and to use explicit signals” (Dontcheva-Navratilova 2007: 128) to guide the prospective academic audience towards an intended interpretation of the text. Coherence – understood here in agreement with Bublitz (1999) as a dynamic and hearer/reader-oriented interpretative notion – and the quality of being coherent are important in particular for the expressing of semantic relations in written academic discourse, including diploma theses produced by novice non-native speakers of English.

Recent studies on written academic discourse have shown considerable cross-cultural variation in academic texts written in English (cf. the Abstract above), in particular now that English has become “the global lingua franca of academia” although there are no native speakers of academic English (Mauranen, Hynninen & Ranta 2010). This variation results from the fact that under the influence of their L1 writing habits many authors from different discourse communities attempt to produce academic texts in English with the intention of achieving native-like fluency by following conventions typical of academic discourse as produced by native-speakers of English.

Although different in the extent to which they have adopted the writing style of English-speaking countries, Czech and German academic writing as representatives of Central European intellectual traditions can be characterized by certain common features which distinguish them from academic writing that is typical of the dominant Anglo-American tradition, in particular on account of differences in educational systems and intellectual styles and attitudes to knowledge and content.

Anglo-American academic texts in general tend to be more dialogic, interactive and reader-oriented, which results from an overall discourse organization through explicit signposting, including text organizers such as DMs. These text characteristics are in contrast to the rather monologic, less interactive and more writer-oriented texts usually connected with Teutonic intellectual traditions (Galtung 1985) and connected with academic texts written in Czech and German. These academic backgrounds prefer a more impersonal style of writing with fewer reader-friendly devices such as text organizers and explicit clues concerning content (e.g. Clyne 1987, Chamonikolasová 2005). Instead, intellectual effort and a readiness to process demanding text are required of the academic audience.

## 2 The discourse markers under investigation

Discourse markers are viewed here, in agreement with Fraser (1990, 1999), as a class of commentary pragmatic markers signalling relationships between adjacent or more distant segments of discourse; hence they contribute to both cohesion and coherence. (For the term ‘discourse segment’ or ‘segment of discourse’, cf. Fraser 1999.) By signalling how the writer intends to relate the current message to the previous discourse segment(s), these markers perform above all text-organizing functions and their meaning is procedural rather than conceptual, as they are “separate from the propositional content” (Fraser 1999: 302) of the utterances into which they are inserted. The writer uses DMs intentionally in order to enable his/her prospective reader(s) to derive meaningful discourse from the text, since “a text is not coherent in itself but is understood as coherent in an actual context” (Bublitz 1988: 32) in which the current reader attempts to interpret the text as discourse by relating it to his/her background knowledge and previous experience of processing similar types of discourse. If a guiding signal is absent, the propositional content of the respective discourse segments remains the same, as would be the case in Example 1 after omitting the marker *in consequence*; however, without any guiding signal it could be more difficult for the reader(s) to arrive at the interpretation intended by the writer.

- (1) *Pupils at upper-primary school are considered to be basic users and their knowledge of the second language is not at high level. In consequence, feedback can be in their mother tongue.*  
(BrnoCorpusMeth 1)

The causal and contrastive DMs selected for this study are expected to occur rather frequently in the data analysed, since their frequent use in written academic discourse results from the author's need to present his/her arguments and standpoints to the academic audience in a straightforward and comprehensive way; moreover, the frequent use of DMs reflects "the characteristic choice of [the] register [of academic prose] to mark the links between ideas overtly" (Biber et al. 1999: 880).

The study is concerned with causal and contrastive relations obtaining at clausal and higher levels of discourse because it is assumed that at these levels the marker relates two separate messages (Fraser 1999: 939-940), thus functioning as a DM (Example 2), while at a lower level it serves as a conjunction within a single message (Example 3) and hence is outside the scope of the present analysis. However, it should be noted that conjunctions functioning within single messages are not at all frequent in the analysed data.

- (2) *These differences have strong linguistic correlates, for both vocabulary and grammatical features (Biber; 2006, p. 226). As a consequence, science and engineering articles are found to be much more impersonal than papers written in the humanities and social sciences (Hyland, 2006),*  
...  
(ChemCorpusLing 19)
- (3) *Many advertisements often become a popular topic for conversation or even enrich the language by various "catchy" slogans, although usually just temporarily.*  
(BrnoCorpusLing 5)

From a morphological viewpoint, the DMs under examination are drawn primarily from conjunctions (e.g. *as, because, since, although, but, while*), adverbs (e.g. *therefore, thus, however, nevertheless, yet*), or prepositional phrases (e.g. *in contrast, on the other hand*). From a syntactic point of view, both causal and contrastive DMs can be divided into markers associated with either hypotactic relations or paratactic relations. The reason for this division is above all an expected difference between the two syntactic groups in frequency of occurrence and the two semantic classes of DMs selected for the analysis. The hypotactic relation (i.e. the relation between two discourse segments one of which is dependent on the other) is typically expressed overtly by certain markers, such as *although, as, because, since, and while*. By contrast, the paratactic relation (i.e. the relation between two segments which are not dependent on each other), apart from being

indicated by certain markers, such as *consequently*, *however*, *on the other hand*, *thus* and *therefore*, can often remain unexpressed explicitly; this does not mean that there cannot be semantic clues in the respective discourse segments, such as nouns, verbs, adjectives (e.g. *contrast*, *contrasting*), prepositional phrases (e.g. *because of*, *on account of*, *in spite of*, *in contrast to*, *as a result of*; for the last two prepositional phrases listed, cf. Examples 4 and 5 respectively), or some other ways of expressing cause or contrast; these possibilities, however, have remained outside the scope of the present inquiry.

- (4) *The study moreover revealed striking similarities between German and American writers in their use of connectors in contrast to British connector usage.*  
(ChemCorpusLing 19)
- (5) *The problem of getting students to express themselves freely in the foreign language has come into prominence in recent years as a result of the growing emphasis on communicative abilities.*  
(BrnoCorpusMeth 4)

As regards the frequency rate of DMs associated with hypotactic and paratactic relations, it should be stated that since the paper deals with written academic discourse, in which clear argumentation and support of the authors' own arguments and standpoints play an important role, a high number of explicitly expressed markers, in particular those occurring in hypotactic relations, is expected, because these are mostly marked overtly, and, as Taboada claims, DMs expressing cause and concession are "typically expressed through subordination" (2006: 576). Moreover, subordination is considered a characteristic of formal written rather than informal spoken discourse, in which 'loose' coordination is more typical (cf. also Leech & Svartvik 1994: 14).

Concerning concession, it should be noted that it is viewed here as a special case of contrast, namely that between the expected/usual causal relationship and the actual situation (Dušková et al. 1988, Fraser 1999), and therefore contrastive DMs subsume markers expressing contrast as well as concession, since, as mentioned in Biber et al., "in some cases, elements of contrast and concession are combined in uses of linking adverbials" (1999: 878), and it is not always possible to draw a clear borderline between these two semantic classes.

### 3 Material and methods

The investigation is based on a sample taken from a larger corpus comprising diploma theses written by students of English in their final year of study at the Faculty of Education, Masaryk University, Brno, Czech Republic, and students at Chemnitz University of Technology, Chemnitz, Germany. For the purposes of this paper only a sample amounting to about 352,000 words has been used, since it is argued, in agreement with Flowerdew (2004: 18), that specialized corpora are more appropriate than general corpora for a comparative study of academic discourse, especially when analysing particular language features in one text type. The corpus can be divided into two corpora, each representing students' diploma theses from a different discourse community and further subdivided into works written in the areas of linguistics and methodology. The individual subcorpora vary slightly in length (85,000-90,500 words) as it has been necessary to exclude from the analysis all parts of the texts which comprise tables, figures, graphs, references, sources, examples, long quotations, and in the case of methodology theses lesson plans prepared by students as part of their research.

Since the main objective of the inquiry is not a detailed comparison of the four subcorpora but rather of the ways in which advanced learners of English apply the selected markers for the expression of causal and contrastive relations in academic texts, differences in the length of the individual texts are not considered relevant here. Moreover, in order to get comparable data for the analysis, all the results discussed and exemplified in this article have been normalized for the frequency of occurrence of selected DMs per 1,000 words, actual numbers being mentioned only occasionally. It is important to stress that there are also marked differences between students' diploma theses written within the same area, which supports my assumption that variation between fields of study, in this case linguistics and methodology, can be caused by individual students' knowledge of the selected markers and consequent preferences in writing habits.

As regards the methodology applied in the analysis, all the texts were first computer-processed using the AntConc concordancer and then sifted manually for both qualitative and quantitative results since some of the language features under examination can perform functions other than those of DMs in written discourse.

## 4 Results from the Czech corpus

### 4.1 Causal DMs

No. of words in the texts	87,636	88,628	Total: 176,264	
Hypotactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<i>as</i>	<b>0.62</b>	<b>0.99</b>	<b>0.81</b>	<b>142</b>
<i>because</i>	0.94	0.77	0.85	150
<i>Since</i>	0.34	0.18	0.26	46
All hypotactic DMs	1.89	1.94	1.92	338
Paratactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<i>as a result</i>	0.10	0.02	0.06	11
<i>consequently</i>	0.19	0.17	0.18	32
<b>(and) so</b>	<b>0.41</b>	<b>0.23</b>	<b>0.32</b>	<b>56</b>
<i>so that</i>	0.06	0.12	0.09	16
<i>therefore</i>	0.96	0.41	0.68	120
<i>thus</i>	0.99	0.47	0.73	129
<i>Then</i>	0.13	0.07	0.10	17
All paratactic DMs	2.85/2.93	1.56/1.64	2.16/2.28	381/402
All DMs	4.83	3.58	4.20	740

**Table 1a: Most frequent causal DMs used by Czech students**

Table 1a indicates that in diploma theses written by Czech students causal relations expressed by explicit hypotactic DMs, although realized by three different markers only (*because* (0.85; 150 occurrences), *as* (0.81; 142) and *since* (0.26; 46); listed by decreasing frequency), are relatively frequent (1.92; 338) when compared to those expressed by explicit paratactic markers (2.28; 402). This result does not correspond with my expectation that hypotactic relations are more frequently expressed by an overt marker in academic discourse. Nevertheless it can be stated that all the three possible hypotactic markers have been found relatively frequently in the data, while paratactic markers, although more frequent in total number, are rather unevenly distributed, some of them having a frequency rate of less than ten tokens in the Czech corpus; owing to their rather limited use, the markers *accordingly*, *as a consequence*, *hence*, *for*, *in consequence*, *now*, *of course* and *somehow* have been excluded from Table 1a, although they are counted in the total number in the last two lines of the table. (Cf. e.g. 381/402 in the last column, where 381 is the number of paratactic DMs listed in the table, while 402 equals the total number of paratactic DMs found in the Czech corpus.) Of the fifteen different types of paratactic markers searched

for during the analysis, only seven (fewer than a half) occur with noteworthy frequency, i.e. ten or more occurrences in the Czech corpus; these are all listed in the table. It should be noted here that all the DMs which are highlighted in the tables signal some marked differences between the Czech and German corpora; thus, for instance, the highlighted results indicate that the total number of the DMs *as* and *so* is much higher in the Czech corpus than in the German corpus (cf. Tables 1a and 2a).

As can be seen from the table, some markers are very frequent, such as *because* (0.85), *as* (0.81), *thus* (0.73), and *therefore* (0.68) (listed by frequency of occurrence), or relatively frequent, i.e. having more than thirty tokens in the corpus, such as *so* (0.32), *since* (0.26), and *consequently* (0.18). When put together the most frequent hypotactic and paratactic markers of all, namely *because*, *as*, *thus* and *therefore*, represent more than two thirds of all markers found in the corpus. Hence it can be concluded that Czech students use a relatively limited repertoire of the more common DMs rather than resorting to the wider choice of markers at their disposal, in particular when expressing paratactic relations. The frequent application of the same marker is illustrated in Example 4, in which the writer prefers using the same DM repeatedly rather than applying a different DM (cf. Example 9 below):

- (4) *According to Harmer (1991) we can divide reasons into two broad categories: Instrumental – people listen to something because it helps them to achieve a clear goal. E.g. when we travel by air we usually know the number of flight but we have to listen to information about announcement of our flight. People listen to the news because they have a general interest what is going on, they are going on holiday and listen to a weather forecast to know what the weather will be like.*  
(BrnoCorpusMeth 2)

Finally, it is worth noting that there are differences between the fields of study in terms of both types and tokens of DMs students use to express causal relations, the most striking difference being the far more frequent use of paratactic DMs in the linguistics subcorpus (4.83; by more than 100 tokens) than in the methodology one (3.58). These differences can be caused by individual students' preferences and writing habits, which in the case of linguistics-oriented theses can be enhanced by students' stronger awareness of marker use in the organizing of discourse and the probable influence of overt instructions provided by teacher(s) of academic writing.



## 4.2 Contrastive DMs

No. of words in the texts	87,636	88,628	Total: 176,264	
Hypotactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<i>although</i>	0.37	1.83	0.32	57
<b><i>even if</i></b>	<b>0.06</b>	<b>0.96</b>	<b>0.09</b>	<b>16</b>
<b><i>even though</i></b>	<b>0.21</b>	<b>0.01</b>	<b>0.17</b>	<b>30</b>
<b><i>though</i></b>	<b>0.24</b>	<b>0.11</b>	<b>0.18</b>	<b>32</b>
<i>while</i>	0.19	0.10	0.15	26
<i>whereas</i>	0.21	0.06	0.13	23
All hypotactic DMs	1.27/1.35	0.82/0.82	1.04/1.08	184/191
Paratactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<b><i>but</i></b>	<b>2.05</b>	<b>1.83</b>	<b>1.94</b>	<b>342</b>
<i>however</i>	0.66	0.96	0.81	143
<i>nevertheless</i>	0.08	0.01	0.25	44
<i>nonetheless</i>	0.39	0.11	0.09	16
<i>on the contrary</i>	0.09	0.05	0.07	12
<b><i>on the other hand</i></b>	<b>0.11</b>	<b>0.07</b>	<b>0.33</b>	<b>58</b>
<i>Still</i>	0.41	0.25	0.18	31
<i>though</i>	0.23	0.12	0.10	18
<b><i>yet</i></b>	<b>0.18</b>	<b>0.02</b>	<b>0.32</b>	<b>57</b>
All paratactic DMs	4.77/5.13	3.42/3.59	4.09/4.36	721/768
All DMs	6.48	4.41	5.42	959

**Table 1b: Most frequent contrastive DMs used by Czech students**

The results concerning contrastive DMs as they are used by Czech students are given in Table 1b. It provides evidence that, contrary to my expectation and as with causal relations, contrastive relations expressed by hypotactic DMs are considerably less frequent (1.08; 191 occurrences) in the data than those expressed by paratactic markers (4.36; 768), although hypotactic relations are usually marked overtly in academic discourse. The most striking result is the uneven distribution of the contrastive markers (altogether 959) in terms of both types and tokens, in particular with paratactic markers. (For comparison with causal markers, cf. Table 1a above.) Of the thirty-eight different types of contrastive DMs included in the investigation, only fifteen occur with noteworthy frequency, i.e. ten or more occurrences in the Czech corpus: this concerns only six hypotactic markers out of the nine selected for the analysis, and nine paratactic markers out of the twenty-nine searched for in the data.

It follows from the table that some markers are very frequent, such as *but* (1.94; 342 occurrences) and *however* (0.81; 143), or relatively frequent, i.e. having more than thirty tokens in the Czech corpus, such as *on the other hand* (0.33; 58), *although* (0.32; 57), *yet* (0.32; 57), and *nevertheless* (0.25; 44) (listed by frequency of occurrence). Of these, the paratactic marker *but* is six times more frequent than the most common hypotactic marker *although*. (For similar results based on native speakers' writing habits, cf. e.g. Altenberg 1986.)

The DMs that follow have been excluded from Table 1b, although as with Table 1a above they are all counted in the lines which give total numbers. The exclusion concerns the hypotactic markers *albeit*, *despite the fact (that)*, *except (that)*, *in spite of the fact (that)*, and *notwithstanding*, and the paratactic markers *actually*, *after all*, *all the same*, *alternatively*, *anyhow*, *anyway*, *at any rate*, *at the same time*, *besides*, *by comparison*, *by contrast*, *conversely*, *in any case*, *in comparison*, *in contrast*, *in spite of that*, *instead*, *on the contrary*, *on the other side*, *oppositely*, and *or else*. With markers such as *albeit*, *notwithstanding* and *oppositely*, zero occurrence is not at all surprising, since, as stated in Altenberg (1986), these markers are not likely to appear in any corpus of contemporary English.

Concerning the overall frequency of occurrence and the distribution of individual types of markers, it can be postulated that, as with the choice of causal DMs, students do not frequently resort to the whole repertoire of DMs at their disposal; the markers novice writers usually use when expressing contrastive relations (15 types) is broader than that of causal markers (10 types); this difference can be influenced by a wider availability of contrastive (38 types) than of causal markers (18 types) and also by students' ability to use a relatively wider repertoire of causal markers, i.e. ten of 18 different types.

As regards the differences between the Czech and German corpora, all the markers highlighted in Table 1b are more typical of the Czech rather than the German corpus; this concerns the hypotactic DMs *even if* (0.09; 16), *even though* (0.17; 30), *though* (0.18; 32) and the paratactic *though* (0.10; 18); *even though* is shown in Example 5, together with one token of *although*, i.e. the most typical hypotactic marker of all (0.32; 57) in the Czech corpus:

- (5) *Even though pictures may refer to some kind of socially accepted priorities or stereotypes, their creators cannot be sure to evoke the same reactions in every reader. Various co-interpretations are possible although at least some accompanying text is always present.*

(BrnoCorpusLing 5)

As for the frequency rate of contrastive DMs, Czech novice writers use considerably more frequently the hypotactic marker *though* (often considered a more informal equivalent of *although*; cf. e.g. Gethin 1992: 67) and the paratactic markers *but*, *yet* and *on the other hand*, the last one possibly under the influence of the similar Czech phrase *na druhé straně*. Example 6 illustrates the marker *though* when used in the hypotactic relation, which, according to Biber et al. (1999: 850), vastly outnumbers its use in the paratactic relation (cf. Example 7). Example 6 also illustrates a mistake, namely the use of a comma after the hypotactic *although*. This type of mistake occurs in both the Czech and the German corpora, although not frequently. Example 8, which comprises the DMs *but* and *on the other hand*, provides evidence that some students tend to cumulate certain markers, sometimes even those expressing the same semantic relation.

- (6) *Though the performances were not filmed again, the final discussion was very interesting and lively, as the students fully identified with their new identities and situations. Although, the organization of the project in this mixed ability class was rather complicated and difficulties were also encountered with making the students start the activities, finally, they got fully involved in and enjoyed the project.*  
(BrnoCorpusMeth 4)
- (7) *Structure is present in particular conceptual spheres and there is a system into which ideas denoted by vocabulary can be classified and organized. Other views, though, were brought into semantics, for example that of division of meaning into “a series of component functions” (ibid.: 76).*  
(BrnoCorpusLing 1)
- (8) *Advertising has a very important role in every capitalist society. From the economic point of view it involves mainly positive aspects but, on the other hand there are many people that have very strong objections towards it.*  
(BrnoCorpusLing 5)

Finally, it is necessary to note that more differences between the ways in which Czech and German students use causal and contrastive DMs will be commented on in Section 5 when discussing results from the German corpus.

## 5 Results from the German corpus

### 5.1 Causal DMs

In diploma theses written by German students of English (cf. Table 2a), causal relations are expressed by some explicit DMs much more frequently (7.36; 1,292) than in diploma theses written by Czech students (4.20; 740); this difference is especially striking with paratactic markers which German students use in 833 cases (4.75) in contrast to the 402 cases (2.28) in which Czech students use this syntactic group. As for the hypotactic relations which can be expressed by three different markers only, i.e. *since* (1.43; 251), *because* (0.88; 155) and *as* (0.30; 53) (listed by decreasing frequency), there is a great difference, too, because German students unambiguously prefer the marker *since*, which they tend to use repeatedly in order to introduce every other discourse segment (cf. Example 9), while Czech students, who give preference to *because*, use *since* in far fewer cases (0.26; 46). (For the repeated use of *because* in the Czech corpus, cf. Example 4 above.)

- (9) *Note, however, that some linguists consider aspect as equally deictic, since it is concerned with the relation of E to a reference frame R. For this reason, Dorfmueller-Karpusa regards aspect as internally deictic, since it involves the relation of an event to an internal point of reference, and to tense as externally deictic, since it involves the relation of an event to an external point.*

(ChemnitzCorpusMeth 5)

As with the distribution of the paratactic causal markers in the Czech corpus, in the German corpus these DMs are rather unevenly distributed, some of them having lower frequency of occurrence than ten tokens in the corpus, meaning that they are not explicitly listed in Table 2a; however, in the case of German corpus this concerns only the markers *for*, *in consequence*, *now*, *of course* and *somehow*. Of the fifteen different types of paratactic markers searched for during the analysis, ten occur with noteworthy frequency, i.e. ten or more tokens; these DMs, which are all listed below, represent two thirds of all DMs German students apply. It can thus be concluded that German novice writers make use of a slightly wider range of paratactic DMs when expressing causal relations, and, unlike Czech novice writers, they often resort to the DMs *accordingly* (0.20; 35 occurrences), *as a consequence* (0.10; 18), and above all *hence* (0.70; 123) (illustrated in Example 11 below); in addition, German students apply the marker *consequently* more frequently (0.31) than Czech students (0.10). (All the most

striking differences between the German and Czech corpora in which German novice writers give preference to a certain DM are written in bold in Table 2a.)

No. of words in the texts	90,810	84,712	Total: 175,522	
Hypotactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<i>as</i>	0.45	0.14	0.30	53
<i>because</i>	1.22	0.52	0.88	155
<b><i>since</i></b>	<b>0.81</b>	<b>2.09</b>	<b>1.43</b>	<b>251</b>
All hypotactic DMs	2.49	2.75	2.62	459
Paratactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<b><i>accordingly</i></b>	<b>0.26</b>	<b>0.13</b>	<b>0.20</b>	<b>35</b>
<b><i>as a consequence</i></b>	<b>0.10</b>	<b>0.11</b>	<b>0.10</b>	<b>18</b>
<i>as a result</i>	0.09	0.14	0.11	20
<i>consequently</i>	0.14	0.48	0.31	54
<b><i>hence</i></b>	<b>0.43</b>	<b>0.99</b>	<b>0.70</b>	<b>123</b>
<i>(and) so</i>	0.13	0.09	0.11	20
<i>so that</i>	0.11	0.17	0.14	24
<b><i>therefore</i></b>	<b>1.37</b>	<b>1.75</b>	<b>1.55</b>	<b>272</b>
<b><i>thus</i></b>	<b>1.26</b>	<b>1.10</b>	<b>1.18</b>	<b>207</b>
<b><i>then</i></b>	<b>0.26</b>	<b>0.30</b>	<b>0.28</b>	<b>49</b>
All paratactic DMs	4.15/4.24	5.25/5.29	4.68/4.75	822/833
All DMs	6.73	8.04	7.36	1,292

**Table 2a: Most frequent DMs for causal relations in German students' theses**

As can be observed above, some markers are very frequent in the German corpus, such as *therefore* (1.55; 272 occurrences), *since* (1.43; 251), and *thus* (1.18; 207) (listed by frequency of occurrence), or relatively frequent, i.e. having more than thirty occurrences, such as *because* (0.88; 155), *hence* (0.70; 123), *consequently* (0.31; 54), *as* (0.30; 53) and *accordingly* (0.20; 35). In contrast to the Czech corpus, where the most frequent marker of all is the hypotactic DM *because* (0.85), in the German corpus the most frequent paratactic marker is *therefore* (1.55), followed by the hypotactic *since* (1.43) and paratactic *thus* (1.18; cf. Example 11). These findings testify not only that German students apply paratactic DMs, among which *therefore* and *thus* clearly predominate, much more frequently, but also that there is a stronger tendency on the part of German novice writers to apply the natural ordering of discourse segments in academic texts; according to this strategy (cf. e.g. Altenberg 1987) the segment of discourse which introduces new and/or unexpected information or a new aspect within already known information comes second, i.e. after the discourse segment

with known information, which is, of course, the default case with paratactic markers, as shown in Example 10, in which *therefore* introduces a subsequent segment providing new information. (Example 10 also illustrates *but*, one of the two most common contrastive DMs in the German corpus, discussed in 5.2 below.)

- (10) *Certainly, it is advantageous that English is widespread all over the world, has a simple structure, and is very suitable as a common scientific language, but various dangers such as a mishmash of languages, the disadvantage of non-English-speaking scientists, and the reduced variety of languages should not be forgotten. Therefore it is very important to guarantee a certain quality of scientific English, even by foreign authors; otherwise it could degenerate into a pidgin language.*  
(ChemnitzCorpusMeth 14)

Finally, it is necessary to mention that although there are differences between the fields of study in terms of both types and tokens of DMs German students use when expressing causal relations (e.g. a slightly more frequent use of paratactic DMs in methodology-oriented theses), these are less marked than in the case of Czech students. As already stated, these differences can be caused by individual students' preferences and writing habits resulting from overt field-specific instructions provided by teacher(s) of academic writing and/or thesis supervisors.

## 5.2 Contrastive DMs

The results concerning contrastive DMs as used by German novice writers are given in Table 2b. Contrary to my expectation, contrastive relations expressed by hypotactic DMs are less frequent (2.44; 428 occurrences) in the German corpus than those expressed by paratactic markers (4.83; 846), although, as stated above, hypotactic relations are usually marked overtly in written academic discourse and, in addition, subordination is more typical of written discourse. (For an even greater difference between two syntactic groups of contrastive DMs, cf. findings from Czech students' theses in Table 1b above.) The most interesting finding recorded in Table 2b is the uneven distribution of the contrastive DMs under scrutiny (altogether 1,274) in terms of both types and tokens, above all when expressing paratactic relations. Of the thirty-eight different types of contrastive DMs included in the study, sixteen occur with noteworthy frequency, i.e. having ten or more occurrences in the German corpus. However, this concerns only four hypotactic DMs out of the nine selected for the analysis, in contrast to six in

the Czech corpus, and as many as twelve paratactic DMs out of the twenty-nine searched for in the data, in contrast to nine in the Czech corpus.

No. of words in the texts	90,810	84,712	Total: 175,522	
Hypotactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<b><i>although</i></b>	<b>0.69</b>	<b>0.70</b>	<b>0.70</b>	<b>122</b>
<i>though</i>	0.18	0	0.09	16
<b><i>while</i></b>	<b>0.95</b>	<b>0.68</b>	<b>0.82</b>	<b>144</b>
<b><i>whereas</i></b>	<b>0.51</b>	<b>1.02</b>	<b>0.75</b>	<b>132</b>
All hypotactic DMs	2.32/2.40	2.40/2.48	2.36/2.44	414/428
Paratactic DMs	Linguistics	Methodology	Norm. rate	Raw No.
<i>actually</i>	0.11	0	0.06	10
<i>at the same time</i>	0.19	0.01	0.10	18
<i>but</i>	1.55	1.37	1.46	257
<b><i>however</i></b>	<b>1.56</b>	<b>1.55</b>	<b>1.56</b>	<b>273</b>
<i>in contrast</i>	0.14	0.05	0.10	17
<i>instead</i>	0.13	0.01	0.07	13
<b><i>nevertheless</i></b>	<b>0.39</b>	<b>0.81</b>	<b>0.59</b>	<b>104</b>
<i>nonetheless</i>	0.13	0.07	0.10	18
<i>on the contrary</i>	0.12	0.04	0.08	14
<i>on the other hand</i>	0.19	0.06	0.13	22
<i>still</i>	0.20	0.15	0.18	31
<i>yet</i>	0.18	0.16	0.18	31
All paratactic DMs	4.89/5.09	4.30/4.53	4.60/4.83	808/846
All DMs	7.49	7.01	7.26	1,274

**Table 2b: Most frequent DMs for contrastive relations in German students' theses**

It is evident that German students, like Czech students, use some markers, such as *however* (1.56; 273) and *but* (1.46; 257), very frequently or frequently, i.e. in more than one hundred cases, namely *while* (0.82; 144), *whereas* (0.75; 132), *although* (0.70; 122) and *nevertheless* (0.59; 104) (all of which occur much more frequently in the German corpus). If the three most common paratactic markers, i.e. *however*, *but* and *nevertheless* (the last two illustrated in Example 11) are put together, they represent more than two thirds of all paratactic DMs in the corpus; it follows that German novice writers, like Czech students, tend to use a rather limited repertoire of DMs that they know well and are able to use correctly, above all to express paratactic relations.

Example 11 illustrates yet another tendency, namely to introduce almost every other discourse segment with a DM, probably under the influence of overt instructions to apply explicit guiding signals for the prospective readers when organizing written discourse in a clear way. A similar tendency has been recognized in the Czech corpus, too. Apart from two paratactic contrastive DMs often used in the German corpus, namely *but* and *nevertheless*, Example 11 comprises two causal DMs *thus* and *hence*, discussed in 5.1 above.

- (11) *The present perfect progressive is, according to Kirsten, usually used with activity verbs and can thus be viewed as the opposite of the resultative perfect. Nevertheless, the present perfect progressive can also be used with achievement and accomplishment verbs if emphasis is put on the situation itself and not on the ongoing process of the situation. The use of the present perfect progressive in texts is similar to the use of the present perfect, but due to the fact that it is an experiential form it is used to make descriptions of eventualities more vivid for the reader/listener. It can be used to express emotions or a subjective interpretation of the situation, hence it may express sympathy, antipathy, disappointment, irony, etc.*  
(ChemnitzCorpusMeth 5)

The following DMs have been excluded from Table 2b, although, as with Table 2a above, they are counted in the lines which give total numbers of DMs. However, in the case of German students the exclusion concerns slightly fewer markers, namely the hypotactic *albeit*, *despite the fact (that)*, *except (that)*, *in spite of the fact (that)*, and *notwithstanding*, and the paratactic *after all*, *alternatively*, *anyhow*, *anyway*, *at any rate*, *at the same time*, *besides*, *by comparison*, *by contrast*, *conversely*, *in any case*, *in comparison*, *in spite of that*, *on the other side*, *oppositely*, and *or else*. Some of the markers listed occur only sparingly or not at all in the data.

Concerning the overall frequency of occurrence and the distribution of individual types of markers, it can be postulated that, as with the choice of causal DMs, German students do not frequently resort to the whole repertoire of DMs at their disposal; the repertoire of markers novice writers usually use when expressing contrastive relations is slightly broader (16 types) than that of causal markers (13 types); the reason for this could be the same as with the Czech students, i.e. a wider choice of contrastive (38 types) than of causal DMs (18 types) and the students' ability to use a relatively wider repertoire of causal markers, i.e. thirteen of 18 different types, especially in German novice writing.



## 6 Some results drawn from native speakers' writing habits

In order to compare the data from the present investigation with data drawn from native speakers' use of the selected DMs, some results mentioned in Biber et al. (1999: 887) have been used. Table 3, in which the average frequency rate of the selected DMs per 1,000 words is given, compares the results drawn from the Czech and German corpora included in the investigation with those given in Biber et al. (1999: 887); consequently, only those causal and contrastive DMs that are listed in the latter source are mentioned.

Causal and contrastive DMs in paratactic relations	Czech Ss (theses)	German Ss (theses)	Native speakers of English (RAs)
<i>hence</i>	0.01	<b>0.70</b>	0.1
<i>so</i>	<b>0.32</b>	0.11	0.2
<i>then</i>	0.09	0.28	<b>0.4</b>
<i>therefore</i>	0.68	<b>1.54</b>	0.6
<i>anyway</i>	0.02	0.01	<b>0.05</b>
<i>however</i>	0.81	<b>1.56</b>	1.1
<i>nevertheless</i>	0.25	<b>0.59</b>	0.1
<i>on the other hand</i>	<b>0.33</b>	0.13	0.1
<i>though</i>	<b>0.10</b>	0.02	0.05
<i>yet</i>	<b>0.32</b>	0.18	0.1

**Table 3: Comparison of frequency per 1,000 words between diploma theses written by novice non-native writers and research articles by native speakers of English**

All the results highlighted in Table 3 indicate the highest frequency rate of a certain DM in a particular corpus in comparison with the other corpora. These findings clearly indicate that when writing academic texts native speakers of English tend to use most of the markers less frequently than the novice non-native writers included in the study, which, in my opinion, is due above all to the fact that the repertoire of DMs used by the former is much broader than that applied by non-native speakers in the writing of diploma theses. There are, of course, differences between writers in the two discourse communities caused by field-specific advice and overt instructions provided by thesis supervisors and teachers of academic writing. (For similar conclusions based on the study of German and Finnish writers' habits, cf. Wagner 2011 and Ventola & Mauranen 1991 respectively.)

## 7 Conclusion

It is worth summarizing now some of the most typical problems Czech and German students face when using DMs to express causal and contrastive relations in academic texts:

1. Some students are not able to distinguish between hypotactic and paratactic DMs (e.g. *though* as a conjunction and *though* as an adverb);
2. Some students tend to overuse certain selected DMs while introducing every other discourse segment with an explicit marker;
3. Some students tend to apply certain DMs repeatedly (e.g. the repeated use of the DMs *because* and *since*, discussed above);
4. Most students use only a limited repertoire of the selected DMs, which they mostly know well and thus prefer to use;
5. Some students sometimes use markers which are appropriate in informal rather than formal discourse (e.g. *though* as a subordinator in the Czech corpus, cf. Table 1b above).

When expressing relationships between adjacent or more distant segments of discourse, causal and contrastive DMs are used above all to organize discourse, which is important in all written academic discourse, including diploma theses written by novice non-native writers. Although there are some differences between the works of students writing their theses in different fields of study and in particular between students from the two discourse communities, Czech and German students, the results clearly show that it is crucial to pay sufficient attention to the study of DMs and their appropriate use in academic discourse. Hence, apart from providing students with lists of DMs that include their semantic functions and the use of mostly gap-filling exercises for practice, more time should be spent on (1) recognizing DMs in texts and identifying their semantic meaning in a given context, (2) using substitution exercises in different genres, (3) providing overt instructions on the stylistic appropriateness of certain DMs, (4) working with authentic language corpora, and (5) practising students' written performance, thus enhancing their self-confidence and promoting their communicative language competences. These competences are becoming more and more important now that English has become the lingua franca of international academic communication and the process of internationalization of scholarship opens up new challenges for our understanding of academic communication (Duszak 1997).

## References

- Altenberg, B. (1986) 'Contrastive linking in spoken and written English.' In: Tottie, G. and Bäcklund, I. (eds) *English in Speech and Writing. A Symposium*. Uppsala: Almqvist and Wiksell. 13-40.
- Altenberg, B. (1987) 'Causal ordering strategies in English conversation.' In: Monaghan, J. (ed.) *Grammar in the Construction of Texts*. London: Francis Pinter Publishers. 50-64.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (eds) (1999) *Longman Grammar of Spoken and Written English*. London: Longman.
- Bublitz, W. (1988) *Supportive Fellow-Speakers and Cooperative Conversations. Discourse Topics and Topical Actions. Participant Roles and 'Recipient' Action on a Particular Type of Everyday Conversation*. Amsterdam and Philadelphia: John Benjamins.
- Bublitz, W. (1999) 'Introduction: Views on coherence.' In: Bublitz, W., Lenk, U. and Ventola, E. (eds) *Coherence in Spoken and Written Discourse*. Amsterdam: John Benjamins. 1-7.
- Chamonikolasová, J. (2005) 'Comparing the structures of academic texts written in English and Czech.' In: Huttová, M., Böhmerová, A., Keníž, A. and Tandlichová, E. (eds) *Slovak Studies in English I*. Bratislava: Comenius University. 77-84.
- Clyne, M. (1987) 'Cultural differences in the organization of academic texts: English and German.' *Journal of Pragmatics* 11/2, 211-247.
- Čmejrková, S. and Daneš, F. (1997) 'Academic writing and cultural identity: The case of Czech academic writing.' In: Duzsak, A. (ed.) *Culture and Styles of Academic Discourse*. Berlin and New York: Mouton de Gruyter. 40-62.
- Dontcheva-Navratilova, O. (2007) 'On coherence in written discourse.' In: Schmied, J., Haase, C. and Povolná, R. (eds) *Complexity and Coherence. Approaches to Linguistic Research and Language Teaching*. REAL Studies 3. Göttingen: Cuvillier Verlag. 127-145.
- Dontcheva-Navratilova, O. (2012) 'Cross-cultural differences in the construal of authorial voice in the genre of diploma theses.' In: Berkenkotter, C., Bhatia, V. K. and Gotti, M. (eds) *Insights into Academic Genres. Linguistic Insights. Studies in Language and Communication*. (Vol. 160). Bern: Petr Lang. 301-328.
- Duzsak, A. (1997) 'Cross-cultural academic communication: A discourse-community view.' In: Duzsak, A. (ed.) *Culture and Styles of Academic Discourse*. Berlin and New York: Mouton de Gruyter. 11-39.
- Dušková, L., Strnadová, Ž., Knittlová, D., Peprník, J. and Tárnyiková, J. (1988) *Mluvnice současně angličtiny na pozadí češtiny*. Praha: Academia.
- Flowerdew, L. (2004) 'The argument for using English specialized corpora to understand academic and professional language.' In: Connor, U. and Upton, T. (eds) *Discourse in the Professions*. Amsterdam and Philadelphia: John Benjamins. 11-33.
- Fowler, R. (1986) *Linguistic Criticism*. Oxford: Oxford University Press.
- Fraser, B. (1990) 'An approach to discourse markers.' *Journal of Pragmatics* 14/3, 383-395.
- Fraser, B. (1999) 'What are discourse markers?' *Journal of Pragmatics* 31/7, 931-952.
- Galtung, J. (1985) 'Struktur, Kultur und intellektueller Stil.' In: Wierlacher, A. (ed.) *Das Fremde und das Eigene*. München: Iudicum Verlag. 151-193.
- Gethin, H. (1992) *Grammar in Context. Proficiency Level English*. 3<sup>rd</sup> ed. Edinburgh: Nelson.

- Halliday, M. A. K. and Hasan, R. (1989) *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. 2<sup>nd</sup> ed. Oxford: Oxford University Press.
- Kortmann, B. (1991) *Free Adjuncts and Absolutes in English*. London: Routledge.
- Leech, G. and Svartvik, J. (1994) *A Communicative Grammar of English*. 2<sup>nd</sup> ed. London: Longman.
- Mauranen, A., Hynninen, N. and Ranta, E. (2010) 'English as an academic lingua franca: The ELFA project.' *English for Specific Purposes* 29/3, 183-190.
- Mur-Dueñas, P. (2008) 'Analysing engagement markers cross-culturally: The case of English and Spanish business management research articles.' In: Burgess, S. and Martín-Martín, P. (eds) *English as an Additional Language in Research Publication and Communication. Linguistic Insights. Studies in Language and Communication*. Vol. 61. Bern: Peter Lang. 197-213.
- Povolná, R. (2007) 'Aspects of coherence in spoken discourse.' In: Schmied, J., Haase, C. and Povolná, R. (eds) *Complexity and Coherence: Approaches to Linguistic Research and Language Teaching*. REAL Studies 3. Göttingen: Cuvillier Verlag. 107-125.
- Povolná, R. (2009) 'Exploring interactive discourse markers in academic spoken discourse.' In: Dontcheva-Navratilova, O. and Povolná, R. (eds) *Coherence and Cohesion in Spoken and Written Discourse*. Newcastle: Cambridge Scholars Publishing. 60-80.
- Povolná, R. (2012) 'Cross-cultural differences in the use of discourse markers by Czech and German students of English in the genre of Master's theses.' In: Berkenkotter, C., Bhatia, V. K. and Gotti, M. (eds) *Insights into Academic Genres. Linguistic Insights. Studies in Language and Communication*. (Vol. 160). Bern: Petr Lang. 329-351.
- Schmied, J. (2011) 'Academic writing in Europe: A survey of approaches and problems.' In: Schmied, J. (ed.) *Academic Writing in Europe: Empirical Perspectives*. REAL Studies 5. Göttingen: Cuvillier Verlag. 1-22.
- Seidlhofer, B. and Widdowson, H. G. (1999) 'Coherence in summary: The contexts of appropriate discourse.' In: Bublitz, W., Lenk, U. and Ventola, E. (eds) *Coherence in Spoken and Written Discourse*. Amsterdam: John Benjamins. 205-219.
- Stašková, S. (2005) 'Options of identity: Authorial presence in research article abstracts.' In: Huttová, M., Böhmerová, A., Keníž, A. and Tandlichová, E. (eds) *Slovak Studies in English 1*. Bratislava: Comenius University. 201-207.
- Swales, J. M. (2004) *Research Genres. Explorations and Application*. Cambridge: Cambridge University Press.
- Taboada, M. (2006) 'Discourse markers as signals (or not) of rhetorical relations.' *Journal of Pragmatics* 38/4, 567-592.
- Ventola, E. and Mauranen, A. (1991) 'Non-native writing and native revising of scientific articles.' In: Ventola, E. (ed.) *Functional and Systemic Linguistics. Approaches and Uses*. Berlin and New York: Mouton de Gruyter. 457-492.
- Wagner, S. (2011) 'Concessives and contrastives in student writing: L1, L2 and genre differences.' In: Schmied, J. (ed.) *Academic Writing in Europe: Empirical Perspectives*. REAL Studies 5. Göttingen: Cuvillier Verlag. 23-48.