His eyes narrowed — her eyes downcast: contrastive corpus-stylistic analysis of female and male writing[1]

Anna Čermáková — Lenka Fárová (Praha)

ABSTRACT

This study presents a comparative and cross-linguistic (English and Czech) examination of female and male characters in contemporary British and Czech fiction texts written by female and male authors. We approach the texts within the corpus stylistics framework using four broadly comparable corpora of British and Czech fiction. Our focus is on the most frequent nouns occurring with the possessive pronouns *his* and *her* and *jeho* and *její,* respectively. By examining the nouns we look at how female and male characters are linguistically represented: what are the main differences between their representation and to what extent they are similar. We also aim to look at whether the author's gender plays a role in the process of characterisation. Further, we aim to pinpoint the differences in characterisation by comparing the British and Czech fiction texts.

KEYWORDS

gender and language, gender in literature, corpus stylistics, comparable corpora

1. INTRODUCTION

In 1922 the Danish linguist Otto Jespersen made one of the first attempts to describe gender-based differences in language use (Jespersen, 1922). Jespersen (and many others to follow) proposes, as Deborah Cameron notes in her blog: "a very traditional view of women's writing — that it differs from men's in being less forceful, less daring, less logical in its structure and less individual in its style".[2] Cameron (2016) further notes that Jespersen sees women as "linguistically less innovative and less adventurous than men":

> "Women move preferably in the central field of language, avoiding everything that is out of the way or bizarre, while men will often coin new words or expressions." (Jespersen, 1922, 248)

Jespersen proposes that in terms of syntactical complexity, men organise their argument in an elaborate way using various logical connectors — he compares male sentences to "a set of Chinese boxes, one within another" (Jespersen, 1922, 252) —

2   https://debuk.wordpress.com/2016/03/06/do-women-and-men-write-differently/ (accessed 17.1.2017).

while women in Jespersen's view rely mostly on a simple coordination (i.e. *and*) and he uses another poetic comparison to a "a set of pearls joined together on a string of ands and similar words" (ibid.). These ideas, though based on very little evidence, went basically unquestioned for a long time supporting some other popular stereotypes.

Today, it seems to be easy to conduct a reasonably large study to confirm or refute these arguments. A simple search for one of the most frequent logical conjunctions *because* in fiction texts written by men and women[3] shows that *because* is more often used by female writers (and the difference is statistically significant) and the conjunction *and* is indeed used more often by women than men (with statistically significant difference)[4] — this could lead to a hasty simple conclusion that in general women actually use more conjunctions than men. But what, if anything, does that tell us about their writing styles?

This topic continually fascinates both researchers and the general public with the underlying assumptions that there are, indeed, differences between how women and men speak and write. There are even specialised online applications that based on your writing will guess your gender.[5] As Cameron (2016), who has followed an online discussion of the users of this application, notes:

> "Women who'd been misidentified as men often put this down to being ex-tomboys or geeks who had no truck with 'girly' things: none of them seemed offended by being told they wrote like men, and sometimes they appeared to be flattered. Men who were miscategorized as women, by contrast, more often expressed bafflement, annoyance or discomfort. They also got teased by other people in the comments: had they been writing poetry again? Were they secretly gay?" (Cameron, 2016)

As Cameron (ibid.) further notes, these comments and opinions show that "gender isn't just a difference, it's a hierarchy". Cameron refers to Criado-Perez's book *Do It Like a Woman* (2015), which discusses the common "basic deficit model, in which men set the standard of excellence and whatever women do is somehow deficient, weak and inferior" (Cameron, 2016).

Research into language and gender is vast (for a very brief overview see Section 2.1, for a variety of topics being researched see e.g. Holmes and Meyerhoff, 2003) and includes linguistic considerations of gender in literary texts. Gender in literary texts is traditionally approached either as "a comparison of the fiction created by male and female authors and is typified by the search for 'the female sentence'

---

3   For data see Section 3.

4   *because* occurs in our corpus of female writing (for data, see Section 3) with the normalised frequency of 945.6 ipm while in male writing it occurs with the frequency 792.9 ipm and this difference is significant at p < .001 (crit. 10.82757) (X^2 = 78.85954). *and* occurs in female writing with the relative frequency of 2.672% and in male writing with the frequency 2.572%, the difference is significant at p < .001 (crit. 10.82757) (X^2 = 114.09645)

5   E.g. http://www.hackerfactor.com/GenderGuesser.php

or a specifically female style of writing" or as "a study of the uses to which the linguistic gender system of different languages has been put in literary works" (Livia, 2003, 142). Linguistic analysis of literary texts far too often "focuses on the micro level of text construction, leaving such 'literary' concerns as plot, character, theme, and moral or ideological point to literary theorists" (Livia, 2001, 8). However, as Livia (ibid.) notes, an important overlap exists in the fields of narratology and stylistics. At the same time, we can note a very significant shift in narratology in recent years: "the recognition that literary narrative fiction can be defined not by event but by character" (Stockwell and Mahlberg, 2015, 129).

In this study, by focusing on characterisation, we aim to examine the representation of female and male characters to see how/whether they differ. We aim to do so in British and Czech contemporary literature to see whether there are any specific features that would in this respect distinguish these two literary traditions. Our analysis is focused on nouns that are in our corpora frequently associated with male and female characters. "[T]hat male and female characters in fiction receive very different treatment is not particularly controversial, but the claim that women's writing differs in some essential way from that of men is more tendentious" (Livia, 2003, 144), therefore we will also aim to account for the differences (as well as sameness) between female and male authors.

The structure of the paper is as follows: Section 2 briefly introduces the theoretical background of language and gender research and gender in literary studies. It further briefly discusses the importance of genre considerations in analysing literary texts and introduces our corpus stylistic approach. Section 3 discusses our comparable corpus of British and Czech fiction. Section 4 explains our methodology. Sections 5 and 6 present the data analysis. Section 7 offers conclusions and discusses directions for further research.

2. THEORETICAL BACKGROUND

2.1 LANGUAGE AND GENDER

It was Robin Lakoff (1975) who introduced the view of "male dominance" into linguistics. She claimed that men dominate women through language and this view was characteristic of the early academic feminist debate in linguistics (for an overview see e. g. Baker, 2014, 2–3 and Wodak, 2015). Later approaches (late 80s and early 90s) started to emphasize the differences between women and men rather than male dominance (Tannen, 1990). This 'difference paradigm' is still prevalent in popular culture; however, in the academic debate it has also been strongly disputed (Cameron, 2007). Within the post-structuralist paradigm, it has been proposed that the connection between sex and gender is less direct and Butler in her influential feminist work *Gender Trouble* (1990) sees gender as 'performative'. Feminist efforts have certainly had an influence on language use and certain overtly sexist language use has been somewhat reduced. However, as Baker (2014, 5) says, this does not mean that sexist language no longer exists, only that "sexist discourse has become increasingly more complex, sophisticated and ambiguous".

Current research (esp. social constructionism) sees gender as "linguistic dealings *with* (individual, and groups of) women, men, boys and girls [...] what is said *to* them and, more importantly what is said and written *about* them" (Wodak, 2015, 699). It stresses that 'gender' is not to be perceived as a dichotomy between 'female' and 'male' sex. As Wodak puts it:

> "By focusing only on the biological sex, the whole group of women is perceived as homogenous and contrasted to the whole group of men — [...] without taking differences in these groups into account." (Wodak, 2015, 698).

The importance of other social variables such as race, ethnicity, age etc. is being stressed. However, as Livia (2001) points out, from the linguistic point of view these categories are not the same, "whether or not the author considers gender an important variable, the structure of language makes it a required information" (Livia, 2001, 36). Depending on the language, there is a number of linguistic features where it is difficult, or even impossible, to avoid "gender", e.g. possessive pronouns in English, or verbal endings in Czech.

2.2 GENDER IN LITERARY STUDIES

The connection between language and gender is of interest not only to linguists but also to literary scholars (for a brief overview see e.g. Showalter, 2009, xi — xxix). Work by feminist thinkers in the 1970s (e.g. Cixous, Irigaray and Kristeva in France or Showalter in the USA) had prepared the ground and "by the early 1980s it was clear that feminist literary criticism and attention to female writers had gained institutional legitimacy" (Wolfson, 1997, 12). Feminist literary criticism does not have a uniform perspective on literature in the same way that feminist studies represent many various approaches to what they arguably have in common, i.e. "looking at how both overt and covert practices in different societies and groups within those societies function to empower many men rather than many women in a range of ways" (Sunderland, 2011, 8).

One of the ways of looking at linguistic representation of gender in fiction texts is to examine character identity. Jockers and Kiriloff (2016) conducted a quantitative study of gender and character agency using a large corpus of 19th century fiction texts (3 329 novels), "[a] study of character action may serve as a proxy to not only demarcate character types, but also to investigate what behaviors, and *types* of behaviors, were conventionally aligned with different groups of characters" (Jockers & Kiriloff, 2016). A large quantitative study such as this makes it possible to look at textual trends and confront them with other scholarly research: e.g. Welter (1966) stresses that 'submission' was the feminine virtue in the 19th century and Markovits (2006, 103) considers the Victorian period characterised by "women's limited sphere of action". Gilbert and Gubar in *The Madwoman in the Attic* (1979) notoriously claim that during the Victorian period women were categorized as either angels or monsters depending on how well they conformed to social norms. However, the appearance of significant female writers during this period also suggests that gender roles are developing and changing.

Jockers and Kiriloff (2016) analyse gendered pronouns and verbs while also considering possible effects of literary genre, Gothic novel in their case. A similar study was conducted by Baylog et al. (2014), which examined gendered pronouns (*she* and *he*) and the verbs following them. As both studies show, there is a "strong correlation between character gender and verbs in the 19th century novel" (Jockers & Kiriloff, 2016). This finding suggests that "representations of behavior, or agency, understood in terms of the kinds of actions that are associated with particular pronouns, are an important element of characterization" (Jockers & Kiriloff, 2016). Their findings also indicate that "what characters are doing is a key component of how we understand them" while "other elements of character identity, such as speech patterns and visual appearance" (ibid.) seem in their statistical analysis less indicative.

When we examine the verbs that Jockers and Kiriloff and Baylog et al. identified as the most significantly associated with gender, we can indeed see that they confirm some of the stereotypes suggested by literary criticism. In Jockers and Kiriloff's study the verbs most frequently associated with men are *took, found, rode, walked, come,* while verbs most associated with women are *wept, sat, looked, felt* and *cried.* Baylog et al. give a slightly longer list: the "male" verbs are *strode, took, muttered, manded, shouted, rode, tined, termined, shook* and *clared*; and the "female" verbs were *countenance, gaze, sobbed, bless, trembling, face, embrace, knows, wept, blushed, weeping, walked, loved, marry, accept, burst.*

Another quantitative study was conducted by Bamman et al. (2014), where the authors focus on computational modelling of character "personas". Their study suggests an association between some personas and certain genres, and shows that some personas were "clearly gendered" (Bamman et al., 2014, 377). As has been repeatedly stressed, genre considerations are important in the analysis of literary texts, linguistic or other. As Cameron puts it:

"If you find differences between men and women in a sample of fiction where the male texts are mostly thrillers and the female texts are mostly romances, it can be hard to disentangle the effects of gender from those of genre" (Cameron, 2016).

In relation to gender, it has even been claimed that "certain genres were considered more appropriate for male or female authors" (Jockers & Kiriloff, 2016). However, genre, useful as the concept is, is also a highly contested one. Traditional 'genre' distinctions as used in literary studies, where they have been described in terms of conventions of form and content, have become blurred. The label 'genre' is no longer applied only to literary texts and the category is much broader (Freedman & Medway, 1994). As Freedman and Medway note "'genre' has been able to connect a recognition of regularity in discourse types with a broader social and cultural understanding of language in use" (Freedman & Medway, 1994, 2).

Luke (1994, viii) defines 'genre' as "typified rhetorical actions based on recurrent situations". This genre definition is broadly in line with the use of the term in corpus linguistics, where usually three perspectives on texts are considered: register, genre and style (Biber & Conrad, 2009). Biber and Conrad (2009, 2) define the genre perspective as including a "description of the purposes and situational context of a text

variety", with its linguistic analysis focusing on "the conventional structures used to construct a complete text within the variety". It is important to bear in mind that 'genre' is an external label and once a text has this label, it raises expectations tied to the genre categorisation.

It is clear that genre influences not only written style but also the subject matter, which may have a gendered dimension. This obviously does not hold only for written language. The subject matter is governed by the situational context, which in turn governs the way we speak, conforming more or less to the roles we play. Coates (1998, 295) notes "the 'me' that changes a baby's nappy or mashes a banana for a toddler is a different 'me' from the one who participates in a committee meeting". Therefore, the choice of data (already at the level of individual texts) influences the analysis in a very substantial way and the interpretation of the results should always be tied to the analysed data. As Baker (2014, 29) notes in his analysis of the spoken component of the BNC, generally thought of as the representative corpus of English — the speech samples in the corpus are not balanced, i.e. already the data are to a degree skewed: "males were more likely to be recorded in public contexts, while females were more likely to be recorded in private ones".

2.3 APPROACHING FEMALE AND MALE WRITING FROM A CORPUS STYLISTIC PERSPECTIVE

Our study looks at the representation of female and male characters in women's and men's writing from a corpus stylistic perspective. Corpus stylistics analyses literary texts and "employs corpus linguistic methods to support the analysis of textual meaning and the interpretation of texts. As such, corpus stylistics research makes it possible to focus on individual texts and even text extracts — as the places where the aesthetic effects of language are best analyzed" (Mahlberg, 2015, 358, see also Leech and Short, 2007, 11). Quantitative and qualitative methods of corpus linguistics are used with "intrinsic explanatory purpose" and the linguistic analysis engages with the theories and concerns of literary criticism (Mahlberg, 2015, 358). Thus, we use frequency information derived from the texts as our starting point for further analysis.

Our perspective is also cross-linguistic; we examine contemporary British and Czech fiction to see whether there are culturally specific differences. Our research questions are:

How are female and male characters linguistically represented in contemporary English and Czech fiction?

Is this 'representation' different when the author's gender is considered?

In view of the above discussion of 'genre' another important methodological research question arises:

Is it possible to create a truly comparable corpus?

3. DATA

We work with four, broadly comparable, fiction corpora: British women writers (6.5 mil. tokens, BrWW corpus) and British male writers (5.4 mil., BrMW corpus),

Czech women writers (3.6 mil., CzWW corpus) and Czech male writers (8.5 mil., CzMW corpus). The English data are based on a manual selection of texts from the BNC (for the list of individual texts, see Appendix 1) and the Czech data were manually selected from the SYN-v4 corpus (see Appendix 2).

The task of preparing a comparable corpus has proved challenging. In a crosslinguistic analysis comparable corpora are expected to be comparable on the basis of several parameters — these usually involve the size of the corpus, the time when the texts were produced and often also the genre. In our case, we needed to include another parameter, namely the gender of the author of the text. Table 1 shows the composition of our corpora in terms of the size parameters.

Corpus	Size (in tokens)	Number of texts	Number of authors
BrWW	6,453,021	150	131
BrMW	5,440,612	135	98
CzWW	3,550,644	47	37
CzMW	8,518,232	103	103

TABLE 1. Composition of the four corpora in terms of their size.

As Table 1 shows, the four corpora are not of equal size. We selected our texts from much larger corpora (BNC and SYN), where fiction is only one of the components. The upper limit was thus imposed by the availability of texts in these reference corpora. In order to minimize possible idiolect influences, we decided not to include more than three texts by one author. The smallest corpus is our corpus of Czech women writers (CzWW). Although the overall size of the SYN v4 corpus is 4,349,023,692 tokens and the fiction component with adults as target audience contains 125,345,686 tokens, there are only 52 texts written by women, of which 47 were selected.[6] The selection of texts written by men (CzMW corpus) had, on the other hand, "ideal" parameters, where we selected 103 texts by 103 authors.

The most complicated issue in terms of the corpus composition was the question of genre. We have manually added the 'genre' label to each text in our corpora. For English we were mostly guided by the Amazon classification and for Czech by the Kosmas or "book databases" classification.[7] The genres for our English and Czech corpus data overlap only partly and genre overlap between female and male writers is also only partial; for details see Figures 1 to 4. Had we aimed for more exact overlap, our data would have been too small and we also felt that we should respect the existing differences (for discussion, see Section 2.2).

---

6   5 texts were excluded because they are in fact non-fiction or experimental prose.

7   For English, an additional source consulted, where Amazon (www.amazon.com) classification was not clear, was https://www.goodreads.com/. For Czech the sources are available at: www.kosmas.cz, www.databazeknih.cz and www.cbdb.cz. The classification was carried out in October and November 2016.
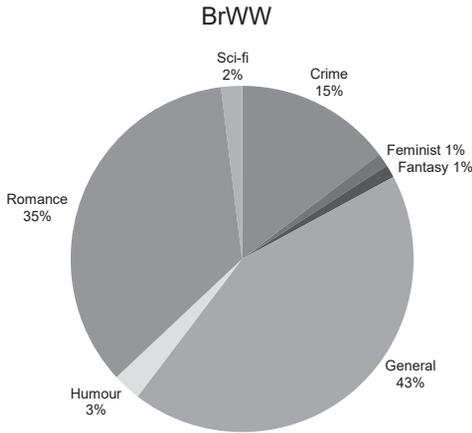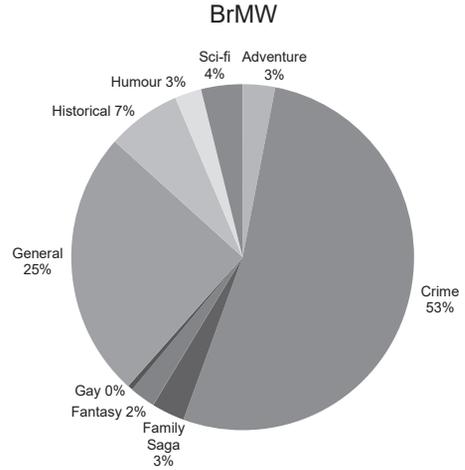
BrWW

Sci-fi 2%
Crime 15%
Feminist 1%
Fantasy 1%
Romance 35%
Humour 3%
General 43%

FIGURE 1. Genre distribution in BrWW corpus.

BrMW

Humour 3%
Sci-fi 4%
Adventure 3%
Historical 7%
General 25%
Crime 53%
Gay 0%
Fantasy 2%
Family Saga 3%

FIGURE 2. Genre distribution in BrMW corpus.

CzWW

Romance 5%
Humour 4%
Historical 3%
General 23%
Autobiography 60%
Fantasy 3%
Crime 3%
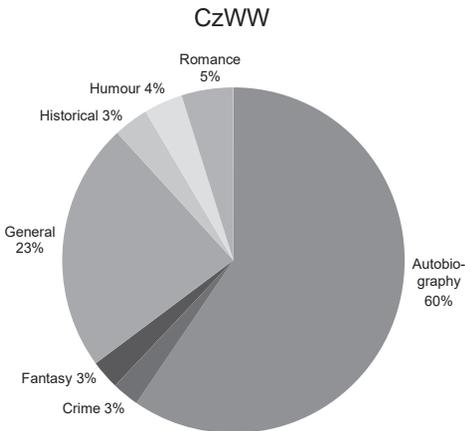
FIGURE 3. Genre distribution in CzWW corpus.

CzMW

Crime 12%
Fantasy 1%
Autobiography 11%
Adventure 3%
Sci-fi 1%
Romance 2%
Humour 5%
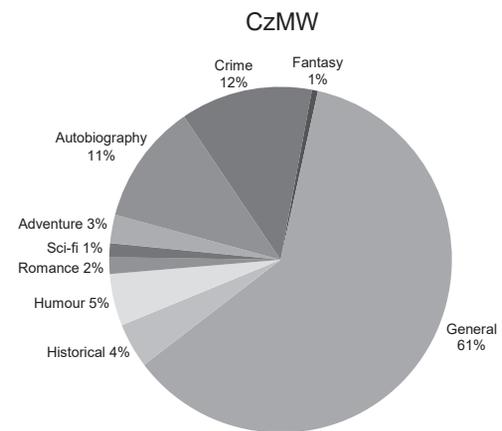Historical 4%
General 61%

FIGURE 4. Genre distribution in CzMW corpus.

In terms of publication date, there are also discrepancies between the English and the Czech component. While the BNC contains texts only up to 1993, with SYN we could aim for more contemporary texts. Thus *contemporary* is for us defined in a very broad sense with some texts going as far back as the 1960s; however, all the older texts selected are still read and many of them have been reprinted. The 'time' variable (i.e. publication date) needs to be considered in relation to gender and the feminist movement as well. The modern cultural and historical development was strikingly different in Britain and the Czech Republic (Czechoslovakia), respectively. While feminism and public discussion of gender issues in Britain have been a topic of main-stream discussion since the 1960s, this was not the case in Czechoslovakia. The atti-

tudes towards gender issues in the Czech Republic are still outside the mainstream political, cultural and academic agenda.[8] Table 2 shows the composition of the four corpora in terms of their publication date; the result is a best compromise between the desirable and the possible.

English	1960–1974	1975–1984	1985–1993	Czech	1960–1975	1976–1989	1990–2000	2001–2015
BrWW	9%	3%	88%	**CzWW**	2%	4%	26%	68%
BrMW	7%	11%	82%	**CzMW**	13%	8%	32%	47%

TABLE 2. The distribution of texts in terms of the publication date in the English and Czech data.

4. METHODOLOGY

Previous quantitative studies, such as Jockers and Kiriloff (2016) and Baylog et al. (2014) (see Section 2.2) have focused on verbs, as verbs are often considered to be the most indicative of how we perceive and understand fictitious characters, i.e. the "doing" is one of the key elements of characterisation. Other studies have acknowledged the importance of other elements of characterisation, including those that are less immediately obvious for the reader. Froehlich (2012) in her study of women characters in Shakespeare's plays, for instance, focuses on possessive pronouns. In our analysis we focus on nouns frequently occurring with gendered possessive pronouns. The aim of the very rough quantitative analysis of our data — we analysed about 50 most frequent nouns[9] occurring with *his/her* and *jeho/její*[10] — is to highlight lexemes suitable for further qualitative analysis, where other variables, such as genre, can be taken into account.

For the analysis of the English language corpora, we analysed occurrences of possessive pronouns *his* and *her* immediately followed by a noun (word form, i.e. thus distinguishing between singular and plural forms) in both BrWW and BrMW cor-

8   This is also apparent in the genre classification. While in BrWW there are two works (HJH, BP8) classified as 'Feminist' (a category, which we kept in our genre classification, see Figure 1), there are also two works (A0L, HGF) that had additional classification labels 'Gay&Lesbian' and 'Jewish&Lesbian'; we subsumed those under the label 'General', however. Other novels (CCM, FET, G06, G1D, H7W, FNT, A6J, F9R) in this corpus had additional labels such as 'Contemporary Women' or 'Women's Fiction' (not to be confused with 'Romance'), and we kept all those under the label 'General'. In BrMW, there was only one novel which carried the label 'Gay&Lesbian" (for further details, see Appendix 1). The category 'Feminist' or 'Gay&Lesbian´ is not present in the categorization of the Czech texts. As far as we know, only one female author (with two texts in CzWW) speaks publicly about her lesbian orientation.

9   The cut-off point is arbitrary. The Czech lists are slightly longer, to include all the nouns with the same frequency at our cut-off point: CzWW thus includes 53 nouns for *její* and 52 nouns for *jeho*; CzMW includes 55 nouns for *její* and 56 nouns for *jeho*.

10   We do not include in our analysis the reflexive possessive pronouns *svůj/svoje*.

pora[11] (Sections 5.1 and 5.2). We have analysed the 50 most frequent nouns occurring immediately after *his* and *her,* respectively. We compared the noun frequencies in the two lists and we have divided the nouns into "masculine" and "feminine", i.e. nouns that occur significantly more often with the pronoun *his* and *her*, respectively.[12] The "top" 50 nouns in the two noun lists overlap to a substantial degree. To compare the frequencies for nouns that occur in one list only (among the "top" 50), we looked up their frequency of occurrence with the other pronoun, e.g. in BrWW *his shirt* occurs 155 times (rank 45 on the "*his*" list), while *her shirt* does not occur among the 50 most frequent nouns occurring with *her*; it occurs only 38 times.

The distribution of *his* and *her* is uneven in both corpora, e.g. in BrWW the more frequent pronoun is *her,* which means that the number of nouns occurring statistically significantly more often with *her* is greater. In order to select nouns that are typically associated with *her* and much less with *his* in this group, we have further compared frequencies of occurrence and selected nouns with substantial difference in frequency (a difference of more than 2.5 times in their raw frequencies), while also considering the number of texts in which they occur, e.g. *her hair* occurs 929 times, while *his hair* occurs only 331 times. The quantitative overview is followed by a case study of the noun *eyes*, which is one of the most frequent nouns occurring with both *his* and *her* in both corpora (Section 5.4).

The analysis of the Czech data focused on possessive pronouns *jeho* [his] and *její* [her] immediately followed by a noun in CzWW and CzMW corpora. Considering the rich Czech morphology and in order to make the analysis of English and Czech comparable, we decided to perform a part of lemma search for Czech, i.e. we searched for the lemma form of the possessive pronouns followed by a noun in all its singular or plural forms,[13] which is indicated in the text by capital letters, e.g. JEHO OČI in fact covers various plural forms: *jeho oči, jeho očích, jeho očí, jeho očím* etc. (see Section 6.1).

We should also bear in mind that there is a considerable difference in the use of possessive pronouns between English and Czech generally. English possessive pronouns (*his, her*) are in a complementary distribution with determiners, deictic and indefinite pronouns; their function overlaps with Czech possessives only partly. English frequently uses possessives when referring to body parts or personal belongings, while in Czech this relation need not be expressed or is expressed in a different way (e.g. dative case, reflexive possessive pronoun) (Dušková, 2006), *cf.* the following examples (http://emsa.ff.cuni.cz/4.21):

a) He dropped **his** stick. — Upustil hůl. (no possessive pronoun in Czech)
b) She raised **her** eyes from **her** book. — Zvedla oči od knihy. (no possessive pronoun in Czech)

---

11   Using http://bncweb.lancs.ac.uk/ search interface.
12   For testing the statistical significance we have used the online tool: http://sigil.collocations. de/wizard.html.
13   [lemma="jeho"] [tag="N..S.*"] resp. [lemma="jeho"] [tag="N..P.*"] using www.kontext.cz search interface.

c) **My** hand was trembling. — Třásla se **mi** ruka. (possessive *my* in English, dative pronoun *mi* in Czech)
d) She pulled on **her** gloves. — Navlékla **si** rukavice. (possessive in English, reflexive *si* in Czech)

Consequently, the overall frequencies of the possessive pronouns in English are naturally much higher, e.g. the pronoun JEHO [his] occurs in the CzWW corpus with the frequency 2,092 ipm[14], while *his* in the BrWW corpus occurs with the frequency 8,957 ipm (for details, see Sections 5 and 6). In practice this means that in Czech we deal with much lower noun frequencies (nouns occurring with gendered pronouns) and the results of our pilot qualitative studies in Czech were not indicative at this point. In other respects the analysis of Czech mirrors the analysis of English. In the next step of our analysis, we aim to construct a "typical" female and male character in fiction texts written by women and men, respectively (Sections 5.3 and 6.3).

5. WOMEN AND MEN IN BRITISH CONTEMPORARY FICTION

5.1 'HIS/HER + NOUN' IN BRWW CORPUS

The pronoun *his* occurs in the BrWW corpus 57,797 times (8,957 ipm), followed by 4,513 different nouns (types) with a total number of occurrences of 46,468. The pronoun *her* occurs twice as often: 115,655 times (17,923 ipm), followed by 4,638 different nouns 62,236 times. The frequency distribution between *his* and *her* is thus strikingly unequal; however, the difference in ratio between '*his* + NOUN' and '*her* + NOUN' is less pronounced, '*her* + NOUN' occurring 1.3 times more often, and the difference in noun types is even in favour of the pronoun *his.* While the difference in noun types is not statistically significant, the difference in the frequency of occurrence is statistically significant for the token value[15] (see Table 3).

	Total	+NOUN (types)	+NOUN (tokens)
his	57,797	4,513	46,468
her	115,655	4,638	62,236

TABLE 3. Pronouns *his* and *her* in BrWW corpus.

Many of the 50 most frequent nouns occurring with *his* and *her* overlap. The three most frequent nouns — *eyes, head, face* — are the same. As explained in Section 4, we have compared the frequencies of the nouns in the two lists and tested them for statistical significance. Table 4 below shows the nouns based on whether they tend to occur with the masculine or feminine possessive pronoun. The figures in brackets in the first column indicate the number of texts in which these nouns occurred.

14   ipm — normalized frequency per million tokens (words including punctuation).
15   X^2 = 2,306.35440 *** at p < .001 (crit. 10.82757)

Nouns occurring significantly more frequently with *his*	No significant difference in frequency between *his* and *her*	Nouns occurring significantly more frequently with *her*
arms (128), *wife* (132), *name* (121), *gaze* (72), *chest* (80), *tone* (74), *words* (89), *expression* (67), *car* (71), *pocket* (81), *shirt* (68), *office* (53), *son* (54)	hand, voice, hands, mouth, arm, back, chair, brother, tongue, attention, desk, work	eyes, head, face, mother, mind, father, heart, hair, feet, body, life, lips, husband, way, fingers, throat, shoulders, breath, sister, cheeks, neck, shoulder, daughter, thoughts, room, legs, stomach, skin, friend, breasts, cheek, teeth, chin, side, parents, lip, family, bag, nose

TABLE 4. "Masculine" and "feminine" nouns in the BrWW corpus (=150 texts) (within the table, the ordering of nouns is based on their frequency; the figure in brackets in the first column indicates the number of texts in which the noun occurred with the possessive *his*).

Considering overall frequency distribution of *his* vs. *her* as discussed above, it is not surprising that there are fewer "masculine" nouns. All of these are frequent and occur in at least a third of the texts (see column 1, Table 4). Thus, the most "masculine" nouns in BrWW are *arms, wife* and *name,* while the most "feminine" nouns are, based on these results, *eyes, head* and *face.* However, *eyes, head* and *face* are also, as mentioned above, the three most frequent nouns occurring with *his.* Therefore, to pinpoint the "feminine" nouns (that is those occurring significantly more often with the pronoun *her,* column 3 in Table 4) more accurately, we have compared the frequencies of occurrence of these nouns with the possessive *her* to their frequency of occurrence with *his* and selected nouns, where we identified substantial differences in their frequency distribution (more than 2.5 times), e.g. *her cheeks* occurs 354 times while *his cheeks* occurs 48 times, see Table 5 below. In the table, the first figure in round brackets is the frequency of occurrence with *her,* the second figure is the frequency with *his. T*he figure in square brackets indicates the number of texts in which the noun occurred with the pronoun *her;* most of them occurred in more than half of the texts.

mind (1323 vs. 506) [139], *heart* (1005 vs. 275) [120], *hair* (929 vs. 331) [135], *husband* (629 vs. 0) [124], *cheeks* (354 vs. 48) [89], *thoughts* (309 vs. 81) [91], *stomach* (288 vs. 71) [84], *breasts* (253 vs. 32) [70], *lip* (219 vs. 50) [64], *bag* (207 vs. 21) [68]

TABLE 5. "Feminine" nouns in the BrWW corpus [=150 texts] (within the table, the ordering of nouns is based on their frequency, the first figure in the round brackets indicates the frequency of the noun with the possessive pronoun *her* and the second number the frequency with *his;* the figures in square brackets indicate the number of texts in which the noun occurred with the possessive *her*).

Even a quick glance at Tables 4 and 5 shows that there are differences in how women writers represent their female and male characters. For the physical description of male characters, *arms, chest* and *shirt* are the most important nouns. It is also the manner of speaking that seems prominent, as nouns like *tone* and *words* are frequent. Male characters often possess or drive a *car* and find themselves in an *office*. The *name* of male characters is also important. For men, *sons* are (statistically) important (but *mother* is still the 10th most frequent noun occurring with *his*), while for women the family/social circle is wider (*mother, father, sister, daughter, parents, family, friend*). The essential characteristics of female characters created by women writers are *mind, heart, thoughts* and body parts: *hair, cheeks, stomach, breasts, lip, throat, shoulders, legs, cheek* and *chin*. The most important female possession seems to be a *bag* (cf. *handbag* in Section 5.2). It seems that these nouns are in line with some of the "traditional" stereotypes, but a detailed collocate analysis (Section 5.4) shows that the differences are much more complex.

5.2 'HIS/HER + NOUN' IN BRMW CORPUS

The pronoun *his* occurs in the BrMW corpus 54,905 times (10,092 ipm), followed by 4,952 different nouns (types) in 44,751 cases. The pronoun *her* occurs 39,189 times (7,203 ipm), followed by 2,973 different nouns 20,779 times. In comparison to the BrWW corpus, where the difference in the frequency of occurrence of *his* vs. *her* was double, in the BrMW corpus it is lower (it is still 1.4 higher in favour of *his*). The difference in ratio between '*his* + NOUN' and '*her* + NOUN' is, however, more than double and the difference in noun types is also substantial (1.6 times in favour of '*his* + NOUN') (see Table 6). All these differences are significant.

	Total	+NOUN (types)	+NOUN (tokens)
his	54,905	4,952	44,751
her	39,189	2,973	20,779

TABLE 6. Pronouns *his* and *her* in BrMW corpus.

Again, many of the most frequent nouns occurring with *his* and *her* overlap. The three most frequent nouns are the same in both corpora: *eyes, head, face*. We have again compared the two lists. Table 7 below shows the nouns occurring statistically significantly more often with *his* vs. *her* and nouns where no significant difference in frequency was identified. Considering the overall frequencies, again unsurprisingly, there are very few nouns (*hair, breasts, handbag, dress*) that occur significantly more often with the possessive pronoun *her*.

OPEN
ACCESS

Nouns occurring significantly more frequently with *his*	No significant difference in frequency between *his* and *her*	Nouns occurring significantly more frequently with *her*
head, eyes, face, hand, hands, wife, father, way, voice, feet, mind, mouth, lips, life, back, arms, arm, body, fingers, chair, name, shoulders, desk, office, nose, pocket, neck, throat, chest, shoulder, son, teeth, knees, breath, car, heart, room, brother, jacket, watch, glass, horse, family, forehead, side	*mother, legs, daughter, sister, parents, cheeks, skin, tongue*	***hair*** (100), ***breasts*** (50), ***handbag*** (43), ***dress*** (44)

TABLE 7. "Masculine" and "feminine" nouns in the BrMW corpus (=135 texts) (within the table the nouns are ordered based on their frequency, the figure in brackets in the third column indicates the number of texts in which the noun occurred with the possessive *her*).

The group of nouns occurring significantly more often with *his* is fairly extensive. We have thus again pinpointed nouns with the most substantial difference in frequency of occurrence between *his* and *her* (see Table 8): *wife* (occurring in 91% of texts), *pocket* (17 times more frequent with *his*, occurring in 66% of texts), *office* (15.2 times more frequent, occurring in 46% of texts), *jacket* (13.9 times more frequent, occurring in 47% of texts) and *horse* (9.4 more frequent, occurring in 19% of texts). The nouns *pocket* and *shirt* are "typical" masculine nouns also in BrWW corpus. Other nouns that occur more than 4 times more frequently with *his* are *desk* and *chest* (see Table 8).

hands (881 vs. 337) [127], ***wife*** (693 vs. 0) [123], ***way*** (606 vs. 183) [125], ***feet*** (599 vs. 207) [122], ***chair*** (301 vs. 96) [97], ***desk*** (243 vs. 52) [74], ***office*** (228 vs. 15) [61], ***nose*** (223 vs. 80) [93], ***pocket*** (222 vs. 13) [89], ***chest*** (210 vs. 50) [94], ***son*** (195 vs. 65) [59], ***teeth*** (195 vs. 72) [89], ***breath*** (187 vs. 56) [88], ***car*** (185 vs. 54) [60], ***jacket*** (167 vs. 12) [64], ***watch*** (167 vs. 44) [73], ***glass*** (161 vs. 43) [67], ***horse*** (159 vs. 17) [26], ***family*** (147 vs. 59) [59], ***forehead*** (145 vs. 53) [67]

TABLE 8. "Masculine" nouns in the BrMW corpus [=135 texts] (within the table, the ordering of nouns is based on their frequency, the first figure in the round brackets indicates the frequency of the noun with the possessive pronoun *his* and the second figure the frequency with *her*; the figures in the square brackets indicate the number of texts in which the noun occurred with the possessive *his*).

Upon examination of the above lists, we can again see that there are some differences in how male authors represent their female and male characters. The male characters are not only described more often but they are portrayed using a greater variety of body parts: *hands, feet, nose, chest, teeth, forehead* — in contrast to female characters, where only *hair* and *breasts* seem to be important. Among male clothes the most prominent is *jacket*. Pockets[16] are mentioned often and male characters seem to fre-

---

16   *his pocket* occurs in 89 texts with three most significant collocates *from, into* and *in*. For comparison, *her pocket* occurs only 13 times in 11 texts.

quently use a *watch* (73 texts) and a *car* (60 texts[17]). Another typical noun associated with male characters is a *glass*[18] (67 texts).[19] Male characters are frequently situated in a work environment (*chair, desk, office*), which is similar in the BrWW. In terms of frequency, the female characters seem to be reduced to a description of their appearance; in addition to their *hair* and *breasts,* their *dress* and *handbag* get noticed.

5.3 "TYPICAL" FEMALE AND MALE CHARACTER IN BRWW AND BRMW CORPORA
There is a telling overlap in the lists of the most frequent nouns determined by the pronoun *his/her* in the two corpora but there are also differences which suggest that some of the stereotypical character portrayals occur in the texts. To confirm the claim, this would, however, require detailed analyses of these nouns. We attempted to construct a "typical" female and male character in these two corpora. We have focused on both shared characteristics and differences. Tables 9 and 10 summarize the nouns that are significantly associated with male and female characters, respectively in BrWW and BrMW corpora.

	'Man' in BrWW	'Man' in BrMW
same characteristics in both corpora	body parts: *chest* possession: *car* clothing: *pocket* work: *office* family: *wife, son*	
differences	body parts: *arms* clothing: *shirt* speech/looking: *gaze, tone, words, expression* other: *name*	body parts: *hands, feet, nose, teeth, forehead* work: *chair, desk* clothing: *jacket* accessories/possessions: *watch, glass, horse* family: *family* accessories: *watch, glass* other: *way, breath*

TABLE 9. "Typical" male in the BrWW and BrMW corpora.

	'Woman' in BrWW	'Woman' in BrMW
same characteristics in both corpora	body parts: *hair, breasts*	
differences	body parts: *cheeks, stomach, lip* family: *husband* accessories: *bag* other: *mind, heart, thoughts*	clothing: *dress* accessories: *handbag*

TABLE 10. "Typical" female in the BrWW and BrMW corpora.

17 *her car* occurs only in 26 texts
18 With most frequent collocates (span -3/3, based on Log-likelihood, not calculated over sentence boundaries): *raised, drained, refilled, wine, down, put, held.*
19 *horse* is also statistically significant but occurs in only 26 texts.

Considering the overall differences in the distribution of *his* and *her* in the two corpora, it is not possible to draw any definite conclusions from these two tables; however, they do represent a good starting point for further study. Based on the frequencies we have analysed, female characters seem to be somewhat flat in fiction written by men, being reduced to only four nouns that occur significantly frequently with the pronoun *her*. This is clearly due to the fact that there are fewer female characters than male: the pronoun *he* occurs in BrMW with the frequency 17,288 ipm, while the pronoun *she* occurs only with the frequency 8,147 ipm.

Male characters, on the other hand, occur frequently with a much greater number of nouns. The overlap between the characterisation of male characters in fiction written by female vs. male authors is bigger than the overlap in female characterisation. Male authors associate more concrete nouns with their male characters (*chair, desk, jacket, glass, watch, horse* etc.), while women writers seem to concentrate more often on ways of speaking and looking (*gaze, tone, words, expression*).

5.4 WHAT DO THE *EYES* TELL US?

For a more detailed analysis, we have selected the noun *eyes*, which is the most frequent[20] noun associated with both *his* and *her* in both corpora. We have analysed the most frequent collocations of *his/her eyes* in the two corpora and then further analysed the collocations for their specific meanings in the texts. The analysis suggests that some of the typical patterns may be genre-tied.

5.4.1 *EYES* IN BRWW

The most frequent collocates[21] of the phrase ***his eyes*** suggest that one of the most frequent semantic prosodies for women writers when using the phrase is to express menace, defending position, aggression or some other discomfort. Of the 50 most frequent collocates, 15 are nearly exclusively used in these contexts,[22] e.g. *glittered, glinted, gleam, blazed, dangerously, flashed*, see examples 1 and 2.

1) *Anger glittered coldly in **his eyes**.* [JY7]
2) *... but hostility still glinted in **his eyes**.* [H9L]

The most frequent collocate is *narrowed* (occurring 103 times in 40 texts),[23] occurring even more frequently than the typical collocate *closed* (92 times in 60 texts), and the most frequent collocate of the phrase *his eyes narrowed* is *dangerously* (example 3).

3) ***His eyes narrowed dangerously*** *and for a split-second Polly held her breath.* [H7W]

---

20 In the BrMW corpus, the noun *eyes* with the pronoun *his* is the second most frequent after *head*.
21 Collocates are calculated within the span -3/3, based on MI score, top 50 collocations considered occurring in at least 5 texts.
22 As opposed to only 3 that are clearly used positively: *crinkling, twinkled, lingering*.
23 The phrase *his eyes narrowed* itself occurring 87 times. Most of the occurrences of this collocate occur in texts classified as 'romance'.

The phrase **her eyes** occurs in a much greater variety of contexts, e.g.: 'surprise/panic' (*widening*, *widened*), 'looking away' (*averting*, *averted*, *downcast*), or 'tears/crying' (*dabbed*, *welled*, *welling*, *pricked*, *tears*). The most frequent collocate is *closed* (334 times in 98 texts) and the second most frequent collocate is *tears* (161 times in 76 texts). However, the most significant collocate (based on MI) is *widening*, which would be typically used in a context of surprise and often negative surprise (example 4).

> 4) *Perdita gazed at him, her white left cheek slowly turning bright scarlet,* **her eyes widening in horror**. [CA0]

5.4.2 *EYES* IN BRMW

In comparison, the phrase *his eyes* has several functions in the texts written by male authors and very few collocates indicate a clear semantic prosody, among them *blazing* (occurring 9 times in 6 texts, mostly in the context of 'anger'), *tears* (31 times in 24 texts) and *fear* (11 times in 9 texts).[24] The most significant (based on MI value) is again *narrowed* and the most frequent *closed* (149 times in 64 texts). The phrase *his eyes narrowed* itself does not have any significant collocates; among the 18 occurrences it is used in a context of menace or attack only 4 times (example 5).

> 5) *When he did,* **his eyes narrowed** *and for a moment Kelly thought he was going to bound across the room and hit her.* [BP7]

The phrase *his eyes* is used in a variety of contexts, e.g.: 'change' (*narrowed, narrowing, widened*), 'movement/fixation' (*screwed* (*up*), *flickered, darting* (*about*), *fixed* (*on*), *lingered, blinking*), 'looking away' (*averted, closed*), or 'menace' (*glazed, blazing*).

    This seems to be somewhat different from the uses of the phrase in the texts written by female authors, where 'menace' is the dominant semantic prosody when male eyes are mentioned in the text. This less clearly delimited meaning of the phrase in the BrMW corpus is also supported by the fact that the phrase is sometimes accompanied by explicit explanation as to what is meant; see examples 6 and 7.

> 6) *The boy sat there, his legs crossed under him, perfectly still, watching Tuan Ti Fo,* **his eyes narrowed, as if trying to understand** *why the old man was smiling.* [GUG]
> 7) *He* **narrowed his eyes** *and* **tried to look tough**. [HWL]

Similarly *her eyes* have only partly overlapping semantic prosodies in BrWW and BrMW corpora. In both, they often indicate that the female character is 'looking away'; *downcast* is the most salient collocate and *averted* is the fourth. In both corpora, *her eyes* are often associated with tears and crying, in BrMW among the top 10 collocates we find: *dabbing, welled, brimming, dabbed, tears*. In addition, *her eyes* and *his eyes* in BrMW also both indicate a 'change' of state and describe eye 'movement/fixation'.

---

24 In BrWW corpus *his eyes* collocates with *tears* with a very low MI value and it does not collocate with *fear* at all.

6. WOMEN AND MEN IN CZECH FICTION

6.1 'JEHO/JEJÍ [HIS/HER] + NOUN' IN CZWW CORPUS

The pronoun JEHO [his] occurs in the CzWW corpus 7,423 times (2,090.61 ipm), followed by 1,551 different nouns (types) with a total number of occurrences of 5,577. The pronoun JEJÍ [her] occurs 5,381 times (1,515.5 ipm), followed by 1,309 different nouns 4,076 times. The frequency distribution between JEHO and JEJÍ is unequal and all the differences are statistically significant (see Table 11).

	Total	+ NOUN (types)	+ NOUN (tokens)
JEHO [his]	7,423	1,551	5,577
JEJÍ [her]	5,381	1,309	4,076

TABLE 11. Pronouns *jeho* [his] and *její* [her] in CzWW corpus.

In order to match the analysis of the Czech data to the English, we have performed a part of lemma search distinguishing between singular and plural word forms (for details, see Section 4). We have again compared the frequencies of the nouns in the 'his' and 'her' lists and tested them for statistical significance. Table 12 below shows the nouns based on whether they tend to occur with the masculine or feminine possessive pronoun. Additionally, since we are dealing with much lower frequencies in Czech than in English, we have further narrowed the group of nouns occurring statistically significantly more often with JEHO (first column in Table 12) and included only nouns that occur in more than 10 texts, with the minimum frequency of occurrence being 5. Thus we are able to make some generalizations and we are not dealing with nouns specific to only few texts.

Unlike in the BrWW, where the female possessive pronouns were more frequent,[25] in the CzWW the male possessive pronoun JEHO is more frequent (see Table 11). It is therefore not surprising that there are more nouns occurring significantly more often with JEHO than JEJÍ. The biggest group of nouns, however, is the group where no statistical significance was found (see Table 12). The nouns occurring with JEJÍ with substantial difference in frequency (more than 2.5 times) are MUŽ [man/husband] (in 47% of texts) and MANŽEL [husband] (43% of texts), and less well distributed over the corpus: NOHY [feet/legs] (only 26% of texts), KLÍN [lap] (23% of texts) and KRÁSA [beauty] (17% of texts) (marked in bold in the third column in Table 12). In this group, there are two lexemes referring to a woman's partner (*husband* was among the significant nouns occurring with *her* in BrWW as well), 'feet', which occurred in BrWW, and 'lap' and 'beauty', which did not occur among the significant "female" nouns in BrWW.

We have further examined the differences among the frequencies of nouns occurring more often with JEHO. The nouns that exhibit the most substantial difference in

25   In BrWW, the personal pronoun *she* is more frequent than *he* as well (19,900 ipm vs. 15,235 ipm).

Nouns occurring significantly more frequently with JEHO	No significant difference in frequency between JEHO and JEJÍ	Nouns occurring significantly more frequently with JEJÍ
OČI [eyes] [33], **ŽENA** [woman/wife] (138 vs. 0) [32], TVÁŘ [face] [31], HLAS [voice] [29], POHLED [look] [28], RUKA [hand] [33], OTEC [father] [26], JMÉNO [name] [27], SLOVA [words] [18], RUCE [hands] [15], DŮM [house] [13], **SLOVO** [word] (26 vs. 0) [12], **MANŽELKA** [wife] (24 vs. 0) [11], PRÁCE [work] [14], VINA [guilt] [11], BLÍZKOST [closeness] (16 vs. 5) [11], **MOZEK** [brain] (14 vs. 0) [10]	TĚLO [body], ŽIVOT [life], MATKA [mother], SYN [son], HLAVA [head], ZÁDA [back], VLASY [hair], RODIČE [parents], DCERA [daughter], SMRT [death], OBLIČEJ [face], PRSTY [fingers], ÚSMĚV [smile], BRATR [brother], RODINA [family], SRDCE [heart], DECH [breath], BYT [flat], PŘÍTOMNOST [presence], ÚSTA [mouth], MÍSTO [place], DĚTI [children], RTY [lips], STRANA [side], VĚK [age], MAMINKA [mom], RAMENO [shoulder], VÝRAZ [expression], POKOJ [room], STAV [state], POSTEL [bed], DÍTĚ [child], MYSL [mind], KŮŽE [skin], SESTRA [sister], VŮLE [will], POHYB [movement], POVAHA [personality]	**MUŽ** [man/husband] (93 vs. 0) [22], **MANŽEL** [husband] (58 vs. 0) [20], NOHY [feet/legs] (20 vs. 7) [12], **KLÍN** [lap] (17 vs. 0) [11], **KRÁSA** [beauty] (15 vs. 0) [8]

TABLE 12. "Masculine" and "feminine" nouns in the CzWW corpus (=47 texts) (within the table, the nouns are ordered by frequency, the figures in square brackets indicate the number of texts; in the first column, the first figure in round brackets indicates the number of occurrences with *jeho* and the second figure the frequency with *její*; in the third column, the first figure indicates the frequency with *její* and the second figure that with *jeho*).

frequency of occurrence (marked in bold in the first column of Table 12) are ŽENA [woman/wife] (68% of texts), SLOVO [word] (26% of texts), MANŽELKA [wife] (23%), BLÍZKOST [closeness] (23% of texts) and MOZEK [brain] (21% of texts). Again, we have two lexemes referring to a partner. We also have one abstract noun — BLÍZKOST [closeness] — which occurs mostly in the phrase *v jeho blízkosti* [close to him] and in a majority of cases it refers to a physical closeness (both abstract and concrete) of a man and a woman.[26]

6.2 'JEHO/JEJÍ [HIS/HER] + NOUN' IN CZMW CORPUS
The pronoun JEHO [his] occurs in the CzMW corpus 17,274 times (2,028 ipm), followed by 2,968 different nouns (types) with a total number of occurrences of 12,689. The pronoun JEJÍ [her] occurs 8,522 times (1,000 ipm), followed by 1,858 different nouns 6,443 times. The frequency distribution between JEHO and JEJÍ is again unequal and all the differences are statistically significant (see Table 13).

26 In half of the cases (6 out of 12), there is an erotic context and in two cases the context is negative, e.g. *stačilo sedm měsíců jeho blízkosti* [seven months of him being close was enough].

	Total	+ NOUN (types)	+ NOUN (tokens)
JEHO [his]	17,274	2,968	12,689
JEJÍ [her]	8,522	1,858	6,443

TABLE 13. Pronouns *jeho* [his] and *její* [her] in CzMW corpus.

We have again compared the two lists and tested them for statistical significance. Table 14 below summarizes the nouns based on whether they tend to occur with the masculine or feminine possessive pronoun. Additionally, we have again set a limit of a minimum number of texts in which the nouns (occurring significantly more often with JEHO) have to occur to be considered. Since CzMW is much larger than CzWW, the minimum number of texts was set to 20[27].

Nouns occurring significantly more frequently with **JEHO**	No significant difference in frequency between **JEHO** and **JEJÍ**	Nouns occurring significantly more frequently with **JEJÍ**
ŽENA [woman/wife] (64), ŽIVOT [life] (57), OTEC [father] (55), JMÉNO [name] (57), HLAVA [head] (62), RUKA [hand] (56), ZÁDA [back] (41), SLOVA [words] (49), POHLED [look] (42), MANŽELKA [wife] (32), SYN [son] (40), SMRT [death] (33), MÍSTO [place] (40), DCERA [daughter] (28), PŘÍTEL [friend] (26), MOZEK [brain] (23), DUŠE [soul] (28), RODINA [family] (31), PRÁCE [work] (28), PŘÁTELÉ [friends] (26), SLOVO [word] (37), NÁZOR [opinion] (24), PŘÍTOMNOST [presence] (31), BRATR [brother] (21), OSUD [fate] (24), KANCELÁŘ [office] (23), ŘEČ [speech] (21), POZORNOST [attention] (23), PŘÍPAD [case] (23), STRANA [side] (25), BOK [hip/side] (22), POSTEL [bed] (23), ODCHOD [leave] (24), PŘEDSTAVY [ideas] (21)	TVÁŘ [face], HLAS [voice], OČI [eyes], TĚLO [body], MATKA [mother], OBLIČEJ [face], RODIČE [parents], BYT [flat], RTY [lips], RUCE [hands], POKOJ [room], ÚSTA [mouth], RAMENO [shoulder], DECH [breath], KONEC [end], POSTAVA [figure], ÚSMĚV [smile], VĚK [age], OBSAH [content], LÁSKA [love], NOHY [legs/feet], CHOVÁNÍ [behaviour], PRSTY [fingers], DŮM [house], DÍTĚ [child], KLÍN [lap], OTÁZKA [question], TATÍNEK [daddy]	MANŽEL [husband] (46), MUŽ [man/husband] (36), VLASY [hair] (22), KRÁSA [beauty] (18), ŇADRA [breasts] (12), VŮNĚ [smell] (16), PRSA [breasts] (12)

TABLE 14. "Masculine" and "feminine" nouns in the CzMW corpus (=103 texts) (within the table, the nouns are ordered by frequency, the figure in round brackets indicates the number of texts).

---

27 Thus the following nouns were not considered: MILOST [grace] (6 texts), PÁN [lord] (16), SESTRA [sister] (15), MYSL [mind] (18), SPOLEČNÍK [companion] (11), MYŠLENKY [thoughts] (18), VŮLE [will] (17).

Since the frequency of 'JEHO + NOUN' is in the CzMW corpus so much higher, naturally most of the nouns occur statistically significantly more often with the masculine pronoun. The group of nouns where no significant difference was found is fairly extensive as well, and very few nouns occur significantly more often with the female possessive: MANŽEL [husband] (45% of texts), MUŽ [man/husband] (35% of texts), and KRÁSA [beauty] (18% of texts) are the same as in the CzWW corpus. In the CzMW the nouns: VLASY [hair] (21% of texts), ŇADRA [breasts] (12% of texts), VŮNĚ [smell] (16% of texts), and PRSA [breasts] (12% of texts) additionally occur. Among the "feminine" nouns in CzMW are, in addition to 'hair' and 'smell', two lexemes referring to a male partner and two lexemes referring to 'breasts'.

We have further compared the frequencies of occurrence for the nouns occurring statistically significantly more often with the male possessive pronoun. Table 15 summarizes nouns where we have identified a substantial difference in their frequencies; all these nouns also occur in more than 20% of texts. The most "salient" nouns are again the two lexemes referring to a partner, i.e. ŽENA [woman/wife] and MANŽELKA [wife], and MOZEK [brain] and SLOVO [word]; these nouns are the same in the CzWW corpus. In comparison with the BrMW, in CzMW a social circle of friends is significant (PŘÍTEL [friend], PŘÁTELÉ [friends]) but there are no family members other than 'wives'. Apart from the 'brain', there are no other body parts, while in BrMW we had *hands, feet, nose, chest, teeth* and *forehead,* and no nouns referring to clothing (in BrMW we had *jacket, watch, pocket*). There are nouns referring to work environment: PRÁCE [work] and KANCELÁŘ [office] (in BrMW we had *chair, desk, office*). And additionally, there is a group of abstract nouns: DUŠE [soul], NÁZOR [opinion], OSUD [fate], and PŘEDSTAVY [ideas].

Substantial difference in frequency distribution

ŽENA [woman/wife] (282 vs. 0) [64], MANŽELKA [wife] (81 vs. 0) [32], PŘÍTEL [friend] (58 vs. 12) [26], MOZEK [brain] (54 vs. 7) [23], DUŠE [soul] (53 vs. 13) [28], PRÁCE [work] (47 vs. 14) [28], PŘÁTELÉ [friends] (46 vs. 5) [26], SLOVO [word] (45 vs. 0) [37], NÁZOR [opinion] (44 vs. 0) [24], OSUD [fate] (39 vs. 10) [24], KANCELÁŘ [office] (38 vs. 7) [23], ŘEČ [speech] (35 vs. 0) [21], PŘEDSTAVY [ideas] (30 vs. 6) [21]

TABLE 15. "Masculine" nouns in the CzMW corpus [=103 texts] (within the table the ordering of nouns is based on their frequency, the first figure in the round brackets indicates the frequency of the noun with the possessive pronoun *jeho* and the second figure the frequency with *její*; the figures in square brackets indicate the number of texts in which the noun occurred with *jeho*).

Upon examination of the above lists, we can again see that there are differences in how male authors represent their female and male characters. There are more male characters and they are described using a greater variety of nouns, including abstract nouns, while for female characters there was only one significant abstract noun: VŮNĚ [smell]. Again, in terms of frequency, the female characters seem to be reduced to a description of their appearance, as in BrMW.

6.3. "TYPICAL" FEMALE AND MALE CHARACTER
IN THE CZWW AND CZMW CORPORA

We attempted to construct a "typical" female and male character for the two Czech fiction corpora (cf. Tables 9 and 10 for English). We have again focused both on the same characteristics and differences. Tables 16 and 17 below summarize the nouns that are significantly associated with male and female characters, respectively, in the CzWW and CzMW corpora. Given that the male possessive pronoun JEHO occurs in both corpora more frequently, we have considered here only nouns that occur with a substantial difference in frequency (as defined in Section 4), while for the female possessive JEJÍ, we have considered all nouns occurring significantly more often.

	'Woman' in CzWW	'Woman' in CzMW
same characteristics in both corpora	family: MANŽEL [husband], MUŽ [man/husband] other: KRÁSA [beauty]	
differences	body parts: NOHY [feet/legs], KLÍN [lap]	body parts: VLASY [hair], ŇADRA [breasts], PRSA [breasts] other: VŮNĚ [smell]

TABLE 16. "Typical" female in the CzWW and CzMW corpora.

	'Man' in CzWW	'Man' in CzMW
same characteristics in both corpora	body parts: MOZEK [brain] family: ŽENA [woman/wife], MANŽELKA [wife] speech: SLOVO [word]	
differences	other: BLÍZKOST [closeness]	place: KANCELÁŘ [office] social circles: PŘÍTEL [friend], PŘÁTELÉ [friends] speech: ŘEČ [speech] other: PRÁCE [work], NÁZOR [opinion], OSUD [fate], PŘEDSTAVY [ideas], DUŠE [soul]

TABLE 17. "Typical" male in the CzWW and CzMW corpora.

Female characters seem to be somehow impoverished in both the CzWW and the CzMW corpus and reduced to stereotypical descriptions of their looks (it is difficult not to notice the contrast between 'beauty' (Table 16) and 'brain' (Table 17)). Male characters in CzWW, when considering only nouns with substantial differences in frequency, are reduced to very few nouns as well. If we consider all the significant nouns (as in Table 12), we additionally find descriptions of physical appearance that involve body parts (OČI [eyes], TVÁŘ [face] [31], RUKA [hand], RUCE [hands] and HLAS [voice]). We also find OTEC [FATHER], DŮM [house] and PRÁCE [work]. There is also an-

other abstract noun: VINA [guilt]. Male characters in CzMW are very similar; in addition they involve 'friends' (both in plural and singular) and more abstract nouns: DUŠE [soul], NÁZOR [opinion], OSUD [fate] and PŘEDSTAVY [ideas]. The whole picture is rather different from that in BrMW, where no abstract nouns occur; instead we have several nouns referring to body parts.

7. CONCLUSIONS

In our comparative (female vs. male writers) cross-linguistic (English vs. Czech) study we have aimed to identify nouns that occur significantly more often with gendered possessive pronouns *her/his* and JEJÍ/JEHO. Our analysis has focused only on the most frequent nouns and it shows that female and male characters in fiction are, to a degree, treated in a stereotypical way. However, the quantitative analysis requires further qualitative study to confirm the results. Purely in terms of the frequencies, there seem to be more male characters in fiction texts included in the BrMW, CzWW and CzMW corpora, and therefore we find more nouns characterising them.

In addition, our analysis has also shown that there are certain differences in the way female/male writers treat their female/male characters. These differences may, however, be partly due to genre conventions and expectations. The analysis also suggests that there are differences of a more general nature in characterisation in British vs. Czech fiction. While in both BrWW and BrMW we find a number of body parts mentioned among the most frequent nouns, this is not the case for the Czech language data. In contrast, in the Czech language data, some of the most frequent nouns are abstract.

British women writers "build" their male characters by mentioning their *arms* or *chest,* thereby confirming these as signs of masculinity, whereas female characters are described using nouns such as *hair* or *cheeks*. In connection with male characters, the noun *office* is mentioned frequently as well, again confirming that men are represented as more often situated in a working environment. The most important male possession seems to be a *car,* while for women it is a *bag.* Perhaps somewhat surprisingly, one of the frequent nouns for women is *stomach* (*her stomach* occurs 4 times more often than *his stomach* in BrWW corpus in more than half of the texts). Another interesting gendered word is *pocket*: in the BrWW corpus *his pocket* occurs 2.5 times more frequently than *her pocket*; in BrMW corpus the difference is even more striking — *his pocket* occurs 17 times more often, so *pocket* is clearly a masculine attribute from the perspective of both female and male writers. Other "signs" of masculinity for British male writers are *watch, glass* and *horse.* It is interesting to note that none of these features among the top nouns for male characters in the Czech fiction. British male writers also to some extent confirm the existing stereotypes in their perception of female characters — the most frequent nouns are *hair, breasts, handbag* and *dress*.

The picture that Czech women writers "build" around their male characters is less clear-cut; they primarily mention the women in their lives (ŽENA [woman/wife], MANŽELKA [wife]), they refer to their 'words' and they also mention BLÍZKOST [closeness], which upon closer examination in a majority of cases refers to a relation be-

tween a man and a woman. Female characters in CzWW and CzMW are similar in that they are both "built" around descriptions of physical appearance. The common "attribute" of masculinity for both female and male Czech writers is 'brain', while the attribute for women is 'beauty'. Attempting a comparison of female vs. male characters across all four corpora would, however, require a more extensive study. The very general conclusions that emerge tend to confirm gender stereotypes; for female characters (across all four corpora) description of appearance is crucial, while this is less so for the male characters. For male characters, the common denominator (across all four corpora) is nouns referring to work, while for female characters it is the 'breasts'. The most apparent differences are the higher frequency of nouns referring to body parts in the British fiction and abstract nouns that we find in the Czech fiction.

Methodologically, the study proved challenging. The task of creating a comparable corpus, i.e. a reliable starting point for the analysis, resulted in a number of compromises (see Section 3). However, these methodological issues concern most contrastive studies and in our view do not constitute an insurmountable problem, if acknowledged in the analysis. Thus the results presented above need to be interpreted in the context of the data we have used. In our case, differences in language typology between English and Czech also played their part. As explained in Section 4, the use of possessive pronouns in English and Czech is not directly comparable, which is clearly visible when the frequencies are compared across the four corpora: English possessive pronouns occur much more frequently than the Czech ones.

REFERENCES

Baker, P. (2014) *Using Corpora to Analyze Gender*. London: Bloomsbury.

Bamman, D., T. Underwood and N. A. Smith (2014) A Bayesian mixed effects model of literary character. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics 1*: Long Papers (ACL: 2014).

Baylog, O., L. Dimmit, T. Heller and G. Kiriloff (2014) "More than custom has pronounced necessary": Exploring the correlation between gendered verbs and character in the 19th century novel. *Presentations, Talks, and Seminar Papers — Department of English*. Paper 7. Available at: http://digitalcommons. unl.edu/englishtalks/7 (accessed 21.3.2017)

Biber, D. and S. Conrad (2009) *Register, Genre, and Style*. Cambridge: Cambridge University Press.

Butler, J. (1990) *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Cameron, D. (2007) *The Myth of Mars and Venus*. Oxford: Oxford University Press.

Cameron, D. (2016) *Do Women and Men Write Differently?* Available at: https://debuk. wordpress.com/2016/03/06/do-women-and-men-write-differently/ (accessed 17.1.2017)

Coates, J. (1998). "Thank God I am a woman": The construction of differing femininities. In: D. Cameron (ed.) *The Feminist Critique of Language: A Reader*, 295–320. London: Routledge.

Criado-Perez, C. (2015) *Do It Like a Woman ... and Change the World*. Portobello Books.

Dušková, L., et al. (2006) *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia. Available in an electronic online version at: http://emsa.ff.cuni.cz/ (accessed 21.3.2017)

Freedman, A. and P. Medway (1994) Locating genre studies: antecedents and prospects. In: A. Freedman, P. Medway (eds) *Genre and the New Rhetoric*. London: Taylor and Francis.

Froehlich, H. (2012) Independent women? Representation of gender-specific possession in two Shakespeare plays. *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching.* Available from: http://www.lancaster.ac.uk/fass/events/laelpgconference/papers/v07/Heather.pdf (accessed 23.3.2017)

Gilbert, S. M. and S. Gubar (1979) *The Madwoman in the Attic: The Woman Writer and the Nineteenth-century Literary Imagination.* New Haven: Yale University Press.

Holmes, J. and M. Meyerhoff (eds) (2003) *The Handbook of Language and Gender.* Blackwell Publishing.

Jespersen, O. (1922/2007) *Language, Its Nature, Development, and Origin.* London: Routledge.

Jockers, M. and G. Kiriloff (2016) Understanding gender and character agency in the 19th century novel. *Journal of Cultural Analytics,* published 12.01.2016, available online http://culturalanalytics.org/2016/12/understanding-gender-and-character-agency-in-the-19th-century-novel/ (accessed 21.3.2017)

Lakoff, R. (1975) *Language and Woman's Place.* New York: Harper and Row.

Leech, G. N. and M. Short (2007) *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. Pearson Education.

Livia, A. (2001) *Pronoun Envy: Literary Uses of Linguistic Gender.* Oxford: Oxford University Press.

Livia, A. (2003) "One man in two is a woman": linguistic approaches to gender in literary texts. In: J. Holmes and M. Meyerhoff (eds) *The Handbook of Language and Gender*, 142–158. Oxford: Blackwell Publishing.

Luke, A. (1994) Series Editor's Preface. In: A. Freedman, P. Medway (eds) *Genre and the New Rhetoric.* London: Taylor and Francis.

Mahlberg, M. (2015) Literary style and literary texts. In: D. Biber and R. Reppen (eds) *The Cambridge Handbook of English Corpus Linguistics.* Cambridge.

Markovits, S. (2006) *The Crisis of Action in Nineteenth-century English Literature.* Columbus: Ohio State University Press.

Showalter, E. (2009) *A Literature of Their Own. British Women Writers, from Charlotte Brontë to Doris Lessing.* London: Virago Press.

Stockwell, P. and M. Mahlberg (2015) Mind-modelling with corpus stylistics in *David Copperfield. Language and Literature* 24 (2), 129–147.

Sunderland, J. (2011) *Language, Gender and Children's Fiction.* London: Continuum.

Tannen, D. (1990) *You Just Don't Understand: Women and Men in Conversation.* London: Virago.

Welter, B. (1966) The cult of true womanhood: 1820–1860. *American Quarterly* 18 (2), 151–74.

Wodak, R. (2015) Gender and language: cultural concerns. In: N. J. Smelser and P. B. Baltes (eds) *International Encyclopedia of the Social and Behavioral Sciences* (2ⁿᵈ ed.), vol. 9, 698–703. Elsevier.

Wolfson, S. (1997) *British Literature: Discipline Analysis.* Baltimore: National Center for Curriculum Transformation Resources on Women.

CORPORA USED

British National Corpus, available at http://bncweb.lancs.ac.uk/

SYN, version 4 (Křen, M. et al.), available at: http://www.korpus.cz

APPENDIX 1.

The authors are listed in alphabetical order; the three letter code in brackets is the BNC identifier.

BRWW CORPUS

Anderson, C. (JYB), Andrews, L. (CK0), Anthony, E. (CDY), Appignanesi, L. (FNT), Ash, R. (JY3), Ashe, J. (JYA), Bailey, E. (HGV), Bainbridge, B. (FNU), Banks, L. R. (FEE), Bauling, J. (H9L), Bawden, N. (CEX, HH9), Beauman, S. (C8S), Bedford, S. (F9R), Beechey, W. (AC7), Bennetts, P. (EVC), Binchy, M. (CCM), Bingham, C. (EEW), Blair, E. (AN7), Bowring, M. (JYE), Brayfield, C. (APU), Browning, A. (JY8), Butler, G. (K8V, H85), Byatt, A. S. (APR, FET), Byrne, J. (HH1), Carter, A. (FRC), Cole, M. (CR6, FAB), Conran, S. (FPB), Cookson, C. (AT7, CFY, CK9), Cooper, F. (A0L, HGF), Cooper, J. (CA0), Cox, J. (FPK), Craven, S. (HH8), Darcy, L. (H9H), Dillon, A. (EVG), Domini, H. (ACW), Drabble, M. (EFP, FB0), Dudley Edwards, R. (HTG), Dunnett, D. (BP0, HRC), Elgin, E. (CEH), Elliot, R. (HA9), Ellis, A. T. (G06, G0X, G1D), Ellis, P. (G0M), Evans, J. (JXW), Faibairns, Zoe et al. (BP8), Fitzgerald, P. (H0R), Forrester, H. (CDN), Forster, M. (ADS), Fox, N. (JY4), Galford, E. (HGN), Gay, A. (FP0), George, C. (JYC), Goodman, L. (H8J), Gower, I. (CKD), Granger, A. (CEB), Grant-Adamson, L. (GV2, GWG), Greenwood, D. M. (HA2), Grey, A. (JY0), Hall, A. L. (HGL), Harvey, C. (H8X), Hill, S. (AD1), Holland, S. (JYD), Howard, S. (H97), Jagger, B. (H7P), James, D. (H7W), James, P. D. (C8T, CJF, G3E), James, S. (GV8), Joseph, J. (H9G), Keane, L. (JXY), Keane, M. (H7H), Kingston, K. (HA7), Lee, T. (GUM), Lessing, D. (EV1), Lively, P. (G0Y), Lorrimer, C. (FS1), Macgregor, M. (HHB), Maitland, S. (A6J), Mansell, J. (H8F), Marshall, P. (HGE), Martin, L. (HGT), Marton, S. (JY7), Mildmay, E. (HH0), Miss Read (ASE), Murdoch, I. (APM), Murphy, E. (G16), Murray, A. (HHA), Myers, A. (H8A), Nabb, M. (CJX), Neel, J. (AB9), Neil, B. (FPF), Neil, J. (HA6), Oldfield, E. (JY9), Pargeter, E. (HGG, K8S), Pearson, D. (J19), Pemberton, L. (FRS), Peters, E. (H8L), Pitt-Kethley, F. (AC3), Pope, P. (BP1), Pym, B. (HA4), Raymond, M. (J54), Rayner, C. (AN8), Rendell, R. (A73, FU2), Rhodes, E. (CB5), Richmond, E. (H9V), Roberts, M. (GUK), Rogers, J. (HJH), Rowlands, B. (GVP, HNJ), Scott, M. (A0R), Shepherd, S. (C8D), Smith, J. (HTR), Staples, M. J. (CKE), Steele, J. (JY1), Stewart, M. (CKF), Storm, C. (AD9, GW2), Tanner, J. (BMW), Taylor, E. R. (FAT, AEA), Taylor, J. (JXV), Trollope, J. (CMJ), Vine, B. (CDB, EDN), Warner, M. (GUX, G0S), Weldon, F. (HGJ), Wells, A. (HA5), Wiat, P. (CCD), Wilder, Q. (JY6), Wilkinson, L. (JY2), Williams, C. (H8H), Wilson, P. (HGD), Wood, B. (G10), Wood, S. (H94), Wright, D. (FSC)

BRMW CORPUS

Aldiss, B. (HA0), Amis, M. (FYV, H0M), Anderson, J. R. L. (H0D), Archer, J. (K8T), Banks, I. (FP6, HWC, G0A), Barber, N. (CHG), Barker, C. (CRE), Barnard, R. (H8Y, H9D), Barnes, J. (G1A), Barnes, T. (G15, G1W), Bartlett, N. (AR2), Bowling, H. (EA5), Bragg, M. (FP1), Brett, S. (GUF, H92), Bulmer, K. (CE5), Burley, W. J. (GW3, GWB, HWP), Cairney, J. (ATE), Cartwright, J. (FAJ), Clarke, L. (H82), Clynes, M. (H90, HU0, HH5), Cornwell, B. (CCW, CMP), Craig, D. (A0N), Critchley, J. (HNK), Curtis, J. (FP7), Darvill-Evans, P. (F9X), Deighton, L. (HR7), Dexter, C. (HWM), Dibdin, M. (BMR, HTT), Doherty, P. C. (BMN, H9C), Falconer, P. (FPX), Farrell, J. G. (EFW), Forbes, C. (ARK, CN3), Forsyth, F. (CAM), Fowles, J. (G07, G13), Fox, W. (ACE), Francis, D. (BP9, ADY), Francome, J. (BP7), Frayn, M. (G12, J17), Gallagher, S. (FYY, G03, GWO), Gandolfi, S. (AMU), Gidley, C. (EWH), Gill, A. (H84), Goddard, R. (G0N, H8T), Greenland, C. (CJA), Grey, A. (FU8), Hamer, M. (CS4), Harding, P. (H98, K95), Harrison, R. (ANL), Hayden, T. (B1X), Herriot, J. (G3S), Higgins, J. (HTW), Hill, D. (G3G), Hill, R. (GUD), Holdstock, R. (HTM), Horwood, W. (FP3), Hutson, S. (G01, G0P), Innes, H. (GV6), Ishiguro, K. (AR3), James, R. (H8M, J13), Josipovici, G. (A08), Kearney, P. (GWF), Kippax, F. (HTS), Kureishi, H. (C8E),

Laws, S. (G0E), Ling P. (FPM), Lodge, D. (ANY, GVT), Lyall, G. (H86, HR4), Macdonald, M. (HP0), Mackenzie, D. S. (ASN), MacLean, A. (CKC), MacNeill, A. (ECK, EF1), Maitland, I. (G02), McDowell, C. (CEY), McGahern, J. (A6N), McGrath, E. (A7J), Mo, T. (H9N), Mortimer, J. (FB9), Newman, K. (ALJ, GVL), Pearce, M. (HTX, J10), Pickernell, R. (B3J), Potter, D. (ADA), Pratchett, T. (HA3), Rankin, R. (HWN, HTU), Ripley, M. (HTL, HW8, HWL), Robinson, D. (HRA), Rutherford, D. (HTJ), Self, W. (FR3), Seymour, G. (CJT, CLD), Shah, E. (CML), Strong, T. (CEC), Taylor, A. (CKB, GUU), Thomas, D. M. (AE0), Thompson, E. V. (HHC), Thomson, R. (C86), Trevor, W. (H7A), Watts, N. (A74), Williams, N. (ASS, HR8, HR9), Wingrove, D. (GUG)

APPENDIX 2.

The authors are arranged in alphabetical order, with the title of the book or the first word of the title followed by three dots added in brackets.

CZWW CORPUS

Babická, H. (Jeden …), Bellová, B. (Celý …), Bellová, B. (Mrtvý …), Bidermanová, S. (Dany …), Boučková. T. (Rok …), Brabcová, Z. (Rok …), Brabcová, Z. (Stropy), Brdečková, T. (Slepé …), Červenková, J. (Pozdní …), Dousková, I. (Goldstein …), Dousková, I. (Hrdý …), Dousková, I. (Medvědí …), Dousková, I. (Oněgin …), Fojtová, V. (Zástava …), Frýbová, Z. (Falešní …), Frýbová, Z. (Hrůzy …), Hofmanová, J. (Láska …), Hůlová, P. (Cirkus …), Kačírková, E. (Dívka …), Kadlečková, V. (Mycelium. Jantarové …), Kocábová, N. (Tohle …), Kriseová, E. (Perchta …), Kriseová, E. (Ryby …), Kubešová, B. (Žabky …), Limrová, K. (Dvě …), Mornštajnová, A. (Slepá …), Naušová, E. (Jizvy), Pekárková, I. (Dej …), Pekárková, I. (Kulatý …), Procházková, I. (Otcové …), Procházková, L. (Oční …), Romportlová, L. (Černobílé …), Rytířová, H. (Jsem …), Rywiková, N. (Dům …), Salivarová, Z. (Honzlová), Soukupová, P. (K moři), Správcová, B. (Spravedlnost), Stiborová, V. (Zapomeň …), Svobodová, J. (Rozlet …), Szalaiová, R. (Tajná …), Švecová, J. (Mé …), Švecová, J. (Zrození …), Tučková, K. (Žítkovské …), Zgustová, M. (Peppermint …), Zgustová, M. (Tichá …), Zgustová, M. (Zimní …), Žídková, A. (Devětkrát …)

CZMW CORPUS

Ajvaz, M. (Druhé…), Bajaja, A. (Zvlčení), Balabán, J. (Černý …), Bartuška, V. (Pochybnost), Bažant, J. (Knihovna), Beran, S. (Hliněné …), Bílý, J. (Zostření …), Boček (Poslední …), Bondy, E. (Šaman), Budinský, V. (Gymnázium), Cimický, J. (Poslední …), Dostál, Z. (Ryby), Dotlačil, J. (Jiné …), Drábek, J. (Náhle …), Durek, Z. (Abtil …), Eidler, P. (Sebranka), Exner, M. (Svatoušek), Fahrner, M. (Steiner …), Filip, O. (Valdštejn …), Fischl, V. (Dvorní …), Frýbort, P. (Vekslák …), Fuks, L. (Myši …), Gruša, J. (Dotazník), Grym, P. (Čínský …), Haidler, J. (Jobova …), Hájek, P. (Kráska …), Hájíček, J. (Rybí …), Hejcman, P. (Let …), Hodek, A. (Gangsteři …), Horák, O. (Dvořiště), Hostovský, E. (Cizinec …), Hovorka, J. (Okarína …), Hrabal, B. (Obsluhoval …), Hromek, J. (Chytněte …), Charoust, M. (Ostrov), Chochola, V. (Jako …), Jandourek, J. (Bomba …), Jedlička, J. (Kde …), Jícha, J. (Prevít …), John, R. (Memento), Kahuda, V. (Proudy), Kalenda, F. (Ordál), Klíma, L. (Velký …), Klvaňa, T. (Marina), Kohout, P. (Ten …), Komárek, S. (Opšlstisova …), Körner, V. (Adelheid), Kovtun, J. (Pražská …), Kratochvil, J. (Truchlivý …), Kraus, I. (Vedlejší …), Krhut, D. (Pašerák …), Kundera, M. (Žert), Kupka, J. S. (Čech …), Landsmann, I. (Fotr), Ludva, R. (Jezdci …), Luňák, P. — Pečenka, M. (Hrdinové), Lustig, A. (Dita …), Matoušek, I. (Ego), Mertl, V. (Hřbitov …), Michal, K. (Čest …), Moc, S. (Upadlí …), Moník, J. (Psi …),

Neff, O. (Reparátor), Němec, J. (Dějiny …), Němec, M. — Herz, J. (T.M.A.), Nuska, B. (Padraikův …), Ouředník, P. (Příhodná …), Páral, V. (Mladý …), Pilous, J. (Se srpem…), Pížl, J. (Adrenalin), Polák, J. (Závody), Poštulka, V. (Hřbitovní …), Procházka, J. (Ucho), Prouza, P. (Ctitelé …), Rašek, A. (Podnikatelka), Rotrekl, Z. (Světlo …), Sojka, J. (Rok …), Stančík, P. (Mlýn …), Svejkovský, J. (Mrtvý …), Svěrák, Z. (Obecná …), Šabach, P. (Občanský …), Šimáček, J. (Cahrakter), Škvorecký, J. (Zbabělci), Šmíd, Z. (Proč …), Švejda, J. (Havárie), Topol, J. (Kloktat …), Trefulka, J. (O bláznech …), Třešňák, V. (Klíč …), Vacek, P. (Hověda …), Vaculík, L. (Český …), Valášek, K. (Veselé …), Vaněk-Úvalský, B. (Zabrisky), Velinský, J. (Zmizení …), Verner, P. (Pražské …), Viewegh, M. (Výchova …), Viktora, Z. (Poslední …), Vokolek, V. (Tak …), Vondruška, V. (Olomoucký …), Weil, J. (Na střeše …), Zábrana, J. — Škvorecký, J. (Vražda …), Zapletal, Z. (Půlnoční …), Žáček, J. (Jsem …), Žváček, J. (Lístek…)

Anna Čermáková
Charles University, Faculty of Arts, Institute of the Czech National Corpus
nám. J. Palacha 2, 116 38 Praha 1, Czech Republic
anna.cermakova@ff.cuni.cz

Lenka Fárová
Charles University, Faculty of Arts, Department of Germanic Studies
nám. J. Palacha 2, 116 38 Praha 1, Czech Republic
lenka.farova@ff.cuni.cz