

11-1-2021

Contexts and Consequences of Sentence Splitting in Translation (English-French-Czech)

Olga Nádvořníková

Charles University, Prague, olga.nadvornikova@ff.cuni.cz

Follow this and additional works at: <https://digijournals.uni.lodz.pl/rela>



Part of the [Arts and Humanities Commons](#)

Recommended Citation

Nádvořníková, Olga (2021) "Contexts and Consequences of Sentence Splitting in Translation (English-French-Czech)," *Research in Language*: Vol. 19: Iss. 3, Article 1.

Available at: <https://digijournals.uni.lodz.pl/rela/vol19/iss3/1>

This Article is brought to you for free and open access by the Arts & Humanities Journals at University of Lodz Research Online. It has been accepted for inclusion in Research in Language by an authorized editor of University of Lodz Research Online. For more information, please contact journals@uni.lodz.pl.

CONTEXTS AND CONSEQUENCES OF SENTENCE SPLITTING IN TRANSLATION* (ENGLISH-FRENCH-CZECH)

OLGA NÁDVORNÍKOVÁ

Charles University, Prague

olga.nadvornikova@ff.cuni.cz

Abstract

The present paper examines the contexts and consequences of sentence splitting in English, Czech and French translated fiction. In the data extracted from a parallel (multilingual) corpus, we analyze first a language-specific context of sentence splitting (sententialization of non-finite verb forms in translations from English and French into Czech), and second, contexts of splitting occurring in all directions of translation. We conclude that sentence boundaries are usually introduced at the point of a sentence entailing the fewest modifications in the target sentence, especially between two coordinate clauses; and that a systematic sentence splitting, deeply modifying the style of the source text, involves the effect of simplification and normalization.

Keywords: sentence splitting, translation, non-finite verb forms, coordination, simplification, normalization

1. Introduction

Sentence splitting in translation, i.e. replacing one sentence structure in the source text by two or more sentence structures in the target text, may have various causes: on the one hand, differences in grammatical structure or in stylistic norms between the source and the target languages, and, on the other hand, specific features of the translation process itself, e.g. the tendency to simplification. The aim of this paper is to analyze occurrences of splitting of sentences in translation of fiction in English, Czech and French, and to observe contexts in which the splitting occurs, and the consequences it entails at the sentence level as well as the text level.

The first section of this paper summarizes the previous research of changes of segmentation in sentences and in punctuation in the domain of contrastive linguistics as well as translation studies. The next section introduces the corpus and the data used in the present study: the InterCorp parallel corpus and non1:1 segments extracted from it. (Non1:1 segments are aligned pairs of parallel segments consisting of more than one segment on either side.) The aim of the

* This work was supported by the European Regional Development Fund-Project „Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World“ (No. CZ.02.1.01/0.0/0.0/16_019/0000734).

quantitative analysis carried out on these data is to identify the frequency of splitting of sentences in translation in the three languages under analysis and specific texts where splitting of sentences may be a systematic translation strategy. Section 4 is the main part of the paper, analyzing contexts of splitting of sentences and its consequences in translation. The final part of the paper contains a brief summary and some suggestions for further research.

2. Previous research

Splitting of sentences has drawn the attention of researchers in contrastive linguistics as well as in translation studies.

From the contrastive point of view, important research into splitting of sentences was carried out by Fabricius-Hansen (1996, 1998 and 1999) and Solfjeld (1996), comparing German, Norwegian and partly also English with regard to the information density of non-fictional texts (within the framework of the SDRT theory). Fabricius-Hansen (1998: 198) defines (relative) informational density “as a function of the amount of information conveyed in a text or discourse on the one hand and the number of sentences, clauses, and words contained in the discourse on the other hand”. She shows that German non-fictional texts encode the discourse information in long, hierarchical sentences, whereas Norwegian prefers a more paratactic style. Therefore, in translations from German into Norwegian, both authors observe information (sentence) splitting.

For our research, the most important part of these papers are the analyses of potential causes and consequences of the splitting of sentences. As for the consequences of sentence splitting, both Solfjeld and Fabricius-Hansen analyze in detail so-called information extraction in the information structure of the target sentence(s). As for the causes, Fabricius-Hansen mentions German linguistic features allowing for a condensed information packaging, such as heavy prenuclear and postnuclear noun phrase modification or accumulation of adverbial adjuncts (Fabricius-Hansen 1998). Solfjeld (1996) points out that the preference of Norwegian for finite clauses, in contrast with non-finite ones in German, results in their *sententialization*.

Fabricius-Hansen (1999: 204) also points out that, in comparison with Norwegian, English is able to pack discourse information in larger units thanks to a wider use of participial adjuncts and to a more refined use of punctuation, including the colon and semicolon. A recent research has confirmed this difference in the use of punctuation for Czech on the one hand, and French and English on the other hand: the frequency of the colon, and particularly semi-colon is much lower in Czech than in the other two languages (Nádvořníková, forthcoming1).

Nevertheless, research in translation studies has shown that splitting of sentences in translation may be caused not only by structural and stylistic

differences between languages (such as the use of non-finite clauses or punctuation marks), but also by specific features inherent to translation, so-called translation universals (see e.g. Baker 1993; Mauranen and Kujamäki 2004; Malmkjær and Windle 2012; Nádvorníková 2017a or Robin 2017).

In her seminal study, Baker (1996: 176–177) defines four basic translation universals (simplification, normalization, explicitation and levelling-out). Splitting of sentences is most frequently associated with *simplification* (see e.g. Laviosa 2006 or Chlumská 2017), i.e. “the idea that translators subconsciously simplify the language or message or both” (Baker 1996: 176). Vanderauwera (1985) observed systematic sentence splitting in her corpus of about fifty novels translated from Dutch into English. These changes were accompanied by other shifts making the texts simpler and more “readable”, e.g. the above-mentioned raising of the non-finite structures to finite ones.¹ Similarly, Pagnoulle (2004) shows how the French translator of Joyce’s novel *Finnegans Wake* shortens the markedly long sentences of the source text into *tranches digestibles*, making the target text more clear and elegant, but at the same time erasing the specific “affective” style of the source text.²

Recent quantitative research carried out on more than 130,000 non1:1 segments of the English-French-Czech part of the version 8 of the InterCorp parallel corpus (Nádvorníková 2017a) has confirmed that splitting of sentences occurs more frequently in translations into Czech (from English and French), whereas the joining of sentences in the opposite direction of translation, from Czech into French and English (see Table 2 in Nádvorníková 2017: 452). The present paper is the logical continuation of that mostly quantitative study: focused only on one type of shift (splitting of sentences), we move from the overall description of tendencies to the analysis of its specific consequences in translations. Moreover, we benefit from the new version of the InterCorp parallel corpus (version 10), considerably larger than the one used in the preceding research.

3. Corpus & Data

The InterCorp parallel corpus in its most recent version (version 10) contains more than 2.1 billion tokens. At present, it involves 40 languages, with Czech as the pivot language (see www.korpus.cz/intercorp and Čermák and Rosen 2012 or

¹ Vanderauwera (1985) attributes these modifications to the minority status of the Dutch literature, in comparison with the target (English) literature.

² Laviosa-Braithwaite (1996) considered the average sentence length as a potential indicator of simplification, assuming that it would be higher in translations than in non-translated texts. However, this assumption was confirmed only in journalistic texts; in fiction, the opposite was the case.

Nádvořníková 2016 in French).³ The corpus is divided into a “core” part and “collections.” The “core” consists mostly of fiction and partly of non-fiction. The “collections” are made up of various text types: movie Subtitles, Acquis communautaire, transcripts of debates in the European Parliament and journalistic texts (collections SYNDICATE and Presseurop). In 2017, 18 translations of the Bible were also added to the corpus.

The distinction between the core of the corpus and the collections is important for our research: all core texts are proof-read and alignment in this part of the corpus is checked semi-manually. Another advantage of the core of the corpus for our research is the high quality of translation and the possibility of identifying the direction of translation with certainty.

As it is not possible to extract non1:1 segments from the common corpus interface of InterCorp, the data were provided directly by the Institute of the Czech national corpus.⁴ Nevertheless, as explained previously in Nádvořníková (2017: 450–451), the alignments are limited to segments including Czech, the pivot language, because all the texts are aligned first to the Czech version, and only through this alignment to the other language versions. For this reason, the direct shifts between English and French are not accessible to this study; however, the comparison of the two languages is included in the comments whenever the second translation was available in the corpus. Table 1 shows the composition of the corpus used in this research:

Table 1. Size of the InterCorp parallel corpus (fiction, augmented v10).
(EN, CS, FR – source languages; en, cs, fr = target languages;
in brackets = intersection of three languages)

Sub-corpus (fiction)	N° of tokens	N° of texts
EN-cs(-fr)	18,953,496 (1,134,556)	165 binary (11 tertiary)
CS-en(-fr)	3,420,369 (1,483,802)	38 binary (19 tertiary)
FR-cs(-en)	7,309,594 (573,088)	86 binary (6 tertiary)
CS-fr(-en)	3,706,609 (1,483,802)	48 binary (19 tertiary)

We can see that the intersections between the three languages are more limited than the binary combinations; the difference is the most important in translations from French into Czech and English. This problem is even more visible in the non-fiction sub-corpus; therefore, we decided to exclude it from our research. Nevertheless, we hope that in the future it will be possible to explore the shifts in segmentation in sentences in this text type, too. And because of specific use of

³ The corpus has been used extensively in research both in contrastive linguistics and in translation studies (see www.korpus.cz/biblio).

⁴ I owe thanks to Martin Vavřín and Pavel Procházka from the Institute of Czech national corpus for their help and cooperation.

punctuation, especially in stage directions, we excluded drama from our research, too (e.g. *Ubu roi* by Alfred Jarry or several plays by Václav Havel).

The resulting corpus used for this research therefore includes only fiction in English, French and Czech. On the non1:1 segments extracted from these texts, we carried out a manual analysis of occurrences in which the number of segments in the target language was higher than in the source language. For each direction of translation (FR-cs, CS-FR, EN-cs, CS-en), we analyzed 500 segments, i.e. 2,000 segments in total. We included the second translation whenever it was available (e.g. EN-cs-fr).

Although limited only to fiction (novels) in Czech, English and French, our corpus remains very heterogeneous in style, which inevitably influences the translations, too (cf. the heterogeneity of texts written by Virginia Woolf, James Joyce, J.K.K. Rowling, Marcel Proust, Ferdinand Céline, Michel Houellebecq, Milan Kundera, Bohumil Hrabal or Jaroslav Hašek, all included in the corpus). In non-fiction, the principal aim of the translator is to convey the information introduced in the source text; in fiction, however, the translator is often confronted with creative use of syntax and punctuation, and he/she is supposed to retain these stylistic features in the target text as much as possible. Moreover, the target text may be influenced by specific translators' strategies, their proficiency, linguistic/translation tradition, mother tongue,⁵ editorial guidelines (see e.g. Bisiada 2018), expectations of the target audience, etc.

Table 2 shows the proportion of split segments in the four directions of translation available in our corpus (in relation to the total number of alignments).⁶

Table 2. Proportions of split segments in four directions of translation in the InterCorp parallel corpus (EN, CS, FR – source languages; en, cs, fr = target languages)

Direction of translation	Total N° of alignments	N° of split segments	%
EN>cs	1,168,881	59,639	5,10%
CS>en	212,373	16,824	7,92%
FR>cs	378,750	14,956	3,95%
CS>fr	242,579	15,116	6,23%
TOTAL	2,002,583	106,535	5,32%

The data in Table 2 reveal that on average, sentence splitting in either direction of translation is not a dominating strategy: we find the highest proportions in

⁵ For example, two texts from the corpus of translations from Czech into French are themes, not versions, which means that their translators were Czech, translating into French (*Osudy dobrého vojáka Švejka* by Jaroslav Hašek and *Český snář* by Ludvík Vaculík).

⁶ We consider as occurrences of sentence splitting the non1:1 segments where $x > 0 \wedge x < y$ (x = source text, y = target text). However, these numbers are only approximative, as they include also e.g. misaligned segments or segments split by semi-colon, taken by certain tokenizers as sentence boundary.

translations from Czech (into English 8% of all the alignments, into French 6%), and the lowest proportions in the opposite direction of translation (5% in translations from English into Czech and only 4% in translations from French). In fact, in most of the texts, the proportion of split segments varies between only 1% – 10% of the total number of segments in the source text: 146 texts (90% of the subcorpus) in translations from English into Czech, 67 texts (81% of the subcorpus) in translations from French into Czech, 37 texts (77% of the subcorpus) in translations from Czech into French and 19 texts (50% of the subcorpus) in translations from English into Czech.

The results suggest that Czech translators are more respectful to the sentence boundaries in the source text, despite the structural and stylistic differences that may indicate the contrary (see the use of non-finite clauses and colons and semi-colons in section 2; and the sententialization of non-finite verb forms in section 4.1). Nevertheless, due to the above mentioned heterogeneity of our fiction corpus, the proportion of split segments in translation varies from almost zero⁷ to more than 25% of the segments in the source text (29% in a novel by Nicholas Evans translated into Czech, see (11); more than 53% in a novel by Bohumil Hrabal translated into English, cf. in Appendix an example extracted from another novel by the same author; more than 50% in *Nicholas on Holiday* by René Goscinny (see note 20); and 28% in a novel by Petr Šabach, a contemporary Czech author, translated into French, see section 4.2.4.1).

We can compare the results based on the analysis of non1:1 segments involving splitting of sentences (Table 2) with changes in the frequency of the most frequent final punctuation mark, the full stop. Again, an increased proportion of full stops indicates more splitting of sentences. Table 3 shows the proportions of the increase in the number of full stops in translations in comparison with the source text:

Table 3. Number of translations involving an increased number of full stops (in comparison with the source text). EN, CS, FR = source text; en, cs, fr = translations

Modification of the N° of full stops	EN>cs	CS>en	FR>cs	CS>fr
increase by 0%-10%	87 (53%)	14 texts (37%)	63 texts (73%)	25 texts (52%)
increase by 11%-20%	11 texts	5 texts	4 texts	2 texts
increase by 21%-30%	2 texts	4 texts	1 text	1 texts
increase by >30%	0 texts	6 texts	2 texts	2 texts
Decrease	65 (40%)	9 texts (24%)	16 texts (19%)	18 texts (38%)
TOTAL	165 texts	38 texts	86 texts	48 texts

⁷ E.g. Raymond Chandler *The Curtain* 0,90% in EN>cs, 0,05% and 0,13% respectively in two texts by Milan Kundera in CS>en, 0,40% in a novel by Camille Laurens in FR>cs, and 0,02% in another text by Milan Kundera, translated into French (see (4)).

According to Table 3, translations involving a high proportion of increase in the number of full stops (more than 21% of segments) represent only a minority in the corpus (the most in translations from Czech into English), which corroborates the results obtained from non1:1 segments.⁸ The rare cases showing an extreme increase in the number of full stops indicate important sentence splitting, used as a global translation strategy. These cases will be analyzed in section 4.2.

Nevertheless, we have to point out that the increased number of full stops is not an entirely reliable indicator of sentence splitting: first, it does not take into account the other final punctuation marks (ellipsis dots, question mark and exclamation mark), and second, it does not cover cases where the translator compensates for the splitting of sentences by joining them elsewhere. More importantly, none of these data (either based on non1:1 segments or on the frequency of final punctuation marks) cover differences in contexts of splitting of sentences and their consequences, which is the main topic of this paper. Therefore, the next section will be devoted to this analysis.

4. Contexts and consequences of splitting of sentences in translation

As we have mentioned in section 2, we assume that splitting of sentences may be caused on the one hand by structural and stylistic differences between source and target languages, and on the other hand by general features of translation, especially simplification. We expect that the first type of splitting will be language-specific, which means occurring only in one direction of translation (see section 4.1), and the second type will be universal, i.e. occurring in all directions of translation (see section 4.2). The data show that it is more appropriate to speak about the “contexts”, “types” or “opportunities” of splitting of sentences than “causes”. In fact, it is often difficult to identify the one and only “cause” of the splitting of a sentence. For example, the translation of non-finite verb forms by finite ones (see section 4.1) is often accompanied by an overall complexity of the source sentence, which would be another “cause” of the splitting. For this reason, we do not introduce any exact quantifications of the “causes” of splitting of sentences, but only general tendencies and approximate percentages.

On the basis of the preceding findings, we can formulate the following hypotheses concerning the causes (contexts) and consequences of splitting of sentences in translation.

Hypothesis 1 (H1, see section 4.1)

As was shown by Fabricius-Hansen (1996 and 1999) and Solfjeld (1996) (see section 2), the translation of non-finite verb forms by finite ones may cause their

⁸ Surprisingly, the proportion of texts showing not an increase, but a decrease of the number of full stops (albeit mostly very low) is the highest in translations into Czech. However, this result is out of the scope of this paper, because it concerns sentence joining.

sententialization. In accordance with the contrastive research carried out by Čermák et al. (2015) and Malá and Šaldová (2015), we expect more splitting of sentences caused by the sententialization of non-finite verb forms in translations from Czech into French and English than in the opposite direction of translation, because the use of non-finite verb forms is more limited in Czech than in the other two languages.

Hypothesis 2

We assume that independently of the direction of translation, translators will split long sentences, in accordance with the tendency to *simplification* (see translation universals in section 2).

Hypothesis 3

We assume that splitting of sentences entails an increased amount of explicitness and redundancy in the target segment (cf. this observation in Blum-Kulka 1986), because the cohesive ties between the split segments have to be conveyed by means of anaphoric expressions, lexical repetition or the explicitation of semantic relations.

4.1. Language-specific contexts of sentence splitting

In our data, the most frequent language-specific (or translation direction specific) context of splitting of sentences involved the translation of non-finite clauses (participial adjuncts, gerunds, infinitive and absolute constructions, etc.) by finite ones (see *sententialization* of non-finite verb forms in section 2). The analysis of all types of non-finite clauses would not be possible in this study; therefore, we will focus only on one type of them: gerunds and participial adjuncts in English, and *gérondif* and *participe présent* in French.

Our data show that splitting of sentences caused by the sententialization of non-finite verb forms is specific to translations into Czech, which confirms our first hypothesis (H1). In fact, the use of non-finite clauses is very limited in Czech, and this grammatical (structural) difference makes the translators render the French and English non-finite verb forms by finite ones, which may cause the splitting.⁹ As we will see below, this type of change does not occur in the translations from Czech into French and English. One of the main reasons is the fact that the Czech adverbial non-finite verb form, the transgressive, a potential equivalent of English gerunds and participial adjuncts and French *gérondif* (and partially also *participe présent*), is archaic and almost extinct in contemporary Czech (see Nádvorníková 2010 and 2013). In our corpus, the transgressive occurs

⁹ Verbalization of non-finite verb forms is rare in translations between English and French; the most frequent shift is the translation of the English *-ing* form by a relative clause in French (see the same observation in Ballard 2007).

rarely, especially in old texts (see example in note 10) or in texts in which it contributes to a solemn, archaising stylisation (see (13)).¹⁰

Recent research carried out on the InterCorp parallel corpus has shown that the French *gérondif* has for equivalent in Czech a (coordinate or subordinate) finite verb in 67% of the 2,362 analyzed occurrences (see Čermák et al. 2015). A similar result has been observed in translations from English into Czech: 73% of adverbial participial constructions have a finite clause as their equivalent in Czech (see Malá and Šaldová 2015). Other research has shown that only a small part of verbalized French and English participial adjuncts are involved in splitting of sentences (Nádvorníková, forthcoming²). Nevertheless, in the amount of splitting of sentences in translations into Czech, sententialization of non-finite verb forms is involved in about 10% of splitting of sentences in translations into Czech, both from English and French.¹¹

In the following example, we can see that the French translator rendered the English *-ing* forms in the introductory clause by non-finite structures, too (a *gérondif* and an absolute construction), whereas the Czech translator verbalized both structures. This change may have overweighed the introductory clause, and this is probably the reason why the translator split it in two. Because of the verbalization and sententialization, the hierarchy of information provided by the non-finite verb forms is lost, because all the events are placed on the same level of importance. Another consequence of the verbalization of non-finite forms is the necessary explicitation of the grammatical categories of mood, tense and person. Moreover, so as to retain the meaning of accompanying circumstance, conveyed by the *-ing* forms, the translator had to add the explicitating temporal adverbial *Naráz* (meaning “immediately”):

(1) (EN) ‘You don’t mean – you can’t mean the people who live here?’ cried Professor McGonagall, **jumping** to her feet and **pointing at** number four. (J.K. Rowling, *Harry Potter and the Philosopher’s Stone*, 1997/2004)

(fr) – Vous voulez dire... non, ce n'est pas possible ! Pas les gens qui habitent dans cette maison ! s'écria le professeur McGonagall **en se levant** d'un bond, **le doigt pointé sur** le numéro 4 de la rue. (transl. J.-F. Ménard, 2005)

(cs) "Snad nemyslíte – přece nemůžete myslet ty lidi, co bydlí tady?" vykřikla profesorka McGonagallová. **Naráz byla na nohou** a **ukazovala [she jumped to her feet and pointed at]** na dům číslo čtyři. (transl. P. Medek, 2000)

The adding of explicitating adverbial expressions (usually the adverb *přítom* or the PP *při tom*, both meaning *at the same time*, see (2)) is the most frequent

¹⁰ (FR) "Eh bien, mon ami", lui dit Fix **en l'abordant**, "votre passeport est-il visé ?" (J. Verne, *Le tour du monde en quatre-vingt jours*, 1873); (en): "Well, my friend," said the detective, **coming up** with him, "is your passport visaed?"; (cs): „Nu tak, příteli," řekl Fix, **přistupuje** k němu, „už máte visum v pase?“.

¹¹ Absolute constructions as the source of sentence splitting are more frequent in translations from French; in English, participial adjuncts are more frequent in this context.

modification accompanying the splitting of sentences involving the sententialization of non-finite verb forms (it occurs in this context in about 13% of cases).

- (2) (EN) I checked them immediately and urged them forward again, // calming my foolish heart as best I could and **reproving** myself for such nervousness. (I. Banks, *The Bridge*, 1986/2013)
 (cs) Okamžitě jsem je uklidnil a znovu je pobídl **kupředu**. // **Sám** [myself] jsem **při tom** [at the same time] ze všech sil **tišil** své bláhové srdce a **káral se** za nesmyslnou nervozitu. (transl. J. Kantůrek, 2003)

Other changes, such as adding anaphoric expressions (see the pronoun *sám/myself* in (2)), are exceptional in this context.¹² Thus, our third hypothesis (H3), was confirmed only partially. Our data suggest that the loose adverbial relation between the subordinate (finite or non-finite) clause and the main clause (the meaning of accompanying circumstance) is one of the opportunities for the translator to split the sentence in two. In fact, if the non-finite clause conveys a specific adverbial meaning (temporal, instrumental, conditional, etc.), it is usually rendered in Czech by a subordinate structure, most often by a subordinate finite clause.¹³ In this case, sentence boundaries are maintained. However, if the semantic relation of the non-finite clause to the main clause is more loose, i.e. only a general accompanying circumstance, the most frequent equivalent in Czech is a coordinate clause,¹⁴ or, less frequently, two independent sentences (see (1) and (2)).

However, it is not possible to say that the sententialization/verbalization of the non-finite forms and their loose relation to the main clause are the direct and the only cause of sentence splitting, because other factors play a role, especially the already mentioned overall complexity of the source sentence (e.g. the accumulation of several non-finite verb forms – or relative clauses – in the same sentence). In fact, English or French sentences containing non-finite verb forms and split in Czech translation are usually composed of at least four (finite or non-finite) clauses. The splitting of these sentences is frequent especially when the non-finite clause is a (long) premodifying structure, which is typical in French:

- (3) (FR) **Ayant entendu** dire à la cuisine qu'il fallait se méfier du coq **et craignant** d'avoir eu déjà la langue trop longue, // Marinette en resta là et quitta la remise avec la bûche qu'elle venait de choisir. (M. Aymé, *Les contes du chat perché*, 1939)

¹² Our data indicate that important adding of anaphoric expressions is triggered by sententialization of relative clauses, especially in translations of French non-fiction, where cumulation of several relative clauses is frequent.

¹³ E.g. Et elle nous défiait moi et Sophie en même temps, tout **en disant** ça. (L.F. Céline, *Le Voyage au bout de la nuit*, 1932/2001) (cs) A **když to říkala**, mířila tím taky na nás, na mě a na Žofii. (en) **In saying** that she was defying me and Sophie too.

¹⁴ (FR) Et le petit prince s'en fut, **songeant** à sa fleur. (A. de Saint Exupéry, *Le Petit prince*, 1946/1999); (en) And the little prince went away, **thinking** of his flower.; (cs) Malý princ odešel a **myslel [left and was thinking of]** na svou květinu. (transl. Z. Stavinohová, 1972/1989).

(cs) Marienka v kuchyni slyšela, že si musí dát na kohouta pozor, a tak se polekala, že už toho řekla až **moc**. // V **půli** slova se zarazila, chytila poleno, které si vybrala, a upalovala z kůlny pryč. (transl. T. Sýkorová, 1979)

The French source sentence in (3) is quite complex, as it contains seven clauses (four finite ones and three non-finite ones, if we count also the infinitival complement of the participle *craignant*). The Czech translator turned all the three non-finite forms into finite ones and split the sentence on the border of the premodifying structure and the main clause.¹⁵

The language-specific character of sentence splitting involving the sententialization of non-finite clauses is confirmed by two facts in the opposite direction of translation (from Czech into French or English):

First, the splitting of sentences involving the sententialization of non-finite verb forms is quasi non-existent in translations from Czech into French/English, due to a very low frequency of these verb forms in Czech.

Second, French and English translators resort to the de-sententialization (implication) of some Czech finite verb forms, in accordance with the norms of the target language:

(4) (CS) Pokročil jsem k němu a **chtěl jsem** ho odstrčit. // V té chvíli **se rozpřáhl** a udeřil mne pěští do prsou. (M. Kundera, *Žert*, 1968/1991)

(en) When I went up to him, **intending** to push him aside, // he swung and punched me in the chest. (transl. D. Hamblyn; O. Stallybrass, 1992)

(fr) J'avançai jusqu'à lui et fis mine de l'écarter. // Alors, **balançant** le bras, il me lança son poing dans la poitrine. (transl. M. Aymonin, 1975)

In (4), the source segment, written in paratactic style, is composed of independent or coordinate clauses, placed on the same level of importance. In both translations, however, the segment is more hierarchically structured by the means of non-finite verb forms (participial adjuncts *intending* and *balançant le bras*). Moreover, the English translators join the two source sentences in one, using a subordinate finite clause. Both translations correspond to the syntactic and stylistic norms of the target languages. However, as we will see in 4.2.2, exaggerated and systematic modifications according to the target language norms may erase the idiosyncratic source text style.

4.2. Types of splitting of sentences occurring in all directions of translation

In our manually analyzed samples of non1:1 segments (see section 3), we identified three types of contexts triggering sentence splitting independently of the direction of translation. The first two types are mostly typographical, bringing with them only a slight effect of simplification (structuration); the second two

¹⁵ Fabricius-Hansen (1999: 195) calls this type of sententialization “backward information extraction,” as information is extracted to the left of the Principal Counterpart (see section 2).

types, however, may have an impact on the information structure of the target text. The last part of this section (4.2.2) is devoted to the analysis of the contexts triggering the systematic splitting of sentences in translation.

4.2.1. Change of lowercase to upper case after a final punctuation mark

This first type of splitting occurs especially after ellipsis dots, exclamation marks or question marks: in the source text, this punctuation mark is followed by a word starting with a lowercase letter, in translation(s), the equivalent starts with an uppercase letter. Therefore, the alignment tools identify the source segment as one sentence, whereas the target segment as two sentences.

- (5) (FR) - Qui êtes-vous... qui êtes-vous... qui êtes-vous..., répondit l'écho.
(A. de Saint Exupéry, *Le Petit prince*, 1946/1999)
(en) "Who are you – Who are you – Who are you?" answered the echo. (transl. K. Woods, s.a.)
(cs) „Kdo jste?... Kdo jste?... Kdo jste?“ odpovídala ozvěna. (transl. Z. Stavínohová, 1972/1989)¹⁶

This type of change occurs in all directions of translation, and in either language, it is not motivated by a strict stylistic (typographic) convention. The simplification it brings about is very slight.

4.2.2. Replacing comma by a full stop at the end of the interposed introductory clause

This change resembles the previous one: it occurs in all directions of translation in nearly all texts, it is not motivated by a strict stylistic (typographic) convention, and the simplifying effect is minor:

- (6) "Jó, to mají pravdu," odpověděl pan Palivec, // visel tam a sraly na něj mouchy, tak jsem ho dal na půdu. (J. Hašek, *Osudy dobrého vojáka Švejka za světové války*, 1921–1923, 1996)
– Ça, vous avez raison, riposta le patron. // Mais, comme les mouches chiaient dessus, je l'ai fait enlever et mettre au grenier. (transl. H. Horejsi; C. Ancelot, 1989)
"Yeah, you're right," said Palivec, // "it used to hang there. // And the flies kept shitting on it, so, I put it in the attic. (Z.K. Sadloň, 1930)¹⁷

In (6), only the French translator replaced the comma at the end of the introductory clause by a full stop, indicating clearly the boundary of narrative levels; the English translator, on the contrary, retained the original punctuation. We could hypothesize that this type of splitting is motivated in French by the

¹⁶ Moreover, the Czech translator added the explicating question marks, and the English translators replaced the ellipsis dots by n-dashes.

¹⁷ The French translators explicitate the logical relation (opposition – connective *Mais*) between the two coordinate clauses in Czech, and they turn the consecutive relation (*tak/therefore*) into a causal one (*comme/as*). The English translator replaced the specific introductory verb *odpověděl/answered* by the neutral *said* (see Nádvorníková 2017b).

absence of quotation marks, but the data show that this type of change is independent of this factor (see also example in note 24).¹⁸

However, the English translation in (6) illustrates another type of sentence splitting occurring in all directions of translation:

4.2.3. Sentence splitting before the connectives *and/et/a or but/mais/ale*

One most frequent point of splitting of sentences occurring in all directions of translation is before the connectives *and/et/a or but/mais/ale*. Bisiada (2016: 374) argues that splitting of sentences at this point “may be the least intrusive way of introducing full stops” (see also Nádvorníková 2017a). Effectively, in our data, we observed the splitting of sentences at this point independently of the direction of translation.

From Czech into French and English:

(7) (CS) Myslím na Tomáše už řadu let, // **ale** teprve ve světle této úvahy jsem ho uviděl jasně. (M. Kundera, *Nesnesitelná lehkost bytí*, 1985)

(fr) Il y a bien des années que je pense à Tomas. // **Mais** c'est à la lumière de ces réflexions que je l'ai vu clairement pour la première fois. (transl. Fr. Kérel, 1984)

(en) I have been thinking about Tomas for many years. // **But** only in the light of these reflections did I see him clearly. (transl. M.H. Heim, 1984)

In (7), both English and French translators split the source sentence in two before the connective *but/mais*. This type of splitting of sentences is “mostly harmless”, to quote an author whose novels are included in our corpus. In fact, there is no shift from subordination to coordination; the insertion of explicating connectives or anaphoric expressions is rare, because not necessary, and the only obvious change is the frequent elimination of the connective “a” (“and”) in translations into Czech:

(8) (EN) Gurgling he went under, // and the River closed over his curly head. (J.R.R. Tolkien, *The Fellowship of the Ring*, 1954)

(fr) Il s'enfonça en gargouillant, // et le fleuve se referma sur sa tête bouclée. (transl. F. Ledoux, 1988)

(cs) Zakloktal a ponořil se. // Řeka se zavřela nad jeho kučeravou hlavou.¹⁹ (transl. S Pošustová, 1990)

Thus, the connectives *and/a/et* and *but/ale/mais* seem the ideal “gap” for introducing the full stop, similarly to the loose adverbial relation conveyed by

¹⁸ (FR) – Oh ! J'ai très bien compris, fit le petit prince, // mais pourquoi parles-tu toujours par énigmes ? (en) "Oh! I understand you very well," said the little prince. // "But why do you always speak in riddles?" (cs) "Ó, já jsem ti dobře rozuměl," řekl malý princ. // "Ale proč mluvíš stále v hádankách?"

¹⁹ Cf. the same type of change in translation from French into Czech: (FR) A leur retour, les parents trouvèrent les petites qui mettaient la table en chantant, // et ils en furent choqués. (cs) Když se rodiče vrátili, zastihli holčičky, jak se zpěvem na rtech prostírají na stůl. // Pobouřilo je to.

participial adjuncts in 4.1.²⁰ However, the change may involve less visible shifts, at the level of the information structure. In fact, it may emphasize the split part (cf. the same observation in Bisiada 2016), especially when the split segment is short and placed at the end of the sentence:

(9) (EN) He assumed that this must be his great grandfather's doing, // **but** why? (D. Adams, *The Restaurant at the End of the Universe*, 1980)

(cs) Dospěl k závěru, že v tom musí být pradědečkův záměr. // **Ale** proč? (transl. J. Hollanová, 1999)

Splitting before the connective *a/and/et* is particularly frequent in (long) sentences, where the last (split) segment conveys the consequence or the dramatic conclusion of the preceding narrative sequence:

(10) (CS) Když se lupička snažila položit nohu do zasněženého klínu tvrdého y, jež se nacházelo přesně ve středu druhého slova a jehož spodní část dole opouštěla římsu a trčela do vzduchu, sklouzla jí noha **a** // Baumgarten s hrůzou viděl, jak žena sjíždí po ypsilonu do propasti. (M. Ajvaz, *Zlatý věk*, 2001)

(en) As the thief was attempting to place her foot in the snow-filled lap of the “y,” which was in the dead centre of the second word, with its lower part protruding from the ledge and into space, she slipped. // Baumgarten watched with horror as the woman slid down the “y” and towards the abyss. (transl. A. Oakland, 2010)

The Czech source sentence is quite complex, including two coordinate clauses and four subordinate ones. The English translator turned one of the finite subordinate clauses into a non-finite one (an absolute construction augmented by *with*), and, more importantly, split the sentence at the conjunction of the two coordinate clauses. The dramatic climax of the scene is thus separated from the rest, and the split adds to the text an explicating effect of suspense. If frequent and systematic – as is the case in English translations of both texts by Michal Ajvaz included in our corpus – the splitting of sentences may alter the overall style of the source text.²¹ As we observed in Tables 2 and 3, such a systematic splitting of sentences occurs in a minority of translations in our corpus, but its consequences are the more revealing.

4.2.4. Splitting of sentences as a global translation strategy

While evaluating the consequences of sentence splitting, it is necessary to take into account the whole text: on the one hand, the idiosyncrasies of the style of the source text, on the other hand, the relative frequency of these shifts in translations. As we have already said, the task of the translator is difficult, because he/she has to find the right balance between respect to the specific style of the source text,

²⁰ Moreover, in our data (non1:1 segments), we observed the *joining* of sentences at the same point (before the connectives).

²¹ In this text, the proportion of sentence splitting is 30% of all the alignments (cf. Table 2) and the English translator added more than 1,000 full stops (shift from 2,560 to 3,594 full stops), cf. Table 3. The shift in the other text by Michal Ajvaz are similar.

and the norms and conventions of the target language. Table 2 in section 3 shows that splitting of sentences (more specifically, adding of full stops) as a global and systematic translation strategy is exceptional, and slightly more frequent in translations from Czech into English than in the other directions of translation.

In translations from English into Czech, we find the highest proportion of split segments and added full stops (see Tables 2 and 3) in a bestselling novel by Nicholas Evans (*The Divide*): the Czech translator splits the source segments in a staccato of short sentences. Thus, the neutral narrative style of the source text is turned into a markedly segmented text in Czech:

(11) (EN) They saw mule deer and coyote // and just as the road turned to gravel a great pale-winged owl swerved from the cottonwoods // and glided low ahead of them as if piloting the beam of their lights. (N. Evans, *The Divide*, 2007)

(cs) Cestou zahlédli jeleny a kojota. // A právě když zatáčeli na štěrkovou cestu hlouběji do lešů, mihla se před nimi světlokřídlá sova. // Ladně se snesla z větve a odletěla před autem pryč, jako by mu ukazovala cestu. (transl. D. Břejlová)

In (11), the translator split the source sentence in three segments, at the point of the connective *and*.²² If these changes were motivated by a local necessity to compensate for a preceding joining of sentences or to adapt the rhythm of the passage, their impact would remain local. Nevertheless, the frequencies of non1:1 segments and of full stops for the whole text show that in this case, the sentence splitting is a global and systematic strategy of the translator. In fact, sentence splitting represents 29% of the whole number of segments in this text; and the translation contains by 3,554 full stops (by 28%) more than the source text.

In translations from French into Czech, three texts exceed the threshold of 20% of added full stops. In two cases, the splitting may be motivated by concern about the target audience (children and young people): the comic book *Le tour de Gaule d'Astérix*²³ and *Les Vacances du petit Nicolas* by René Goscinny. In the latter case, the most frequent type of sentence splitting is the least intrusive one (cf. section 4.2.2): replacing a comma by a full stop in the interposed introductory clause.²⁴ In the third text with a markedly high proportion of sentence splitting in Czech translation, the most frequent trigger of the change is a specificity of the style of the contemporary French author: the recurrent juxtaposition of several clauses describing successive narrative events:

(12) (FR) Elle défit ensuite un tissu bleu et en sortit un bâton à encre, // une forte odeur de santal s'en dégagea, // enfin, elle déroula un napperon en lattes de bambou où dormaient deux pinceaux. (A. Gavalda, *Ensemble, c'est tout*, 2004)

²² The translator also rendered the non-finite verb form *piloting* by a finite one (tendency observed in 4.1), but without sentence splitting.

²³ E.g. (FR) Nageons vers ce bateau, Obélix, // il va vers Lutèce ! (cs) Poplaveme k té lodi, Obelixi. // Pluje do Lutecie.

²⁴ (FR) — Ah non, a crié Mamert, // c'est pas juste! (cs) "To ne!" rozječel se Servác. // "To je nespravedlivé!

(cs) Rozvázala modrý hadřík a objevilo se plnicí pero. // **Najednou** ucítila silnou vůni santalového dřeva. // Nakonec rozhrnula ubrousek z bambusových vláken, v němž měla zabalené dva štětce. (transl. J. Prokop, 2006)

In (12), the translator split the source sentence at the first comma and then before the connective *enfin* (*finally*). In consequence, the fluid narration is segmented in separated events. Moreover, because of the splitting, the cohesive tie between the first two segments has to be explicitated by adding the adverb *najednou* (*suddenly*, cf. (1) and (2)); and the relation of inference between the second and the third clause is lost.

While observing translations from Czech into French and/or English, we find that this type of paratactic syntax, observed in the source text in (12), is the most frequent trigger of sentence splitting in translations in both languages: the specific, markedly paratactic style of the source text is segmented into shorter, simpler portions. This is particularly the case of the translations of novels by Bohumil Hrabal (see in detail 4.2.4.2). However, in a minority of texts hit by the systematic splitting of sentences in translation, the syntax of the Czech source text is not paratactic, but on the contrary, deliberately hierarchical and complex. In both cases, the consequence of the systematic splitting of sentences is a simplification of the target text and an important modification of its style.

4.2.4.1. Splitting of complex, hierarchical sentences

We observed systematic splitting of complex, hierarchical sentences in two texts in our corpus of translations from Czech: *Modlitba pro Kateřinu Horowitzovou* by Arnošt Lustig, where the number of full stops increased by 48% in the English translation (the intervention of the French translator was less important – only 14%), and *Babičky* by Petr Šabach. Especially the style of Arnošt Lustig is complex and solemn, corresponding to the gravity of the topic (the destiny of a young woman in a Nazi concentration camp). The English translator, however, systematically split the complex source sentences in shorter segments. The French translator maintained the complexity of the source sentence, adding only structuring brackets and the explicitating temporal conjunction *tandis que* (meanwhile):

(13) (CS) Pak stál rabín Dajem z Lodže stranou, **nevšímaje si** pana Rappaporta-Liebena, právě tak jako si ho nevšímá pan Brenske, a zpíval, **nechávací** jednu ruku na zatemnění okna kupé, aby je snad někdo z přítomných pánů nechtěl znenadání otevřít násilím, // a druhou ve vzduchu žehnal Kateřině Horovitzové, // která se bála, že rozumí, proč zpíval zpěv za mrtvé.

(en) Then Rabbi Dajem of Lodz stood to one side, **ignoring** Mr. Rappaport-Lieben just as he had ignored Mr. Brenske, and he went on singing, **keeping** one hand against the blackout curtain over the compartment window in case one of the gentlemen might suddenly try to force it **open**. // **With** the other hand, he pronounced a benediction over Katerina Horovitzova's **head**. // **She** was afraid she knew why he was singing the chant for the dead.

(fr) Alors le rabbin Dayem de Łódź s'écarta et, *ignorant* M. Rappaport-Lieben comme M. Brenske aussi l'ignorait, chanta encore, *une main sur* le store d'obscurcissement (au cas où il prendrait envie à l'un de ces messieurs de l'arracher de force), // tandis que de l'autre il bénissait avec de grands gestes Katarzyna Horowitz, // qui craignait de deviner pourquoi il chantait la prière des morts.

By putting all the aspects of the scene into one sentence, the author creates in one move one complex scene where everything happens at the same time. By splitting the sentence, the English translator fragments the scene, moreover shifting it from simultaneity to succession (by the past perfect in the second sentence). Being frequent and systematic, such a change of segmentation inevitably alters the style of the source text.

4.2.4.2. *Splitting of long sentences written in paratactic style*

Cumulation of several (mostly) coordinate clauses with only a loose logical relation between them is the trait of a specific, idiosyncratic style in fiction (see (12)). This specific syntax may be motivated for example by the intention of the narrator/author to imitate spontaneous, spoken narration. Among the texts manifesting systematic splitting of sentences in translation in our corpus, we find the work by two Czech authors whose style is known for such an approach: Bohumil Hrabal and Jáchym Topol. As shown in Table 3, the splitting of sentences exceeding 20% of increase concerns especially the translations into English, as may be illustrated by both texts by Jáchym Topol. In (14) for example, the English translator split the source sentence in six segments, whereas the French translator maintained the original structure:²⁵

(14) (CS) Teta Fridrichová ani žádná z mých tet u samotného porodu nebyly, // tajný porod řídily starší zkušené ženy, které jsou už mrtvé, // litoval jsem, že mé tety byly tehdy tak mladé, mohly by říct, kdo byla Lebova maminka, ale vlastně je to fuk!, // dívka, která porodila Leba, pak nejspíš zahynula ve shonu válečných dní, // snad zmizela v některém z posledních transportů na východ nebo podle tet skončila nejspíš v tyfovém hrobě, // za ilegální porod by stejně dostala kulku, to mi teta Fridrichová vysvětlila. (J. Topol, *Chladnou zemí*, 2009)

(en) None of my aunts, including Aunt Fridrich, was actually at the birth **itself**. // **It** was overseen by older, experienced women, who are all dead **now**. // **If** only my aunts hadn't been so young, they could've told me who Lebo's mum was, but who **cares!** // **The** girl who had Lebo probably lost her life during the **war**. // **Maybe** she went off on one of the last transports to the East, or maybe, like my aunts said, she met her end in a typhus **grave**. // **If** she'd been caught giving birth illegally, it would've meant a bullet for her anyway, Aunt Fridrich explained. (transl. A. Zucker, 2013)

(fr) Ni tata Fridrichová ni aucune de mes tantes n'était présente à l'accouchement // qui avait été opéré par des femmes plus âgées et expérimentées, et mortes à présent, // je regrettais que mes tantes aient été si jeunes à l'époque, sinon elles auraient pu dire qui était la mère de Lebo, même si ça ne changeait rien ! // la jeune femme qui avait donné naissance à Lebo avait dû mourir dans la tourmente des jours de guerre, // elle avait peut-être disparu dans un des derniers transports de

²⁵ Another contemporary Czech author, Petra Hůlová, is known for her long, paratactic sentences imitating spoken language. Šotolová (2013: 30) reports the German translator of Hůlová's novel *Stanice tajga* systematically split these sentences and added brackets.

femmes vers l'Est, ou alors, comme pensaient les tantes, fini dans une fosse de morts du typhus, // de toute façon, comme me l'avait expliqué tata Fridrichová, elle aurait eu droit à une balle dans la nuque pour avoir accouché illégalement. (transl. M. Canavaggio, 2012)

As the English translator split the source sentence at the syntactic and logical “gaps” between the coordinate clauses, the change does not entail any other modifications than simplification (e.g. no adding of anaphoric expressions, as observed in 4.1).²⁶ Nevertheless, similarly to (13), the style of the source text is modified. The French translator, however, respected the sentence boundaries of the source sentence, turning only the second coordinate clause into a subordinate relative clause.

In translations of the texts by Bohumil Hrabal, the effect of systematic splitting of sentences is even more visible, because the sentences are longer and – as we will see – the changes more complex. In fact, the relative frequency of the full stop in the novel *Příliš hlučná samota*, in which both English and French translators introduced important changes in segmentation, is 8,747 ipm, whereas in the reference corpus of Czech contemporary non-translated fiction, this number is almost six times higher (51,367 ipm, see Nádvorníková, forthcoming 1). In the example given in Appendix, both English and French translators split the long source sentence in seven sentences. The Czech sentence is structured only by commas and recurrent symptomatic words (*dívám se/I watch – vidím/I see* and *dívám se – myslím si/I think*); otherwise, the passage is a free flux of associations and thoughts. Both translators, however, structure the text and explicitate the logical relations in it using various devices: non-finite verb forms (-ing forms and *participe présent*, see section 4.1), subordinate clauses with explicating connectives (in English – *as, when, etc.*), and various types of punctuation marks – ellipsis dots (in French), dashes (in English) and colons and semi-colons (in both translations). The resulting target texts are more structured, more logical, conforming more to the stylistic norm of the target language(s), but the marks of the specific style of the source text are lost (cf. a similar observation in May (1997)). Thus, similarly to the translation of Jáchym Topol, the resulting effect is not simplification, but especially normalization (see Baker (1993) in section 2).²⁷

The analysis of the English and French translations of Bohumil Hrabal's texts reveals also the limitations of our research of sentence splitting: adding full stops is not the only way of structuring the target text, as other punctuation marks, especially the semi-colon, may play a similar role. In fact, translators use the semi-colon when they want to structure the target sentence without splitting it. This tendency is particularly strong in translations from Czech into French (e.g. in the

²⁶ We do not comment on changes within the sentence, such as the explicitation of the Czech PP *za ilegální porod* by the conditional *if*-clause in the final English segment, or the shift from the second Czech coordinate clause to a subordinate relative clause in the French translation.

²⁷ Gouanvic (2007: 73–74) observed the same effect while analyzing French translations of Hemingway: added punctuation marks and important modifications of the syntax erase the simplicity of the source text.

translation of another novel by Bohumil Hrabal, *Obsluhoval jsem anglického krále*, the French translator added only 15 full stops, but 106 semi-colons). More importantly, French translators add large amounts of semi-colons even in translations of Czech texts where there are none (e.g. in texts by P. Šabach or K. Legátová). It is true that in Czech, the relative frequency of the semi-colon in non-translated fiction is more than twice lower than in French (see Nádvorníková, forthcoming1), but massive adding of this punctuation mark in texts where it is absent may represent a considerable modification of the style of the source text.

It is also important to point out that Czech translators use the semi-colon in the same purpose as the French ones: so as to avoid the abrupt break that would be caused by the insertion of a full stop, they opt for a semi-colon (see this observation in Nádvorníková and Šotolová 2016). For this reason, the relative frequency of the semi-colon is five times higher in translations from French into Czech than in non-translated text in the Czech reference corpus (4,195 ipm in translations against 824 ipm in non-translated texts, Nádvorníková, forthcoming1). All these explanations indicate clearly that sentence splitting using the full stop has to be completed by the analysis of the use of other punctuation marks – but this is already the outline for further research.

5. Conclusion and further work

The aim of this paper was to analyze contexts of sentence splitting in translations of fiction in English, French and Czech and to identify consequences these shifts entail at the sentence level as well as at the text level. The quantitative analysis of non1:1 segments extracted from the InterCorp parallel corpus showed that on average, sentence splitting represented only 5% of the alignments. However, the variance in the data goes from almost zero to more than 50%, because of the differences in style of source texts, in translators' strategies and in translation traditions.

The most frequent language-specific context of sentence splitting, occurring only in translations into Czech, was the sententialization of non-finite verb forms. Our analysis, focused on English –ing forms and French *gérondifs* and *participes presents*, showed that translators opt for splitting especially in case of a loose semantic relation between the non-finite clause and the main clause (the meaning of accompanying circumstance). By raising the non-finite clause to a finite one, the hierarchy of information is modified; explicitation of cohesive ties, however, is rare (with the exception of adding of the adverb *přítom/in the same time*, explicitating the relation of simultaneity).

The most important context of sentence splitting occurring in all directions of translation was the introduction of a full stop before the connectives *and/et/a* and *but/mais/ale*. This point represents an ideal “gap” for introducing the full stop in the sentence with minimum of other modifications (except of the adding of emphasis), similarly to the loose semantic relation between the non-finite clause

and the main clause. In both cases, translators use sentence splitting in order to simplify the source sentence. Surprisingly, sentence splitting as global translation strategy was observed in our data more frequently in translations from Czech into English than in the opposite direction; and it concerned more frequently texts written in a paratactic style than those characterized by complex, hierarchically structured sentences. Systematic sentence splitting, erasing the idiosyncrasies of Czech source texts, may be explained by two universal features of translation: simplification and normalization, i.e. the tendency to meet as much as possible the expectations of the target audience.

Nevertheless, our findings are preliminary and contain many loose ends:

- what characterizes sentence splitting involving not full stop, but the other, less frequent final punctuation marks (ellipsis dots, question mark and exclamation mark)?
- what are other language-specific contexts of sentence splitting and what changes do they entail? (e.g. the sententialization of relative clauses involves the adding of anaphoric expressions);
- what are the differences in sentence splitting and its consequences in fiction and non-fiction (and in other text types)?
- what other factors influence the final amount of sentence splitting in a text? (e.g. editorial guidelines, interventions of text revisers, etc.)
- what are the contexts and consequences of the complementary shift – i.e. the joining of sentences – in translation?

By the present study, we hope, at least, to have demonstrated that an apparently minor change – the splitting of a sentence in translation – may involve numerous causes and consequences.

References

- Baker, Mona 1993. Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis and E. Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*, 233–250. Amsterdam-Philadelphia: John Benjamins. <https://doi.org/10.1075/z.64.15bak>
- Ballard, Michel. 2007. Étude traductologique sur corpus : La relative dans les traductions d'une nouvelle de Joyce. In Michel Ballard and Carmen Pineira-Tresmontant (eds.), *Les Corpus en linguistique et en traductologie*. 175–198. Arras: Artois Presses Université.
- Bisiada, Mario. 2016. 'Lösen Sie Schachtelsätze möglichst auf': The Impact of Editorial Guidelines on Sentence Splitting in German Business Article Translations. *Applied Linguistics* 37(3). 354–76. <https://doi.org/10.1093/applin/amu035>
- Bisiada, Mario. 2018. The editor's invisibility. Analysing editorial intervention in translation. *Target* 30(2). 1–23. <https://doi.org/10.1075/target.16116.bis>
- Blum-Kulka, Shoshana. 1986. Shifts of Cohesion and Coherence in Translation. In Juliane House and Shoshana Blum-Kulka. *Interlingual and Intercultural Communication*, 17–35. Tübingen: Narr.

- Čermák, František and Rosen, Alexandr. 2012. The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics*, 17(3). 411–427. <https://doi.org/10.1075/ijcl.17.3.05cer>
- Čermák, Petr et al. 2015. *Románské jazyky a čeština ve světle paralelních korpusů* [Romance languages and Czech in the light of parallel corpora]. Praha : Karolinum.
- Chlumská, Lucie. 2017. *Překladová čeština a její charakteristiky*. Praha: Nakladatelství Lidové noviny.
- Fabricius-Hansen, Cathrine. 1996. Informational density: a problem for translation and translation theory. *Linguistics* 34. 521–565. <https://doi.org/10.1515/ling.1996.34.3.521>
- Fabricius-Hansen, Cathrine. 1998. Informational density and translation, with special reference to German-Norwegian-English. *Language and Computers* 24. 197–234.
- Fabricius-Hansen, Cathrine. 1999. Information packaging and translation: aspects of translational sentence splitting (German–English/Norwegian). In Monika Doherty (ed.), *Sprachspezifische Aspekte der Informationsverteilung*. 175–214. Berlin: Akademie Verlag. <https://doi.org/10.1515/9783050078137-008>
- Gouanvic, Jean-Marc. 2007. *Pratique sociale de la traduction. Le roman réaliste américain dans le champ littéraire français (1920–1960)*. Arras: Artois Presses Université.
- Laviosa, Sara. 2006. *Corpus-based Translation Studies. Theory, Findings, Applications*. Amsterdam/New York: Rodopi.
- Laviosa-Braithwaite, Sara. 1996. *The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation*. PhD Thesis. Manchester: Centre for Translation and Intercultural Studies UMIST.
- Malá, Markéta and Šaldová, Pavlína. 2015. English non-finite participial clauses as seen through their Czech counterparts. *Nordic Journal of English Studies* 1. 232–257. <https://doi.org/10.35360/njes.346>
- Malmkjær, Kirsten and Windle, Kevin (eds.). 2012. *The Oxford Handbook of Translation Studies*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199239306.001.0001>
- Mauranen, Anna and Kujamäki, Pekka (eds.). 2004. *Translation Universals : Do they exist?* Amsterdam/Philadelphia : John Benjamins. <https://doi.org/10.1075/btl.48>
- May, Rachel. 1997. Sensible Elocution: How Translation Works in & upon Punctuation. *The Translator* 3(1). 1–20. <https://doi.org/10.1080/13556509.1997.10798985>
- Nádvorníková, Olga. 2010. The French gérondif and its Czech equivalents. In Čermák, František, Corness, Patrick, Klégr, Aleš (eds.), *InterCorp: Exploring a Multilingual Corpus*. 83–96. Praha: NLN/ÚČNK.
- Nádvorníková, Olga. 2013. Francouzský gérondif a český přechodník: kontrastivní analýza a jazykové korpusy. *Jazykovědné aktuality* [online] 50(3–4). 80–95. Available from: http://jazykovednesdruzeni.cz/JA_2013_3_4.pdf. [Accessed: 22nd July 2018].
- Nádvorníková, Olga. 2016. Le corpus multilingue InterCorp et les possibilités de son exploitation. In Éva Buchi, Jean-Paul Chauveau and Jean-Marie Pierrel (eds.), *Actes du XXVII^e Congrès international de linguistique et de philologie romanes (Nancy, 15–20 juillet 2013)*. 223–237. Société de linguistique romane/ÉLiPhi, Strasbourg. Available from: <http://www.atilf.fr/cilpr2013/actes/section-16.html>. [Accessed: 6th May 2018].
- Nádvorníková, Olga. 2017a. Parallel Corpus in Translation Studies: Analysis of Shifts in the Segmentation of Sentences in the Czech-English-French Part of the InterCorp Parallel Corpus. In Joseph Emonds and Markéta Janebová (eds.), *Language Use and Linguistic Structure*. 445–461. [Online] Olomouc: UPOL. Available from: <http://olinco.upol.cz/wp-content/uploads/2017/06/olinco-2016-proceedings.pdf> [Accessed: 22nd May 2018]
- Nádvorníková, Olga. Forthcoming1. The use of punctuation marks (English, Czech and French) in Reference, Parallel and Comparable Corpora: Question of Methodology. *EURALEX*.
- Nádvorníková, Olga. Forthcoming2. Non-finite forms translated by finite ones: Impact on the Syntax of fictional translated texts. *Proceedings of the Olomouc Linguistics Colloquium* (07/06/2018-09/06/2018). Olomouc: UPOL.

- Nádvorníková, Olga and Jovanka Šotolová. 2016. Změny v segmentaci na věty v překladových textech: analýza dat z francouzsko-českého paralelního korpusu. In Anna Čermáková, Chlumská, Lucie and Markéta Malá (eds.), *Jazykové paralely*. 188-235. Praha: ÚČNK/NLN.
- Pagnouille, Christine. 2004. Traduire les points et les virgules. In Michel Ballard and Lance Hewson (eds.), *Correct/Incorrect*. 33–40. Arras: Artois Presses Université.
- Robin, Edina. 2017. Translation Universals Revisited. *Forum*, 15(1). 51–66. <https://doi.org/10.1075/forum.15.1.03rob>
- Solfjeld, Kåre. 1996. Sententiality and translation strategies German-Norwegian. *Linguistics* 34. 567–590. <https://doi.org/10.1515/ling.1996.34.3.567>
- Šotolová, Jovanka. 2013. Sur le point-virgule et autres détails éphémères. *Etudes romanes de Brno* 34(1). 27–41.
- Vanderauwera, Rita. 1985. *Dutch Novels Translated into English: The Transformation of a "Minority" Literature*. Amsterdam : Rodopi. <https://doi.org/10.1163/9789004490284>

Appendix

(CS) Vrátil jsem prázdnou sklenici a přešel jsem koleje, písek parku křupal a vrzal jak zmrzlý sníh, ve větvích cvrlikali vrabci a pěnkavy, díval jsem se na kočárky a na maminky, jak sedí na lavičkách ve slunci a s vyvrácenými hlavami nastavují tváře léčivým paprskům, dlouho jsem stál před oválným bazénem, ve kterém se koupaly nahé děti, vidím bříška těch děcek, jak jsou označována proužkem od trenýrek a kalhotek, haličtí židé, chasidi, nosívali pásy jako znatelné a výrazné pruhy, které dělily tělo na dvě části, tu krásnější se srdcem a plicemi a játry a hlavou, a tu část lidského těla se střevy a pohlavními orgány jako část pouze trpěnou a tedy nedůležitou, vidím, jak katoličtí kněží ten pruh posunuli ještě výš, kolárkem si na krk dali viditelné znamení, které zdůrazňuje pouze hlavu jako miskou, do které si smáčí prsty sám Bůh, dívám se na ty koupající se děti a na jejich nahých tělech viditelné pruhy od kalhotek a trenýrek a vidím, jak řádové sestry krutým pruhem vykrojily z hlavy pouze tvář, obličej sevřený krunýřem škrobených kukel, tak jak to mají automobiloví závodníci formule jedna, dívám se na ty **stříkající a pohybující se** nahé děti a vidím, jak ty děti nic nevědí o pohlavním životě , a přeci už jejich pohlaví je v tiché dokonalosti , jak mne o tom poučil Lao-c', dívám se na ty pruhy kněží a řádových sester a pásy chasidů a myslím si, že lidské tělo jsou přesýpací hodiny, co je dole, to je i nahoře, a co je nahoře, to je i dole, dva do sebe zapasované trojúhelníky, pečtní znak krále Šalamouna, úměra mezi knihou jeho mladosti, Písni písní, a výsledkem jeho pohledu starého pána, marnost nad marnost, knihou Kazatel. (B. Hrabal, *Příliš hlučná samota*, 1994)

(en) **After returning** my empty glass, I crossed the tram tracks and walked on, the sand in the park **crunching** underfoot like frozen snow, the sparrows and finches **chirping**. // **I looked** at the babies in prams and the mothers on benches in the sun, their faces turned toward its healing **rays**; I stood before the oval pool, where naked children were playing , and noticed the stripes across their midriffs from the elastic in their **pants**. // **Hasidic** Jews in Galicia used to wear belts of loud, vivid stripes to cut the body in two, to separate the more acceptable part, which included the heart, lungs, liver, and head, from the part with the intestines and sexual organs, which was barely **tolerated**. // **Catholic** priests raised the line of demarcation, **making** the clerical collar a visible sign of the primacy of the head, where God in Person dips His **fingers**. // **As** I watched the children playing naked and saw the stripes across their midriffs, I thought of nuns, who sliced head from face with one cruel stripe, **stuffing** it into the armor of the starched coif like Formula One **drivers**. // **Those** naked children splashing away in the water didn't know a thing about sex, yet their sexual organs,

as Lao-tze taught me, were serenely **perfect**. // **And** when I considered the stripes of the priests and nuns and Hasidic Jews, I thought of the human body as an **hourglass** – **what** is down is up and what is up is **down** – a pair of locked triangles, Solomon's seal, the symmetry between the book of his youth, the Song of Songs, and the vanitas vanitatum of his maturity in the Book of Ecclesiastes. (transl. M.H. Heim, 1993)

(fr) Je rends mon verre vide et traverse les rails du **tramway** ; **le** sable dans le parc craque sous les pas, il crisse comme de la neige **gelée** ; **dans** les branches les moineaux et les fauvettes vocalisent à tout **rompre**. // **Je** regarde les landaus, les jeunes mères assises sur des bancs au soleil, la tête renversée, **offrant** leur visage aux rayons bienfaisants, je reste longtemps devant le bassin ovale où se baignent tout nus des petits enfants, **intrigué** par la marque des caoutchoucs de culotte sur leur **ventre**... // **En** Galicie, les juifs hassidiques portaient des ceintures éclatantes et vives, bandes colorées qui leur coupaient le corps en deux zones bien **tranchées** : **la** plus belle, celle du cœur, des poumons, du foie et de la tête, puis le reste, le négligeable, ce qu'on supporte, les boyaux et **le sexe**... // Cette ligne de démarcation, les prêtres catholiques l'ont fait remonter plus haut, jusqu'au **cou** : **leur** petit collet, ce n'est qu'un signe sensible de la primauté de la tête, où Dieu en personne se rince les **doigts** . // **Je** regarde ces petits enfants qui se baignent, leurs corps nus avec la trace visible des culottes et des shorts, **mais** je ne vois plus que des religieuses qui, d'un trait cruel, détachent leur visage de leur crâne pour l'encadrer dans la cuirasse de coiffes amidonnées, **tiens**, c'est comme les coureurs de formule 1 sous leur **casque** ... // Ces enfants nus qui gigotent et s'éclaboussent, je vois qu'ils ne savent rien des réalités sexuelles, et **pourtant** leur sexe est déjà dans une perfection tranquille, comme me l'enseigne **Lao-tseu**... // Je reviens au trait qui coupe en deux le corps des prêtres et des bonnes sœurs, je regarde les ceintures des juifs et je pense que le corps humain est un sablier, ce qui est en bas est en haut, et vice versa, deux triangles communicants, le sceau du roi Salomon, la moyenne de son œuvre de jeunesse et du bilan de l'âge sénile, le Cantique des Cantiques et l'Éclésiaste, vanitas vanitatum. (transl. M. Keller, 1983)