

## TRANSLATING CHILDREN'S LITERATURE: SOME INSIGHTS FROM CORPUS STYLISTICS

Anna Čermáková\*

University of Birmingham, UK

### Abstract

In this article I explore the potential of a corpus stylistic approach to the study of literary translation. The study focuses on translation of children's literature with its specific constraints, and illustrates with two corpus linguistic techniques: keyword and cluster analysis—specific cases of repetition. So in a broader sense the article discusses the phenomenon of repetition in different literary (stylistic) traditions. These are illustrated by examples from two children's classics aimed at two different age groups: the *Harry Potter* and the *Winnie the Pooh* books—and their translations into Czech. Various shifts in translation, especially in the translation of children's literature, are often explained by the operation of so-called “translation universals”. Though “repetition” as such does not belong to the commonly discussed set of translation universals, the stylistic norms opposing repetition seem to be a strong explanation for the translation shifts identified.

**Keywords:** Children's Literature; Translation Universals; Corpus Stylistics; Repetition; Keywords; Clusters

---

\* Dr. Anna Čermáková is a Marie Skłodowska-Curie Fellow at the Department of English and Applied Linguistics at the University of Birmingham, UK. She holds a PhD in Corpus Linguistics from Charles University, Prague. Her main research interests are in corpus linguistics and particularly in corpus stylistics. She is also interested in literary translation, contrastive corpus based linguistics and lexicology. Her e-mail address is [isa.cermakova@bham.ac.uk](mailto:isa.cermakova@bham.ac.uk)



## 1. Introduction

“Classic” children’s literature is often exceptional in its intertextuality, since it may transcend national literary traditions in ways adult literature more rarely does. In this sense, translation, as a form of adaptation and intertextuality, plays an extremely important role. The importance of translation, however, varies in different national literary traditions. For small languages, such as Czech, translations constitute a substantial part of the canon of children’s literature. Arguably, this canon is to a certain degree international and to some extent common to many literary traditions (e.g. the Swedish *Pippi Longstocking*, the French *The Little Prince*, and the English *Alice in Wonderland* or the *Harry Potter* series have been appropriated by children all over the world). “Smaller” languages in general seem to be more accepting of foreign books, and the proportion of children’s classics that actually are translations, albeit entirely domesticated, may be fairly high if not in the majority<sup>1</sup>. The situation for English is entirely different; as Lathey (2011: 203) says, it is estimated that children’s book production in the UK yearly involves only about 2% of translations.

Translating for children involves a number of specifics; especially developmental stages need to be taken into account — “[t]ranslators should be aware of the stylistic features and modes of address appropriate for different age groups” (Lathey 2011: 199). And as Lathey further says: “Fortunately, the specific demands of translating for children are now the subject of discussion and debate in both professional and academic circles” (ibid). The following work is situated in the field of corpus stylistics and aims to show how the corpus linguistics methods used in the analysis of literary texts can be also beneficial in the analysis of literary translation and even support the translation process itself. A translator performing a stylistic analysis of the text to be translated should not only look at “how the text means what it does” but ideally also proceed to a stage “which seeks to explain, not only *how* the text means what it does, but also *why* a writer may have chosen to shape the text in a particular way” (Malmkjær 2003: 38).

I will illustrate this with two case studies of two frequently used methods in corpus linguistics, keyword and cluster analysis, and these will be examined as specific cases of repetition. In the first study I revisit my earlier work on keywords in *Harry Potter*; the second study of clusters will be illustrated by examples from Milne’s *Pooh* books. The texts were chosen on the basis of their target audience: while with *Harry Potter* older children are the assumed readership, the *Pooh* books are traditionally aimed at smaller children and often read aloud to them by their parents. I will examine and discuss repetition in these texts and their translations in relation to differing stylistic traditions and translation universals.

## 2. Translating Children's Literature

The label “Children’s Literature” may seem self-evident but in fact it is not straightforward to define. The position of children’s literature within the field of

literary criticism is not entirely clear and it is far too often regarded as somewhat inferior to “general” literature. As Nikolajeva (2005: xiii) says (my emphasis):

Some common prejudices about children’s literature, maintained not only by its adversaries but even by its most ardent supporters, include the claims that children’s literature is simply action oriented rather than character oriented, optimistic and with happy endings, didactic, and *repetitive*.

It is especially its didactic nature that is one of its distinguishing features. And as Nikolajeva (1996: 3) further says: “The very emergence of children’s literature on a large scale is due to the fact that sometime in the seventeenth century society became conscious of childhood as a special period of life” and children’s literature has become to be regarded as an important educational resource. However, the notion of childhood has been substantially changing with time and place (Hunt 2005: 3; Grenby & Immel 2009: xiii) and no doubt that the role of literature in children’s lives keeps changing as well.

The “simplicity” of children’s literature is perceived both at the narrative and discursal levels (Nikolajeva 2005). Characters are expected to be uncomplicated, the storyline straightforward, preferably linear, the setting familiar—the same goes for the language: the language of children’s literature is assumed to be “a ‘scaled-down’ version of ‘language in general’, simplified to be made accessible to these young readers” (Thompson & Sealey 2007: 2). In fact, the language of children’s literature has been researched only marginally with a few notable (book length) exceptions (Stephens 1992; Knowles & Malmkjær 1996; Sunderland 2011)<sup>2</sup> and recent corpus based studies by Thompson and Sealey (2007) and Wild, Kilgariff and Tugwell (2013)<sup>3</sup>. This perception of “simplicity” and the didactic character of children’s literature arises from the asymmetrical nature of the communication between the reader and the writer:

The specific feature of communication between a children’s writer and a young reader is that it is asymmetrical. A young reader’s cognitive capacity, life experience, and linguistic skills are normally different from those of an adult writer.... Some scholars claim therefore that children’s literature is always adapted to the needs of its audience... involving subject matter as well as form. (Nikolajeva 2005: xv)

Translation is also a form of an adaptation—through time and space—and it is therefore not surprising that another specific area of research in children’s literature is its translation (e.g. Puurtinen 1995, 1998; Oittinen 2000; Tabbert 2002; Lathey 2006, 2011). In fact, children’s literature has become “a playfield” for some translation scholars aiming at identifying some of the translation universals as it has been assumed that these may manifest themselves in the translation of these texts more clearly than in the translations of adult literature. It is especially the demand for accessibility (i.e. adaptation to the needs of the target reader) that are assumed to cause the emergence of translation universals, namely—

simplification, explicitation, and normalization (for discussion see e.g. Laviosa 1998, Malmkjær 2011, Lapshinova-Koltunski 2015, and Ippolito 2013). Baker (1993) defines “simplification” as “the tendency to simplify the language used in translation” (Baker 1996: 181–182), “explicitation” as the tendency to “spell things out rather than leave them implicit” (Baker 1996: 180) and “normalization” (also called “standardization”, “conservatism”, “conventionalization”), as “the tendency to conform to patterns and practices which are typical of the target language, even to the point of exaggerating them” (Baker 1996: 176-7).

Although “repetition” as such is not mentioned among the “universals”, scholars such as Toury (1977, 1995) and Ben-Ari (1998) see the common translation strategy of avoiding repetition as a “universal” feature of translated language. While the terms simplification, explicitation, and normalization are not generally understood as evaluative, repetition, especially in translation, is judged undesirable. Avoiding repetition is a common translation strategy operating across many languages. Ben-Ari (1998: 1) comments that in the case of repetition “considerations of adequacy come second to considerations of acceptability in the target language”, repetition in the translated text is to be avoided even if the source text uses repetition. He then showcases various translators’ techniques of avoiding repetition. Even though avoiding repetition in translation seems to be a nearly universal translation strategy, there are still considerable differences in tolerating repetition among the various rhetorical and stylistic traditions (Abdulla 2001) and translators’ dealing with repetition is thus a part of the “normalization” process.

A specific case of repetition, which illustrates the difficulty, concerns reporting verbs. The reporting verb *said* is the most frequently used reporting verb in English; however, it is perceived as stereotypical by the translation scholar Levý (2011: 113):

Most professional translators are aware these days that the stereotypical repetition of *said* in English introducing direct speech quite simply belongs to a different literary convention, and as a rule they vary the way they represent this reporting verb in translation.

Corness (2009) investigates translation of *said* into Czech in a parallel corpus of 22 English source novels translated into Czech, i.e. nearly 10,000 occurrences of reporting *said* and finds that it was translated into Czech by no less than 1,323 different translation equivalents (and only 136 occurred 10 times or more). Farová (2016) carries out a similar study in a parallel corpus of English fiction texts and their translations into Czech and Finnish. While in Finnish the translators keep the nearest equivalent (*sanoi*) in the majority of cases, there is a clear tendency in Czech for variation (and avoiding repetition, cf. Bečka 1992: 46). Nádvořníková (2017) examines translations of reporting verbs between Czech, French and English using parallel translation corpora and again confirms on larger data the tendency of Czech translators to avoid repetition and especially the repetition of the reporting verb *said*. However, reporting verbs may, as Mahlberg and Čermáková (forthcoming) show in their analysis of *Alice in Wonderland* and its

Czech translation, substantially contribute to the characterization process and variation of the translation equivalents just for the sake of stylistic variation, i.e. avoiding repetition, may not always be the most suitable translation strategy.

Some researchers argue that readability and naturalness are key expectations in the translations of children's literature, which may provide translators more space for creativity (Oittinen 2000). At the same time, the translations are expected to be ideologically/educationally in line with the target literary setting and other specific features are also relevant; e.g. Alvstad (2010) discusses five issues that need to be considered when translating children's literature. These are:

1. cultural context adaptation, which means modification of the original aimed at adjusting a text for the intended reader's frame of reference (Klingberg 1986), which may include adjusting historical and cultural backgrounds such as measurements or names;
2. ideological manipulation, which Klingberg (1986) calls "purification" and which is basically a form of censorship that may include stylistic changes or even more drastic content adjustments,
3. dual readership (children and adults),
4. features of orality, concerns texts that are intended to be read aloud, which may force the translator to choose between the content and sound, and
5. relationship between text and image as the "coexistence of a verbal and visual code is common in children's literature" (Alvstad 2010:24). Illustrations should support the text's content and in cases of content shifts in the course of the translation process the original illustration may no longer be adequate for the translation.

To this list of issues I would add another one and that is sensitive adaptation of the translation to the stylistic norms of the target language. While it is clear that children's literature, including the translated part, as one of the earliest literacy resources, should more or less conform to the rhetoric tradition with which the children are gradually becoming familiar, this also poses problems if the two rhetoric traditions, i.e. the child's mother tongue and the source text language, are divergent. I will illustrate this (Section 4) with specific cases of repetition and their translation into Czech, as repetition is arguably one of the features of children's literature (see Nikolajeva's quote above).

### **3. Corpus Stylistics in Translation**

The term "corpus stylistics" is most often used as a cover term for the application of corpus linguistics methods to the analysis of literary texts

(Mahlberg 2015)<sup>4</sup>. Corpus stylistic studies may look at a work of a particular author (e.g. Mahlberg 2013) or one particular text (e.g. Mahlberg & McIntyre 2011). Essentially, corpus stylistics, being derived from corpus linguistics, relies on comparisons—the text under study is being compared to another text and often to a large collection of texts; these are called reference corpora, e.g. one text by a particular author may be compared to his/her other texts, or to a corpus of texts of the same period etc. One of the strengths of corpus linguistics is identification of repeated textual patterns and other repetitive textual features.

Repetition—repeated words—are often seen as “the most uncreative of strategies” (Toolan 2012, p. 17). However, as himself Toolan (2012) argues, “repetition is central to literariness and literary creativity” but, clearly “not just any repetition will do: some kinds of repetition work brilliantly, others we generally agree are disastrous”. And how to distinguish them? There are various kinds of repetition and repetitions fulfil various functions in the text; they work differently in poetry (e.g. Toolan 2012) and in fiction texts. A commonly used method both in corpus linguistics and corpus stylistics is a keyword comparison. Keywords are words that occur statistically significantly more often; i.e. they are repeated more often than expected in the text under study in comparison with our reference data. Keywords often show the “aboutness” of the text (Scott 2010) and applied to the study of fiction they may help to identify themes and narrative progression (Toolan 2009), or they may help differentiate among various characters, as in Culpeper's (2009) study of characters' speech in Shakespeare's *Romeo and Juliet*. In the translation, by identifying the keywords, the translator may use them to support the translation process itself (Čermáková 2015). Keywords as a specific form of repetition may create lexical networks that contribute to textual lexical cohesion, and in order to render this kind of cohesion in the translation appropriately, these networks need to be identified (Mastropierro & Mahlberg 2017).

Another specific form of repetition are clusters, i.e. sequences of repeated words.<sup>5</sup> Mahlberg (2013) in her cluster analysis in the work of Dickens identifies several types of clusters and their functions. One type of cluster she calls “labels”: these contain a name or are a part of an expression that is used in a way similar to a name (Mahlberg 2013: 152) and “[i]n addition, Labels relate to the more striking repeated phrases associated with characters that are typically discussed in literary criticism” (Mahlberg 2013: 152). Labels have highlighting and contextualising functions and in Dickens they are often associated with body language, e.g. Monsieur Rigaud in *Little Dorrit* who is brought to readers' memory with the help of a cluster *and his nose came down*. In this sense, the verbatim repetition of five or even more words has an important role and as such we would expect this nature of repetition to be rendered in translation appropriately.

#### 4. Repetition in Translation: *Harry Potter and Pooh*

I will now present for illustration two very brief case studies of repetition and its treatment in translation<sup>6</sup>. In Section 4.1 I revisit an earlier study of keywords

in *Harry Potter and the Philosopher's Stone* (Čermáková & Fárová 2010), and in Section 4.2 I briefly look at clusters in Milne's *Pooh* books.

#### 4.1 Identifying Repetition: Keywords in *Harry Potter*

An earlier study (Čermáková & Fárová 2010) examined the Czech and Finnish translation equivalents of some of the English keywords in *Harry Potter and the Philosopher's stone* (1997) in the translations by Pavel Medek and Jaana Kapari, respectively. The hypothesis of this research was that keywords are sufficiently important words in the original that their translation should show particular consistency. The list of the most frequent keywords itself did not contain any big surprises: it contained mostly proper names and words relating to the novel's theme, e.g. *wand*, *cloak*, *owl*, *wizard*, *broomstick*, *troll*. And as expected, the translator treated most of them with consistent translation equivalents. Some, however, seemed to be problematic, e.g. *troll*. *Troll* is not a part of the “domesticated” (i.e. Czech) fictional and fairy-tale world, and the translator was clearly struggling for an equivalent, which ended up in a somewhat distorted cohesion of the story of *trolls* invading *Hogwarts*. So, in this case, the inconsistency of the translation equivalents was less caused by the repetition in the source text, and more by the translator's strategy for adapting the fictional world at least partly to the target audience. I will now examine another keyword—*wizard*—in more detail.

The word *wizard* appears in the novel 65 times, 14 times as an adjective or in possessive use (e.g. *wizard coins*, *bank*, *gold*, *sport*, *chess*), 51 times as a noun. When we examine the concordance lines and sort them to the left, we can notice a number of positively evaluating adjectives (*best*, *great*, *greatest*, *famous*) in contrast to a single negatively evaluating one, *dark*:

Mad?” said Percy airily. “He’s a genius! **Best wizard** in the world! But he is a bit mad, yes. My dad says it must’ve been a powerful **Dark wizard** to get round Gringotts, but they don’t think Harry looked at the **Famous Wizard** card. “Dumbledore again,” he said. He wasn’t in **Great Wizards** of the Twentieth Century, or Notable Magical widely believed to be the work of **dark wizards** or witches unknown. Gringotts’ goblins today particularly famous for his defeat of the **dark wizard** Grindelwald in 1945, for the discovery of the my master’s instructions - he is a **great wizard** and I am weak - “ “ You mean he was there in “Hermione!” “Harry - you’re a **great wizard**, you know.” “I’m not as good as you, “Professor Dumbledore is a very **great wizard**, Potter, he has many demands on his time -” Considered by many the **greatest wizard** of modern times, Professor Dumbledore is it’s logic - a puzzle. A lot of the **greatest wizards** haven’t got an ounce of logic, they’d be stuck

If we further examine the list of collocates (span -5/5), we get words like *duel*, *witches*, *witch*, *defeat*, *greatest*, *great*, *dark*, *wand*, *bad*. This already gives us some idea about the semantic prosody<sup>7</sup> of the word in this text: wizards may represent the “good” or “evil” side and there is lots of competition going on between these two sides. This is important for the translation as there is no straightforward translation equivalent in Czech.

When we examine the translation solutions in Czech, we find three major equivalents, the most frequent *kouzelník* (29 times - 44.6%), the next *čaroděj* (9 times - 13.8%) and *černokněžník* (5 times - 7.7%). These equivalents have different semantic prosodies in Czech, which I examined in a large corpus of children's literature (12.6 mill. words).<sup>8</sup> The most frequent in the latter was *čaroděj* with a frequency of 54.42 ipm (= instances per million), and occurring in both positive and negative contexts. *Kouzelník* was less frequent (35.46 ipm), often associated with magicians and illusionists as performers. The least frequent, and clearly negative, was *černokněžník* (2.7 ipm).

What is interesting in the Czech translation of *Harry Potter and the philosopher's stone* is the development of translations of the word *wizard* throughout the translation process. In the first five occurrences of *wizard* the translator opts for *čaroděj*. With the sixth occurrence the equivalent changes for *čarodějník* further modified by *zlej* ('bad'), which is not present in the source text, see example 1:

(1)

See, there was this **wizard** who went ... bad. As bad as you could go. Worse. Worse than worse. His name was ...”

Hagrid gulped, but no words came out.

“Could you write it down?” Harry suggested.

“Nah - can't spell it. All right - Voldemort.” Hagrid shuddered. “Don't make me say it again. Anyway, this - this **wizard**, about twenty years ago now, started lookin' fer followers.

To byl totiž **čaroděj**, kterej ... Dal se na špatnou cestu. Byl tak zlej, jak jen to vůbec jde. A horší. Eště horší než horší. Menoval se ...”

Hagrid polkl naprázdno, ale nedostal ze sebe slovo.

“Nechcete to radši napsat?” navrhl Harry.

“To ne ... já nevím, jak se to píše. Ať je teda po tvým - Voldemort,” a Hagrid se zachvěl. “Nechtěj po mně, abych to vopakoval. Ať to bylo jak chtělo, před takovejma dvaceti rokama tendleten - ten **zlej čarodějník** začal na svou stranu přetahovat další.

Here the translator clearly feels that he needs to make a distinction between the good and bad wizards; so he chooses a different translation equivalent (though having the same word stem *čaro-*), and by modifying it with the adjective makes the evaluation explicit (this equivalent is used in the whole book series just once). The

next three occurrences of *wizard* are translated by a newly introduced equivalent *kouzelník*; then we have four more equivalents *čaroděj*. After that, the translator adopts a systematic policy of using *kouzelník* for the good wizards and another newly introduced equivalent *černokněžník* for the bad wizards, see example 2.

(2)

Considered by many the **greatest wizard** of modern times, Professor Dumbledore is particularly famous for his defeat of the **dark wizard** Grindelwald in 1945, for the discovery of the twelve uses of dragon's blood and his work on alchemy with his partner, Nicolas Flamel.

Mnozí ho považují za **největšího kouzelníka** moderní doby. Brumbál obzvlášť proslul svým vítězstvím nad **zlým černokněžníkem** Grindelwaldem v roce 1945, objevem dvanácti způsobů použití dračí krve a svými pracemi o alchymii, jež napsal spolu se svým přítelem Nicolasem Flamelem.

For the adjectival/possessive use of *wizard*, the translator consistently uses the adjective *kouzelnický* (adjective derived from *kouzelník*, which suggests this is the unmarked term, the “norm”) (ex. 3):

(3)

Nah, first stop fer us is Gringotts. **Wizards'** bank.

Nejdřív ze všeho musíme ke Gringottovejm - to je **kouzelnická** banka.

There is one more interesting observation concerning the translation equivalents of *wizard*. The words often occur with their female counterparts *witch(es)*. There are two main translation equivalents in Czech, both derived from the same stem (*čaro-*) as the translation equivalent *čaroděj* (which—as noted above—has been dropped by the translator after the first 13 occurrences of *wizard* in the text): *čarodějka* (usually a “good” witch) and *čarodějnice* (usually a “bad” witch). In these cases the translator opts for the pair *kouzelník* and *čarodějka* for the good ones and the pair *černokněžník* and *čarodějnice* for the bad ones. In both cases he clearly prefers to have a pair of words not based on the same root, though alliteration (still weakly present in the second pair) might have been an interesting stylistic effect (sounding perhaps too repetitive?) and chooses the consistency he adopted earlier, see examples 4 and 5:

(4)

Each house has its own noble history and each has produced **outstanding witches and wizards**.

Každá z nich má vlastní slavnou historii a ze všech vyšli **vynikající kouzelníci a čarodějky**.

(5)

Investigations continue into the break-in at Gringotts on 31 July, widely believed to be the work of **dark wizards or witches** unknown.

Vyšetřovatelé se nadále zabývají vloupáním u Gringottových 31. července, všeobecně připisovaným neznámému **zlému černokněžníkovi nebo čarodějnicí**.

Even though a word like *wizard* in a book like *Harry Potter and the philosopher's stone* is a predictable keyword, the development of the translation equivalents above clearly shows that the translator aims at a consistent approach, which is adopted after the first 13 occurrences of the source word in the text. In this case the consistency amounts to explicitation, as the translator decides to draw a clear line between the good and bad wizards. Keeping the first intuitive equivalent *čaroděj*, would have allowed him to maintain the ambivalent meaning (in terms of the semantic prosodies) of the English *wizard*. However, when we look at the equivalents of *wizard* in the remaining books in the *Harry Potter* series, it seems that the frequent repetition of this word is uncomfortable for the translator, and he drops the translation strategy adopted in the first book: all three equivalents are used — *kouzelník* most frequently (605 times), then *čaroděj* (234 times) and *černokněžník* only rarely (45 times). This suggests the translator aimed to avoid repetition and “normalization” taking over. Considering the intended audience of these books are older children, this is in line with the Czech stylistic tradition, which prefers variation. However, I would argue, if the intended audience were smaller children, consistency (one way or other) would have been a better strategy.

#### *4.2 Identifying Repetition: Clusters in Winnie the Pooh and The House at Pooh Corner*

The following probe into cluster translations loosely follows up on my analysis in Čermáková (2015), in which I examined the translation of John Irving's *Widow for One Year*. Irving is an author who heavily relies on repetition (often of long word sequences—eight or even more words) and, as the analysis has shown, this is stylistically unacceptable (at least from the translators' point of view) both in Finnish and Czech. I aimed to show that performing a basic corpus stylistic analysis may be a useful and powerful procedure for the translators before they embark on the translation to make their translation decisions more objectively informed (by quantitative analysis). The individual translation solutions and decisions are then obviously a complex procedure. Here we will look at two classic stories for younger children, by A. A. Milne: *Winnie the Pooh* and *The House at Pooh Corner*. Both have been translated into Czech by Hana Skoumalová.

I have generated four-word clusters from these two texts and picked two for further brief analysis: *A Very Small Animal* relating to Piglet (occurring 8 times) and the often quoted *Bear of Very Little Brain* (occurring 7 times). They

can be classified in Mahlberg's terminology as "labels". Labels, as clusters closely relating to characters, are important and as Toolan (2001: 113, note 3, quoted in Mahlberg 2013: 152) points out: "[t]here seems almost a compulsion, in both authors and readers, to make character names message-bearing". Names have "trait-evoking effects and trigger recall of the features associated with characters" (ibid.). I believe this is even more important in children's literature discourse, both in the case of the books that children are reading for themselves and books that are being read to them. They help children in sketching the character in their minds and understanding the motivations for a character's actions. For the translation, this means that some sort of consistency in the translation is the preferred translation solution.

In the *Pooh* texts labels are transparent and descriptive ("message-bearing") and are even indicated by the use of capital letters. They play an important role in the narrative, but an exact verbatim repetition is not frequent. Let us now examine the label *Bear of Very Little Brain*, which relates to the main character. The word *brain* itself occurs for the first time already in the introduction (ex. 6, my emphasis):

(6)

Pooh is the favourite, of course, there's no denying it, but Piglet comes in for a good many things which Pooh misses; because you can't take Pooh to school without everybody knowing it, but **Piglet is so small** that he slips into a pocket, where it is very comforting to feel him when you are not quite sure whether twice seven is twelve or twenty-two. Sometimes he slips out and has a good look in the ink-pot, and in this way he has got more education than Pooh, but Pooh doesn't mind. Some have **brains**, and some haven't, he says, and there it is.

We learn that Pooh does not have "any brain" already in the first chapter (ex. 7):

(7)

"Pooh couldn't, because he hasn't any **brain**. Did I catch it?"

The next occurrence appears at the end of Chapter 3 in a conversation between Christopher Robin and Pooh, in which Pooh calls himself *a Bear of No Brain at All* (ex. 8):

(8)

"I have been Foolish and Deluded," said he, "and I am a **Bear of No Brain at All**." "You're the Best Bear in All the World," said Christopher Robin soothingly.

In the next occurrence, we already have the full label that is then repeated several times throughout the text (ex. 9):

(9)

“Terrible and Sad,” said Pooh, “because Eeyore, who is a friend of mine, has lost his tail. And he’s Moping about it. So could you very kindly tell me how to find it for him?”

“Well,” said Owl, “the customary procedure in such cases is as follows.”

“What does Crustimoney Proseedcake mean?” said Pooh. “For I am a **Bear of Very Little Brain**, and long words Bother me.”

In all three instances, the translator chooses various translation solutions—*dost filipa* in example 7, which is closer to English “enough wits”; *a nemám ani špetku rozumu* (“and I don’t have any brain at all”) in example 8; and *přihlouplý Medvěd* (“silly/stupid Bear”) in example 9. Though these convey the meaning, the shift from “very little brain” to “silly/stupid” presents a case of explicitation that I suggest sounds too evaluative and “unkind” to Pooh in the translation. In addition, this variation does not prime the reader for the forthcoming repetition. The repeated cluster *Bear of Very Little Brain* is translated as: *Medvěd s nepatrným rozumem* (“Bear with miniscule brain”), *Medvěd s malým rozumkem* (“Bear with little brain”) on three occasions, *hloupoučký Medvěd* (“silly/stupid Bear”)9.

Due to this variation in translation solutions, the translator is not able to render Milne’s word play around the word *brain*, see ex. 10, in which *a Very Clever Brain* (subject of the sentence) is translated as “someone really clever” and example 11 with *BRAIN OF POOH* (occurring in the text four times) is translated as *Moudrý Pú* (‘wise Pooh’), which seems again a considerable shift and explicitation.

(10)

For he felt sure that a **Very Clever Brain** could catch a Heffalump if only he knew the right way to go about it.

Byl totiž přesvědčen, že **někdo hodně chytrý** by jistě Slona chytil, jenom vědět, kudy na to.

(11)

“I shall call this boat **THE BRAIN OF POOH**,” said Christopher Robin, and **THE BRAIN OF POOH** set sail forthwith in a south-westerly direction, revolving gracefully.

“Nazvu tuto loď **Moudrý Pú**,” řekl Kryštůfek Robin a **Moudrý Pú** vyplul půvabným obloukem na jihozápad.

The second cluster I will look at is the label *A Very Small Animal* used for the Piglet. We find the first mention of Piglet’s size already in the Introduction (see example 6 above). The use of this cluster is different to the *Bear of Very Little Brain*, which in the text of *Winnie the Pooh* gradually builds up (as illustrated above) and occurs fully for the first time in Chapter 4. *A Very Small Animal* occurs three times in Chapter 7 (see example 12 and 13); so it is used fairly late in

the text and then once in Chapter 9 (example 14). We find four more occurrences in *The House at Pooh Corner*.

(12)

“It is hard to be brave,” said Piglet, sniffing slightly, “when you’re only a **Very Small Animal**.”

Rabbit, who had begun to write very busily, looked up and said:

“It is because you are a **very small animal** that you will be Useful in the adventure before us.”

(13)

PLAN TO CAPTURE BABY ROO

1. General Remarks. Kanga runs faster than any of Us, even Me.

...

4. A Thought. If Roo had jumped out of Kanga’s pocket and Piglet had jumped in, Kanga wouldn’t know the difference, because Piglet is a **Very Small Animal**.

...

11. And Kanga wouldn’t discover the difference until Afterwards

(14)

“It’s a little Anxious,” he said to himself, “to be a **Very Small Animal** Entirely Surrounded by Water. Christopher Robin and Pooh could escape by Climbing Trees ....

What is interesting about this phrase is its simplicity and seeming normality. However, if we look into the BNC’s subcomponent of fiction aimed at children, we do not find one single occurrence of *very small animal*. We do not even find the phrase in the whole BNC. We do find *small animal* 29 times but the usage is often quite “technical”, as in the following example.

Since all available records show that barn owls swallow small animal prey whole

The translation equivalents used are based around *malé zvířátko* (‘small animal’, animal in diminutive) variously modified by *tak* (‘so’) (2 times), *docela* (‘quite’), *velmi* (‘very’) and without modification (4 times). The use of the diminutive here makes this perfectly normal sounding for children’s texts (indeed we find the phrase several times in the corpus of the Czech literature for children, see note 8). Also the fact that the core translation equivalent is made of two words only diminishes the repetitive effect in the original text. Thus the translation is smoothed out without this phrase being anyhow extraordinary, making this another normalisation manifestation.

These are indeed small translation shifts; considered jointly they show that one layer of Milne’s narrativity is not being conveyed in the Czech translation. It also suggests that translation ‘universals’ of explicitation and normalization are

operating here, consciously or subconsciously on the part of the translator, with some negative consequences.

## 5. Conclusion

In this article I have aimed to show how corpus stylistics can be used in the analysis of literary translation and further how it can support a translator's stylistic analysis by enlarging the toolkit with corpus linguistics techniques. By identifying keywords and clusters as starting points, we are able to uncover lexical networks that contribute to the construction of literary meaning and cohesion in the text that is of specific importance for the young reader. We have seen that repetition seems to be a source of discomfort for many translators. Lexical repetition is most often compensated for by the use of synonymy (in its broadest sense). The resulting semantic networks convey a similar meaning but the original lexical networks are altered or even lost. Coming back to the quote from Malmkjær (2003) in the introduction, I would argue that in some cases, due to persistent stylistic norms, the translators remain at stage one — “*how* the text means *what* it does” and do not proceed to the next stage “which seeks to explain ... also *why* a writer may have chosen to shape the text in a particular way” (Malmkjær 2003, my emphasis). In the case of children's literature this may be mainly due to constraints imposed by the acceptability of the final text. The various shifts in translations I demonstrated here (and that have been demonstrated elsewhere) may be explained as a natural part of the translation process in which translation universals operate. But, in fact, what many of these shifts seem to show is the translators' strategy of simply avoiding repetition and conforming to the target stylistic norms.

As is the case in all translation, “[t]ranslators are the artisans of compromise” (Lefevere 1992: 6); they are in a position where they have to interpret and mediate between two texts (Lefevere 1992: 10—11), and in translating children's literature they need to take into account their immature reader for whom higher levels of explicitation, normalization and simplification may seem perfectly suitable translation strategy.<sup>10</sup> However, with stylistically carefully crafted source texts, the “*why*” is as important as the “*how*” and “*what*”.

## Notes

1. Based on the statistics from the Municipal Library in Prague for the year 2012, the top twenty list of most borrowed (and presumably read) children's books contains only one book by a Czech author, all the rest being translations.
2. All three books are concerned with the question of the ideological load in children's books and how it is manifested linguistically.
3. Thompson and Sealey (2007) quantitatively explore children's fiction and compare it to fiction texts for adult readers, while the study by Wild et al. (2013) is more lexically oriented and uses keywords as the analytical method (see also Section 3).
4. McIntyre (2015, p. 60) criticises “the tendency to define corpus stylistics rather narrowly as the analysis of literary texts using corpus linguistic techniques”

because “stylistics (of the corpus and non-corpus varieties) is not solely concerned with the analysis of literature” and proposes the following definition of corpus stylistics: “the application of theories, models and frameworks from stylistics in corpus analysis”.

5. The term ‘cluster’ is often interchangeably used with the more general term ‘n-gram’, where the *n* specifies the number of consecutive words that are repeated. Another widely used term in corpus linguistics for repeated sequences of words is ‘lexical bundle’ (Biber et al. 1999).
6. The parallel corpus *InterCorp* available at [www.korpus.cz](http://www.korpus.cz) was used throughout the analysis.
7. For the concept of ‘semantic prosody’ see Sinclair (2004). Here I adapt the concept as text specific (see also Mastropierro & Mahlberg 2017).
8. Texts with the label JUN in SYN\_v4 corpus, available at [www.korpus.cz](http://www.korpus.cz)
9. There is a difference between *přihloupý* and *hloupoučký*. They both mean “a bit silly”, but while the first adjective achieves a toning down with the prefix *při-* but remains still fairly negative, the latter case is a diminutive form, in which the negative evaluation is somewhat weakened.
10. In addition, we should remind ourselves that the final translation product, the printed book we find on the bookshelves, has undergone also an editing process in the publishing house, where editing may take place without consultation of the original text.

## References

- Abdulla, A. K. (2001). Rhetorical repetition in literary translation. *Babel*, 47(4), 289–303. Amsterdam: John Benjamins
- Alvstad, C. (2010). Children’s literature and translation. In Y. Gambier & L. Doorslaer (Eds.), *Handbook of translation studies*, Vol. 1, (pp. 22–27). Amsterdam: John Benjamins.
- Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In H. Somers (Ed.), *Terminology, LSP and Translation*, (pp. 175–186). Amsterdam: John Benjamins.
- Bečka, J. V. (1992). *Česká stylistika* (Czech Stylistics). Praha: Academia.
- Ben-Ari, N. (1998). The ambivalent case of repetitions in literary translation. Avoiding repetitions: a ‘universal’ of translation. *Meta*, 43(1), 68–78. Les Presses de l’Université de Montréal.
- Biber, D., Johansson, S. Leech, G., Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman
- Čermáková, A. (2015). Repetition in John Irving’s novel *A Widow for One Year*: A corpus stylistics approach to literary translation. *International Journal of Corpus Linguistics* 20(3), 355–377. Amsterdam: John Benjamins.
- Čermáková, A. & Fárová, L. (2010). Keywords in Harry Potter and their Czech and Finnish translation equivalents. In F. Čermák, P. Corness & A. Klégr (Eds.), *InterCorp: Exploring a Multilingual Corpus*, (pp. 177–188). Praha: NLN.
- Corness, P. (2009). Shifts in Czech translation of the reporting verb *said* in English fiction. In F. Čermák, P. Corness & A. Klégr (Eds.), *InterCorp: Exploring a Multilingual Corpus*, (pp. 159–176). Praha: NLN.
- Culpeper, J. (2009). Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare’s *Romeo and Juliet*. *International Journal of Corpus Linguistics* 14(1), 29–59.

- Fárová, L. (2016). Uvozovací slovesa v překladech třech různých jazyků. In A. Čermáková, L. Chlumská & M. Malá (Eds.), *Jazykové paralely*, (pp. 145–161). Praha: NLN.
- Grenby, M. O. & Immel, A. (2009). *The Cambridge companion to Children's Literature*. Cambridge: CUP.
- Hunt, P. (2005). Introduction: the expanding world of Children's Literature Studies. In P. Hunt (Ed.), *Understanding Children's Literature*. 2<sup>nd</sup> Edition. London/New York: Routledge, 1-114.
- Ippolito, M. (2013). *Simplification, explicitation and normalization: Corpus-based research into English to Italian translations of children's classics*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Klingberg, G. (1986). *Children's fiction in the hands of the translators*. Lund: CWK.
- Gleerup. Knowles, M. & Malmkjær, K. (1996). *Language and control in Children's Literature*. London: Routledge.
- Lapshinova-Koltunski, E. (2015). Variation in translation: evidence from corpora. In C. Fantionuoli & F. Zanettin (Eds.), *New directions in corpus-based translation studies*, (pp. 93–114). Berlin: Language Science Press.
- Lathey, G. (Ed.). (2006). *The Translation of Children's Literature: A Reader*. Clevedon: Multilingual matters.
- \_\_\_\_\_. (2011). The translation of literature for children. In K. Malmkjær & K. Windle (Eds.), *The Oxford handbook of translation studies*, (pp. 198–213). Oxford: OUP.
- Laviosa, S. (1998). Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Meta: Translator's Journal* 43(4), 557–571. Les Presses de l'Université de Montréal.
- Lefevere, A. (1992). *Translating literature. Practice and theory in a comparative literature context*. New York: The Modern Language Association of America.
- Levý, J. 2011. *The Art of translation*. Trans. by P. Corness. Amsterdam: John Benjamins.
- Mahlberg, M. (2007). Corpora and Translation Studies: Textual Function of Lexis in *Bleak House* and in a Translation of the novel into German. In V. Intouti, G. Todisco & M. Gatto (Eds.), *La Traduzione. Lo Stato dell'Arte. Translation. The state of the Art* (pp. 115–135). Longo: Ravenna.
- \_\_\_\_\_. (2013). *Corpus stylistics and Dickens' fiction*. New York: Routledge.
- \_\_\_\_\_. (2015). Literary Style and literary texts. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 346–361). Cambridge: Cambridge University Press.
- Mahlberg, M. & Čermáková, A. (forthcoming). Patterns in Translation. In M. Mahlberg & A. Čermáková (Eds.), *Corpus As Discourse*. Amsterdam: John Benjamins.
- Mahlberg, M. & McIntyre, D. (2011). A case for corpus stylistics: Ian Fleming's *Casino Royale*. *English Text Construction* 4(2), 204–227.
- Malmkjær, K. (2003). What happened to God and the angels: An exercise in translational stylistics. *Target. International Journal of Translation Studies* 15(1), 37–58. Amsterdam: John Benjamins.
- Malmkjær, K. (2011). Translation Universals. In K. Malmkjær & K. Windle (Eds.), *The Oxford Handbook of Translation Studies*. Oxford: Oxford University Press.
- Mastropierro, L. & Mahlberg, M. (2017). Key words and translated cohesion in Lovecraft's *At the Mountains of Madness* and one of its Italian translations. *English*

- Text Construction* 10(1), 78–105. Amsterdam: John Benjamins.
- McIntyre, D. (2015). Towards an integrated corpus stylistics. *Topics in Linguistics* 16(1), 59–68. Berlin: The Gruyter.
- Nádvorníková, O. (2017). Les proportions des verbes SAY/DIRE/ŘÍCI dans les propositions incisives et leurs équivalents en traduction : étude sur corpus parallèle. *Linguistica Pragensia*, 35-57. Prague: Charles University
- Nikolajeva, M. (1996). *Children's Literature Comes of Age. Toward a New Aesthetic*. Abingdon: Routledge.
- Nikolajeva, M. (2005). *Aesthetic approaches to Children's Literature*. Lanham, Maryland: The Scarecrow Press.
- Oittinen, R. (2000). *Translating for children*. London: Routledge.
- Puurtinen, T. (1995). *Linguistic acceptability in translated Children's Literature*. Joensuu: University of Joensuu.
- \_\_\_\_\_. (1998). Syntax, Readability and Ideology in Children's Literature. *Meta* 43 (4), 524–533. Les Presses de l'Université de Montréal.
- Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In M. Bondi & M. Scott (Eds.), *Keyness in texts* (pp. 43–57). Amsterdam: John Benjamins.
- Sinclair, J. (2004). *Trust the text: language, corpus and discourse*. London: Routledge.
- Stephens, J. (1992). *Language and ideology in children's fiction*. Addison-Wesley Longman Limited.
- Sunderland, J. (2011). *Language, gender and children's fiction*. London: Continuum.
- Tabbert, R. (2002). Approaches to the translation of children's literature: a review of critical studies since 1960. *Target* 14(2), 303–351. Amsterdam: John Benjamins.
- Thompson, P., & Sealey, A. (2007). Through children's eyes?: Corpus evidence of the features of children's literature. *International journal of corpus linguistics*, 12(1), 1–23. Amsterdam: John Benjamins.
- Toolan, M. (2001). *Narrative. A critical linguistic introduction*. London: Routledge.
- \_\_\_\_\_. (2009). *Narrative progression in the short story. A Corpus stylistic approach*. Amsterdam: John Benjamins.
- \_\_\_\_\_. (2012). Poems: wonderfully repetitive. In R. Jones (Ed.), *Discourse and Creativity* (pp. 17–34). London: Pearson.
- Toury, G. (1977). *Translational norms and literary translation into Hebrew, 1930-1945*. Tel Aviv: The Porter Institute for Poetics and Semiotics, Tel Aviv University.
- \_\_\_\_\_. (1995). *Descriptive translation studies and beyond*. Amsterdam: John Benjamins.
- Wild, K., Kilgariff, A., & Tugwell, D. (2013). The Oxford Children's Corpus: Using a Children's Corpus in Lexicography. *International Journal of Lexicography*, 26(2), 190–218. Oxford: Oxford University Press

Recebido em: 29/09/2017

Aceito em: 10/10/2017