Tímea Kovács

Institute of English Studies, Department of English Linguistics
Károli Gáspár University of the Reformed Church in Hungary

Comparative analysis of lexical simplification in Hungarian-English translated and interpreted texts

https://doi.org/10.48040/PL.2024.1.5

Although research on translation and interpreting activities have different focuses, the importance of corpus linguistics in both disciplines has been highlighted over the last few decades. Nonetheless, relatively few intermodal corpora, including written and spoken texts and their (simultaneously) interpreted and translated counterparts, have been devised. The development of the EPTIC corpus was started by Bernardini et al. (2016) to fill this gap in the field of Corpus Linguistics. In the scope of their research, a multilingual intermodal corpus was built including the original speeches of European Parliamentary sessions in 2011, their verbatim (written) reports, the translations of the source language texts as well as the simultaneously interpreted texts. The aim of the EPTIC project is to examine and compare lexical simplification (Laviosa, 1998a, 1998b), in translated and interpreted texts in different language pairs and directions. In a similar vein, I aim to explore whether the texts interpreted in Hungarian–English directions are lexically simpler than the translated ones through the analysis of a bilingual intermodal sub-corpus.

Keywords: *intermodal sub-corpus, interpreted texts, translated texts Hungarian-English, simplification, morpho-syntactic differences*

Introduction

The importance of corpus linguistics in the field of translation studies was first highlighted by Mona Baker (1993, 1995) in two papers during the early and mid-1990s. A few years later, Shlesinger (1998) stressed the importance of corpus-based research in the field of interpreting. Over the last two decades, the relevance of corpus linguistics in both disciplines has been reinforced (Baker, 2000; Mikhailov–Cooper, 2016; Moratto–Li, 2023).

The basic aim of corpus linguistics is to establish regularities in the functioning of languages based on corpus-based or corpus-driven observations of large amounts of text available in electronic form, using computer analysis tools. There are abundant examples of different types of corpora, but inter-modal corpora, containing both translated and interpreted texts in the source and target languages, have hitherto received less attention. Although intermodal corpora have already been compiled (Schlesinger, 2009; Kajzer-Wietrzny, 2012), a relatively novel project, EPTIC (European Parliament Translation and Interpreting Corpus, 2016) is the first compilation that encompasses both a monolingual comparable and an inter-modal perspective (Bernardini et al. 2016). EPTIC is an intermodal corpus, as it includes source language spoken texts, their verbatim transcriptions, and the texts simultaneously interpreted and translated into the target language from the written texts (Bernardini et al., 2018). Therefore, it is a parallel and comparable multilingual corpus. EPTIC started as the continuation and extension of the EPIC (European Parliament Interpreting Corpus) corpus, which was compiled by Russo et al. (2006).

In an effort aimed at extending the EPIC project, Bernardini and her colleagues started the EPTIC project in 2016. EPTIC research and corpus compilation is carried out in collaboration with several universities (University of Bologna, University of Belgrade, the Université Catholique de Louvain, Adam Mickiewicz University, University of Ljubljana). The EPTIC corpus contains the speeches recorded during the 2011 European Parliament plenary sessions, their verbatim source-language transcriptions, target-language translations of

verbatim transcriptions and simultaneous interpreted target-language translations of the speeches in five languages so far: English, French, Italian, Polish, and Slovenian.

The aim of their research was to investigate the phenomenon of lexical simplification introduced by Laviosa (1998a, 1998b) in a comparable and parallel multilingual corpus. Based on their findings, Bernardini et al. (2016) and Ferraresi et al. (2018) established that in both translated and interpreted texts, simplification can be observed, but interpreters tend to simplify more than translators. The degree and nature of simplification is influenced by the languages activated in the language mediation and the direction of mediation.

Research on Lexical Simplification with Hungarian as a Pivot Language

Chesterman (2004) claims that simplification is a translation universal, so target (mediated) language texts, irrespective of the language pairs and the direction of mediation, are always simpler than original (non-mediated) texts. Lexical simplification as one of the translation universals proposed by Baker (1993) is also a feature of Hungarian-English mediation. According to Heltai *"the vocabulary of translated texts is more limited, and the average sentence length is shorter than that of non-translated texts"* (Heltai, 2002: 126). Nevertheless, relatively little focus has been given to the different salient linguistic tendencies in the process of translation and interpreting from Hungarian to English.

The recent period has witnessed an increase in the number of studies in the field of interpreting to or from Hungarian. Bakti has compiled a spoken (interpreting) learner corpus based on the recordings of practical examinations of interpreting students (2010). Márta Puklus (2019) has audio-recorded and analysed the texts of 5 court hearings. Henriette Szegh (2021) examines in her doctoral thesis the characteristics of interpreting texts. She conducted research in a French-Hungarian and Hungarian-French intermodal sub-corpus of the Pannonia corpus (Robin et al. 2017). Based on her findings, she concluded that interpreted texts in the investigated language pairs are lexically simpler. Her findings converge with those of other researchers who found that different lexical, grammatical, and stylistic tendencies can be observed in interpreted and translated texts (Shlesinger–Ordan, 2012; Defrancq et al., 2015; Bernardini et al., 2016; Ferraresi et al. 2018).

Research aims

The aim of this research is to examine whether texts interpreted from Hungarian to English display the same degree of simplification than translated ones and to highlight the most salient differences between interpreted and translated texts. The uniqueness of this research lies in the fact that the process of simplification in translation and interpreting has been subject to analysis involving languages belonging mainly to the Indo-European language family. Nevertheless, Hungarian and English are more distant language pairs in terms of their morpho-syntactic and lexical structure than typologically more related languages. With English belonging to the Indo-European language (Germanic) family and Hungarian belonging to the Finno-Ugric language family, the lexical and grammatical systems of the two languages differ in several aspects. English is characterised by analytical morpho-syntactic and lexical structuring, a synthetic sentence structure, an SVO-based word order, the rightward expansion of noun word structures, e.g. *'the girl standing in the corner'*, and subject prominence. Hungarian, however, can be characterised by different features, such as synthetic morphological and lexical structuring, an analytical sentence structure, a predominantly SVO-based word order with topic prominence, and the leftward expansion of noun word structures e.g. *'sarokban álló lány'* (Kovács, 2020a; Heltai, 2021).

Furthermore, the paper will also seek to examine whether interference has a stronger impact on translated or interpreted texts. According to Toury (1995), the law of interference as a translation universal stemming from the source text impacts the output of the translation process.

Presumably, the findings of this research will reveal more details about the linguistic forces governing the process of translation and interpreting in the interaction of two structurally and typologically distant languages, Hungarian and English. Hence, relying on quantitative data obtained by analysing an intermodal sub-corpus of original Hungarian (source language) and English (target language) written, spoken, and translated and interpreted texts with the help of Sketch Engine, the process of (lexical) simplification in Hungarian-English language direction will be examined. The process of simplification will be discussed from aspects such as word count, mean sentence length, the use of keywords, and the use of grammatical words versus content words, more particularly the use of personal pronouns. In addition to lexical simplification, the effect of interference stemming from structural differences between the Hungarian and English languages will also be investigated. The impact of interference will be examined in such structural differences such as the use of the general subject and the use of personal pronouns. Hence, in this paper, answers will be sought to the following questions:

(1) Are texts interpreted from Hungarian into English (lexically) simpler than translated ones?
(2) Is the impact of interference more observable in translated or interpreted texts?

Methodology

The present research is corpus-based, complemented by a comparative analysis of quantifiable data found in the original (source) Hungarian written and spoken texts and texts translated and (simultaneously) interpreted into the target (English) language. A total of four sub-corpora are used for the analysis. The corpus analysed is part of the Pannonia Corpus that has been compiled by the Research group of the Translation Studies Programme at Eötvös Loránd University (Robin et al. 2017). Its intermodal sub-corpus contains English and Hungarian European Parliamentary speeches (a total of 55 speeches), their verbatim written reports, their translations, the actual spoken texts in English or Hungarian, and their (simultaneously) interpreted versions. As the original sub-corpus contains no metadata as to the mode of delivery of the original speeches, whether the speeches were read out and delivered impromptu, the impact of the mode of delivery on lexical simplification will be excluded from the scope of analysis.

In a quest to find the most salient differences in translated and interpreted Hungarian-English texts, it is worth taking a look at the general characteristics of the four sub-corpora.

As can be inferred from Table 1, more similarities can be observed between the target language (English) texts irrespective of the modes of mediation (translation or interpreting) than in their source language counterparts.

Table 1. General corpus information

	HU_WR	**HU_EN_TR**	**HU_SP**	**HU_EN_INT**
tokens	17,068	21,190	15,546	18,142
words	14,417	19,058	14,294	16,483
documents	55	55	53[1]	55
sentences	795	774	780	740
words/sentences	18.1	24.6	18.3	22.3

[1] The transcription of the spoken Hungarian text is not available.

Sentences both in the English translated (HU_EN_TR) and interpreted (HU_EN_INT) sub-corpora contain more, respectively 19,058 and 16,483, words than the Hungarian source language texts. In the verbatim Hungarian (written) texts of the speeches (HU_WR), there are 14,417 words, while in the spoken Hungarian language texts (HU_SP), there is a total number of 14,294 words. This difference in the number of words in the original and mediated texts could be due to the morphologically analytic nature of the English language versus the synthetic nature of Hungarian. English is an inflectional language with isolating tendencies, hence, each morpheme tends to be a separate word, while in Hungarian, being an agglutinative language, suffixes and prefixes are "glued" to the root morpheme changing its original lexical meaning (Heltai, 2021: 209). Therefore, one whole Hungarian sentence (e.g. *Kerestelek*) could include one single word, while its English counterpart would be *I have been looking for you*, including six separate morphemes (Heltai, 2021: 200).

Nevertheless, there is a sizeable difference between the number of words in the English translated (19,058) and interpreted (16,483) corpora. This difference cannot be solely explained by the structural differences of the source and target languages, but rather by the different linguistic characteristics of translated and interpreted texts. This difference suggests that interpreted texts tend to be shorter or *"more reduced"* (Szegh, 2021) than translated ones.

One of the most relevant indicators of lexical simplicity is mean sentence length. As general corpus information in Table 1 suggests the mean sentence length seems to be more convergent in the same languages. English translated texts are characterised by the highest mean sentence length (24.6) followed by interpreted ones (22.3). Hungarian sentences in the original, non-mediated sub-corpus tend to be shorter. The Hungarian spoken sub-corpus shows a relatively higher mean sentence length (18.3) than its written counterpart (18.1).

Findings

As the above comparative general corpus information suggests, in terms of word count and mean sentence length, interpreted texts show a greater degree of simplification than translated texts, as they contain fewer words and shorter sentences. Sentence length, however, it is not a valid measure for spoken discourse, as spoken discourse contains sentence-like utterances but not 'full' sentences.

The comparative analysis of the use of single keywords could also suggest some salient tendencies. It shall be noted here that keywords for the scope of this analysis are defined "*as words which occur with unusual frequency in a given text […] by comparison with a reference corpus of some kind"* (Scott, 1997: 236) and not in the sense of 'keyness' as introduced by Gabrielatos (2018). In the Hungarian language sub-corpus, keywords have been extracted based on the frequency of their use compared to the Hungarian Web 2012 (huTenTen12) corpus, while for the extraction of English keywords, the general English corpus (enTenTen20) has been used as a reference corpus with the help of Sketch Engine. In view of the data included in Table 2, a more relevant difference can be observed in the use of keywords in texts mediated into English (translated and interpreted) and the original Hungarian written and spoken texts.

Table 2. The most frequently used words in the examined corpora

HU_WR	**HU_EN_TR**	**HU_SP**	**HU_EN_INT**
jelentéstevő ('rapporteur')	*madam*	*jelentéstevő ('rapporteur')*	*'ehm'*
progress-program	*barroso*	*'ö' ('er')*	*madam*
wto	*rapporteur*	*progress-program*	*eib*
vízummentesség	*eib*	*wto*	*rapporteur*

('visa free')			
gabonaár ('grain price')	*visa-free*	*vízummentesség ('visa-free')*	*epp*
dohai ('from Doha')	*inter-country*	*dohai ('from Doha')*	*orbán*
posting	*roma*	*posting*	*euro-zone*
špidla	*orbán*	*špidla*	*smes*
képviselőtárs ('fellow Member of Parliament')	*transposition*	*képviselőtárs ('fellow Member of Parliament')*	*barroso*
csomagterv ('package plan')	*špidla*	*csomagterv ('package plan')*	*'eeehm'*

Table 2 shows that the ten most frequently used keywords in the English interpreted text include '*ehm*' and '*eeehm*', which are filled pauses, typical of spoken language production and interpreting labelled as 'speech disfluencies' by Bakti and Kusztor (2017:1). Hence, they are absent in the written texts. In the Hungarian spoken text, the filled pause '*ö*' is also in the group of the ten most frequently used keywords, also a salient feature of spoken language production. In line with the above definition of keywords, filled pauses are listed among keywords as they are used with relatively more *frequency* in the speech production sub-corpus than in the reference corpus. In the section below, the actual number of single keywords will be examined.

Table 3. The actual number of single keywords

HU_WR	**HU_EN_TR**	**HU_SP**	**HU_EN_INT**
3,178	2,479	3,175	2,095

Table 3 shows that the sub-corpus of Hungarian written texts contains the largest number of single keywords (3,178), closely followed by the Hungarian spoken corpus (3,175). The English translated corpus contains a relatively higher number of keywords (2,479) as opposed to its interpreted counterpart (2,095). Words, functioning as 'fillers' with no substantial meaning can only be found in the spoken texts. Hence, it can be presumed that the number of actual keywords without fillers in the English interpreted corpus is even lower than what the quantitative set of data shows. It also seems to reinforce the assumption that interpreted texts are lexically simpler than translated ones.

Table 4. Wordlists and their frequencies

HU_WR	**FRQ**	**HU_EN_TR**	**FRQ**	**HU_SP**	**FRQ**	**HU_EN_INT**	**FRQ**
a ('the')	1,296	*the*	1,360	*a ('the')*	1,270	*the*	1,008
az ('the')	541	*of*	596	*az ('the')*	682	*to*	605
hogy ('that')	428	*to*	591	*hogy ('that')*	534	*of*	483
és ('and')	360	*and*	528	*és ('and')*	419	*and*	480
is ('also')	213	*in*	395	*is ('also')*	354	*that*	403
nem ('no')	180	*is*	375	*nem ('no')*	211	***we***	**389**
európai ('European')	115	*that*	371	*európai ('European')*	176	*in*	296
ez ('this')	105	*a*	312	*ez ('this')*	112	*is*	284

The use of personal pronouns

In line with previous expectations, relying on data in Table 4, it can be concluded that grammatical words are the most frequently used in the translated and interpreted texts as well (such as *'of', 'to', 'and', 'in', 'of', 'that', 'a'*). Nevertheless, a conspicuous difference can be observed. In the English interpreted text, the personal pronoun of '*we*' occurs 389 times, while in the translated sub-corpus it is not among the ten most frequently used words.

Taking a closer look at the original Hungarian written and spoken sub-corpus, it can be discerned that the Hungarian equivalent ('*mi*') of the English '*we*' does not appear. A plausible explanation of that phenomenon can lie again in the divergent morpho-syntactic nature of English and Hungarian. In Hungarian, the subject can be omitted in a sentence, while in English, the use of a subject is obligatory in any sentence (Heltai, 2021:211).

It should be noted here that in English the singular second-person personal pronoun 'you' is used as the general subject, while in Hungarian, the plural first person form 'mi' (*we*) is used. Looking at the use of the personal pronoun, 'mi' (*we*) in the original, translated and interpreted texts can give further ideas about the different linguistic patterns that emerge in translation and interpretation. Going into more detail about the use of '*we*' in mediated texts, the use of other subjective personal pronouns has also been subject to investigation. Comparative data in Table 5 display the use of personal pronouns in Hungarian and English translated and interpreted sub-corpora:

Table 5. The use of subjective personal pronouns in Hungarian-English sub-corpora

	HU_EN_TR	**HU_EN_INT**
I	196	216
you	58	159
he	5	8
she	5	2
we	265	389
they	58	57

As can be inferred from data shown above, there are significant differences in the use of '*you*' and '*we*' in the translated and interpreted sub-corpora. In the corpus of interpreted texts, '*you*' appears 159 times, while in translated texts, only 58 times. This finding seems to reinforce previous findings demonstrating that pronouns are used more frequently in interpreted than in translated texts (Shlesinger, 2008; Szegh, 2021). A plausible explanation for this is that spoken texts contain more situational, deictic expressions, including pronouns. In the section below, the most frequently used n-grams containing '*you*' are examined.

Table 6. The most frequent n-grams containing *'you'* in the examined corpora

HU_EN_TR	**FREQUENCY**	**HU_EN_INT**	**FREQUENCY**
Thank you	**30**	**Thank you**	**78**
Thank you for	7	you very	21
you for	7	you very much	21
Thank you very much	5	Thank you very	19
you very	5	Thank you very much	19
you very much	5	you President	12
Thank you very	5	if you	12
you for your attention	4	Thank you President	11
you for your	4	you Madam	9
Thank you for your	4	Thank you Madam	8

The most frequently occurring linguistic unit in which '*you*' is used is '*Thank You*' in both corpora. Nevertheless, in the interpreted corpus, it appears 78 times, while in the translated one, 30 times. As this expression is a common way of opening and ending an official speech and can also function as a speech filler, it might be repeated more often in an interpreting situation than in translation.

Upon comparing the above set of data to the original Hungarian written and spoken sub-corpora, i.e. the frequency of the Hungarian equivalent of '*Thank you*' ('*Köszönöm*'), a similar tendency can be observed. It should be noted here though that in the Hungarian equivalent term of '*Thank you*', no personal pronoun is used. It can be explained by the fact that in the Hungarian language, the subject is omitted unless particularly emphasised, as derivational suffixes attached to the main verb ('*köszön*') ('to say thank you') determine the subject. Furthermore, Hungarian is a pro-drop language (Heltai 2021: 211), as opposed to English, so '*you*' as the object complement of the English verb '*Thank*' is also absent in the Hungarian expression.

Nevertheless, in the Hungarian written corpus, '*köszönöm*' (Thank you) is mentioned 41 times, while in the spoken corpus 81 times. The difference emerging in the use of '*Thank you*' and '*Köszönöm*' seems to converge more in the dimension of written and spoken texts than in the same language corpus. Hence, the more frequent emergence of the word '*thank you*' in the interpreted corpus. In spoken texts, repetitions tend to appear more often than in written texts, especially if they are part of a set of linguistic expressions regularly used in public speeches. As has been shown in Table 5, subjective personal pronoun '*we*' is also used more frequently in the interpreted (389) than in the translated corpus (265). In terms of the use of other subjective personal pronouns, no relevant deviations can be found.

The use of subjective personal pronouns

It is worth examining the use of subjective personal pronouns in the source language (Hungarian) texts to see whether the salient differences in written and spoken speech production could have an impact on their distribution in the Hungarian written and spoken corpora as well. Data regarding the use of subjective personal pronouns are included in Table 7:

Table 7. The distribution of subjective personal pronouns in the Hungarian sub-corpora

	HU_WR	**HU_SP**
Én ('I')	25	25
te ('you')	-	-
ő ('he'/'she'/'it')	2	2
mi ('we')	26	27
ti ('you')	-	-
ők ('they')	4	4

The most conspicuous characteristic regarding the use of personal pronouns is the total absence of '*you*' (the second-person singular and plural personal pronouns, in Hungarian '*te*' and '*ti*') in the Hungarian sub-corpora. It is partly due to the different use of subjects in English and Hungarian, as in Hungarian, the use of the subject is not obligatory. There is another syntactic phenomenon that might offer some potential explanation to this phenomenon. While in English the general subject is expressed by the subjective personal pronoun '*you*', in Hungarian, verbal structures with derivational suffixes of first-person plural inflection ('*we*') are used.

Qualitative analysis

The qualitative analysis of examples taken from the sub-corpus (Example 1) could also highlight the different lexical patterns arising due to the relatively huge linguistic distance between English and Hungarian rather than the rules governing written and spoken speech production.

Example 1

(1) a
HU_WRITTEN
... azt gondolom, hogy *köszönettel tartozunk* Berés képviselő asszonynak, hogy kezdeményezte, lehetővé tette részben azt, hogy *rácsatlakozzunk* az ILO kezdeményezésére, ...

(1) b
HU_SPOKEN
Azt gondolom, hogy *köszönettel tartozunk* Berés képviselő asszonynak, hogy kezdeményezte lehetővé tette részben azt, hogy *rácsatlakozzunk* az ILO kezdeményezésére, ...

(1) c
HU_EN_TR
I believe *thanks are due* to Mrs Berés for initiating it, enabling *us* to join the ILO initiative, on the one hand …

(1) d
HU_EN_INT
I think that *we*'ve got to thank Madam Berés for having initiated this debate and here I mean that *we* should join the initiative of ILO …

As can be inferred from Example 1, in the translated Hungarian-English text (1 c), the personal pronoun '*we*' is not used at all, while in the interpreted extract (1 d), it appears two times. In the Hungarian source language written and spoken texts (1 a and b), no subjective personal pronouns are used at all. Nevertheless, the derivational suffixes of the verb, '*tartoz+unk*' in the collocation '*köszönettel tartozni*', imply that the subject of this verbal structure is '*we*'. However, as the subject is omitted in the Hungarian sentence, the main verb containing information comes at the end of the phrase ('*hogy köszönettel tartozunk*'). Therefore, the interpreter probably waits to identify the actual subject relying on the derivational suffixes attached to the root morpheme to start the English sentence. To reduce this source-target language lag (Gile, 2001) in decoding the actual subject of the Hungarian sentence, the interpreter uses fillers such as '*we've got to*' and repeat the subject '*we*' that is expressed in writing with the use of an impersonal finite structure *('thanks are due'*) without reference to the actual person involved in the action. This example demonstrates that in addition to the different nature of written and spoken discourse, the structural distance and the morpho-syntactic differences between Hungarian and English interfere more in the process of interpreting with the target language output than in translation.

Conclusions

In view of the findings of research conducted on an intermodal Hungarian and English sub-corpus, it can be stated that Hungarian-English interpreted texts are simpler on the basis of quantitative data gained through the analysis of the corpus, including word count, mean sentence lengths, the actual number of keywords, and the distribution of grammatical words.

It should be noted that in the Hungarian spoken corpus fewer relevant differences emerge as in the English interpreted corpus as compared to their respective written counterparts. More simplifying tendencies seem to emerge in the examined interpreted English corpus due to the specific interpreting strategies in mediated speech production determined by structural differences between the activated languages. The findings of the quantitative analysis, complemented by the qualitative analysis of one example taken from the corpora, suggest that interpreted English texts are simpler with regard to the number of words and keywords, and the relatively more frequent use of grammatical words and speech fillers. It is partly due to the salient characteristics of spoken speech production that contains more repetitions, false starts, and standard sets of expressions required by the given interpreting situation. However, interpreting strategies activated to overcome the difficulties arising due to the structural and typological distance between Hungarian and English languages also seem to impact lexical simplification.

The scope of this research covers the process of lexical simplicity in Hungarian-English translated and interpreted texts. Nevertheless, in view of the limited set of data, findings of this analysis can only suggest some salient tendencies but not significant results. With a view to understanding the process of simplification, more research should be done comparing translated and interpreted languages in view of the morpho-syntactic differences and structural distance between the activated languages.

References

Baker, M. (1995): Corpora in Translation Studies. An Overview and suggestions for Future Research. *Target.* 7/2. 223–245. https://doi.org/10.1075/target.7.2.03bak

Baker, M. (1993): Corpus Linguistics and Translation Studies. Implications and Applications. In: Baker, M. – Francis, G. – Tognini-Bonelli, E. (eds.) (1993): *Text and Technology: In honour of John Sinclair*. 233–250. Amsterdam: John Benjamins. https://doi.org/10.1075/z.64.15bak

Bakti, M. (2010): *Disharmonies in the target language output of simultaneous interpreters*. Unpublished PhD dissertation, Eötvös Loránd University. Online elérhető: https://doktori.btk.elte.hu/lingv/baktimaria/thesis.pdf

Bakti, M. – Kusztor, M. (2017): Speech errors in simultaneously interpreted German target language texts. A descriptive analysis. *Alkalmazott Nyelvtudomány*. 17/2. 1–17. Online elérhető: http://alkalmazottnyelvtudomany.hu/wordpress/wp-content/uploads/bakti.pdf

Bernardini, S. – Ferraresi, A. – Petrovic, M. M. (2016): From EPIC to EPTIC — Exploring simplification in interpreting and translation from an intermodal perspective. *Target*. 28/1. 61–86. https://doi.org/10.1075/target.28.1.03ber

Bernardini, S. – Ferraresi, A. – Russo, M. – Collard, C. – Defrancq, B. (2018): Building Interpreting and Intermodal Corpora: A How-to for a Formidable Task. In: Russo, M. – Bendazzoli, C. – Defrancq, B. (eds.) (2018): *Making Way in Corpus-based Interpreting Studies*. 21–41. Springer: Singapore. https://doi.org/10.1007/978-981-10-6199-8_2

Chesterman, A. (2004): Hypotheses about Translation Universals. In: Hansen, G. – Malmkjær, K. – Gile, D. (eds.) (2004): *Claims, Changes and Challenges in Translation Studies: Selected Contributions from the EST Congress*, Copenhagen 2001. 1–13. Benjamins Translation Library. https://doi.org/10.1075/btl.50.02che

Defrancq, B. – Plevoets, K. – Magnifico, C. (2015): Connective items in interpreting and translation: Where do they come from? In: *Yearbook of corpus linguistics and pragmatics*. 195–222. Springer: Cham. https://doi.org/10.1007/978-3-319-17948-3_9

Ferraresi, A. – Bernardini, S. – Petrović, M. – Lefer, M.A. (2018): Simplified or not simplified? The different guises of mediated English at the European Parliament. *Meta*. 63/3. 717–738. https://doi.org/10.7202/1060170ar

Gabrielatos, C. (2018): Keyness Analysis: nature, metrics and techniques. In: Taylor, C. – Marchi, A. (eds.) (2018): *Corpus Approaches to Discourse: A critical review*. 225–258. Oxford: Routledge. Online elérhető: https://www.academia.edu/34302240/Gabrielatos_C_2018_Keyness_Analysis_nature_metrics_and_techn

iques In Taylor C and Marchi A eds Corpus Approaches to Discourse A critical review Oxford Routledge 225 258

Gile, D. (2001): Consecutive vs. simultaneous: Which is more accurate? *Interpretation Studies*. 1/1. 8–20. https://doi.org/10.50837/istk.0103

Heltai, P. (2002): "Claims, Changes and Challenges". EST 3rd Congress in Copenhagen. 30 August–1 September 2001. *Fordítástudomány*. 4/1. 123–133.

Heltai, P. (2021): *Hungarian–English Constrastive Linguistic Studies*. L'Harmattan: Budapest.

Kajzer-Wietrzny, M. (2012): Interpreting universals and interpreting style. Unpublished PhD dissertation. Adam Mickiewicz University, Poznań, Poland. Online elérhető: https://www.academia.edu/2440862/Interpreting_Universals_and_Interpreting_Style_doctoral_dissertation

Kovács, T. (2020a): Interferenciajelenségek összehasonlító vizsgálata nem fordított magyar és angolról magyarra fordított jogi szaknyelvi szövegekben (A comparative study of linguistic interference phenomena in non-translated Hungarian and English- Hungarian translated legal texts). In: Szabóné Papp, J. (ed) (2020): *Alkalmazott Nyelvészeti Közlemények, Interdiszciplináris tanulmányok. Publications of the University of Miskolc*. 13/2. 150–165. Miskolc: Miskolci Egyetemi Kiadó. Online elérhető: https://matarka.hu/koz/ISSN_1788-9979/vol_13_no_2_2018/ISSN_1788-9979_vol_13_no_2_2018_150-164.pdf

Kovács, T. (2020b): Korpusznyelvészet és tolmácsolás (Corpus Linguistics and Interpreting). In: Adorján, M. – Kovács, T. (eds.) (2020): *Korpusznyelvészet és nyelvi közvetítés* (*Corpus Linguistics and Interlinguistic Mediation*). 73–91. L'Harmattan: Budapest.

Laviosa, S. (ed). (1998a): Special issue – The corpus-based approach: A new paradigm in translation studies. *Meta*. 43/4. 474–479. https://doi.org/10.7202/003424ar

Laviosa, S. (1998b): Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta*. 43/4. 557–570. https://doi.org/10.7202/003425ar

Puklus, M. (2019): A bírósági tolmácsolás nyelvi és nyelven kívüli tényezői, a bírói elvárások és a tolmács szerepe (Linguistic and extra-linguistic factors in court interpreting, judicial expectations and the role of the interpreter). Unpublished dissertation. Eötvös Loránd University.

Robin, E. – Götz, A. – Pataky, É. – Szegh, H. (2017): Translation Studies and Corpus Linguistics: Introducing the Pannonia Corpus. *Acta Universitatis Sapientiae, Philologica*. 9/3. 99–116. https://doi.org/10.1515/ausp-2017-0032

Russo, M. – Bendazzoli, C. – Sandrelli, A. (2006): Looking for Lexical Patterns in a Trilingual Corpus of Source and Interpreted Speeches: Extended Analysis of EPIC. *Forum*. 4/1. 221–254. https://doi.org/10.1075/forum.4.1.10rus

Scott, M. (1997): PC analysis of key words – And key key words. *System*. 25/2. 233–245. https://doi.org/10.1016/S0346-251X(97)00011-0

Szegh, H. (2021): *Harmadik kód a tolmácsolásban. Vajon létezik-e tolmácsolási szöveg? (The third code in interpretation. Is there an interpreting text?)*. PhD Dissertation. Eötvös Loránd University. Online elérhető: https://edit.elte.hu/xmlui/bitstream/handle/10831/63576/dissz_szegh_henriett_fordtud.pdf?sequence=1

Shlesinger, M. (1998): Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies. *Meta*. 43/4). 486–493. https://doi.org/10.7202/004136ar

Shlesinger, M. (2008): Towards a definition of Interpretese. An Inter-modal, corpus-based study. In: Hansen, G. – Chesterman, A. – Gerzymisch-Arbogast, H. (eds.) (2008): Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile. 237–253. Amsterdam: Benjamins Translation Library. https://doi.org/10.1075/btl.80.18shl

Toury, G. (1995): *Descriptive Translation Studies and beyond*. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/btl.100