Idiosyncratic Features of Interpreting Style

Marta Kajzer-Wietrzny
Adam Mickiewicz University, Poznań, Poland

ABSTRACT
Following a number of studies on translator style, this paper attempts to apply corpus linguistics methods used in translation studies and inspired by stylometry to verify whether interpreters also demonstrate features of individual style. Interpreting style is viewed here as a way of interpreting which is characterized by a consistent prominence of one or many linguistic features in interpretations from different source languages that make one's interpretations distinguishable from those of others, as well as from one's own speaking style. Style is approached from the point of view of authorship attribution and stylometry, i.e. described by quantitative linguistic discriminators, which in the present study are the parameters indicating the tendency to repetitiveness, informativeness and lexical sophistication.

KEYWORDS: corpus based interpreting studies, corpus based translation studies, interpreting style, stylometry.

Introduction

Not only do "translators receive minimal recognition for their work" (Venuti 1995:8), but so do interpreters. As is the case with translators, the work of conference interpreters has for decades been perceived as "invisible". Simultaneous interpreters, in particular, always remain in the shadow of the source speaker (especially when working in the booth). And yet, the presence of translators has been recognised and investigated by translation scholars (Hermans 1996, O'Sullivan 2003, Bosseaux 2004). Hermans is one of the first scholars to clearly acknowledge that the translator's "voice is there in the text itself, in every word of it" (1996:9). At the same time, however, he claims that having translated the text, the translator "disappears without textual trace", and that his or her voice "may remain entirely hidden behind that of the Narrator, rendering it impossible to detect in the translated text" (ibid.).

With the advent of corpus linguistics, it has been possible to trace signs of the translators' linguistic presence in a novel way. Significant progress in corpus-based research on translator style has been made especially by Baker (2000), Olohan (2004) and Saldanha (2011). In her seminal paper investigating translator's style, Baker (2000:246) introduced the stylometric approach to translation studies which "attempt[s] to capture the essence of the style of a particular author by reference to a variety of quantitative criteria" (McEnery and Oakes 2000:548). Such quantifiable features are in stylometry referred to as discriminators or style markers.

The aim of this paper is to extend this approach to the study of interpreting style, which has not yet been researched from the stylometric perspective in present day corpus-based investigations of interpreting. Both translation and interpreting serve the same goal: to enable communication between two parties that would otherwise not be able to communicate. Yet, there are obvious differences in translation and interpreting very much related to the mode of delivery that are important for the analysis of style. Being written, translation is subject to revisions, which gives the translator a greater control over the final work. On the other hand,

changes may be introduced not only by the translator, but also by editors. In the case of simultaneous interpreting, the process is instantaneous and irreversible. It is therefore possible that translator's style may be more difficult to identify than that of interpreter. Nevertheless, there have been successful attempts to uncover features of idiosyncratic translators' style with the use of corpus linguistics tools. It is therefore reasonable to believe that this approach might also help investigate features of interpreting style.

The assumptions discussed above lead to the formulation of the following hypotheses, which will be tested in this paper:

> 1) Individual interpreters are relatively consistent in their use of language with respect to repetitiveness, informativeness and lexical sophistication when interpreting from different source languages and this consistency indicates a strong and stable interpreting style.
> 2) The interpreting style of an interpreter is different from his speaking style.

Literature review and rationale

Baker is the first scholar to apply corpus methodology to the study of translation, arguing that it might be useful to investigate the presence of translators, and in particular the style of a literary translator. She suggests a new definition and a new approach to style, which she understood as "a kind of thumbprint that is expressed in a range of linguistic – as well as non-linguistic features" (Baker 2000:245). Along this line of reasoning, style can be revealed in subtle, unobtrusive linguistic habits and patterns of choice which are predominantly subconscious. In the same paper, Baker demonstrates how this novel approach can be successfully employed by comparing the work produced by two literary translators: these turned out to differ with respect to sentence length, level of repetition and predominant verb patterns, elements which could be perceived as being persistent linguistic features of their style.

Baker's innovative idea gave rise to other studies investigating recurring linguistic and non-linguistic patterns in the works of individual translators. Olohan (2004) uses keyword analysis of Venuti's translations to verify whether the foreignizing translation strategies he advocates are indeed reflected in his translations. Having scrutinized keyword lists of texts translated by Venuti, she identifies a number of individual words and expressions used by the translator to achieve the foreignizing effect, such as archaisms, features of formal discourse or expressions in a foreign language. In a recent study, Saldanha (2011:31) examines different definitions of writers' and translators' styles, and tries to put forward a clear theoretical and methodological framework for the investigation of the latter. She proposes the following revised definition of translator style:

A 'way of translation' which:

- is felt to be recognizable across a range of translations by the same translator,
- distinguishes the translator's work from that of others,
- constitutes a coherent pattern of choice,
- is 'motivated', in the sense that it has a discernable function or functions, and
- cannot be explained purely with reference to the author or source-text style, or as a result of linguistic constraints

(Saldanha 2011:31).

Saldanha then sets out to test her definition by comparing the styles of two different translators through a corpus-driven investigation. Her analysis shows that the translators under scrutiny differed with respect, for example, to the use of emphatic italics and of the *that* connective. Saldanha argues that these choices are motivated by different global translation strategies.

Since the studies discussed above managed to successfully point to differences in the style of individual translators, it seems natural to assume that there are also features that make interpreter style distinctive. Van Besien and Meuleman have already observed that interpreters' styles may be revealed in their "preferred and habitual approach to interpreting as evident in his or her selection of problem solving strategies" (2008:135). They analyse major global and local interpreting strategies in a qualitative and comparative case study, and conclude these were used differently by the two interpreters whose performance was investigated. One interpreter produced a 'lean' text, i.e. one consisting in less material than the source speaker due to shorter formulations and the omission of redundancies, meta-communications and all irrelevant comments provided at a lower speech rate and neutral intonation. By contrast, the output of the other interpreter was what the authors call more 'abundant', i.e. consisting of additions to the input in the forms of, among others, repairs and clarifications, delivered at a faster rate with more lively intonation (Van Besien and Meuleman 2008:152).

Knowing that interpreters may approach the interpreting task differently, the present paper focuses on spotting the linguistic characteristics of individual interpreters, or, to paraphrase Hermans' words (1996:6), the interpreter's voice and its 'textual trace'. Following the path set by Baker (2000:45) and Saldanha (2011:31) and the approach guiding stylometry, interpreting style is viewed here as a way of interpreting manifested by a consistent prominence of one or many linguistic features in interpretations from different source languages that makes one's interpretations distinguishable from those of others, as well as from one's own speaking style. The literature, which constitutes the major point of reference for the present research, is very much oriented towards the style of literary translators. It is believed, however, that interpreters also leave linguistic traces (style markers) in their work, but so far this notion has not yet been taken up in corpus-based interpreting studies, which makes the previous corpus-based research on translator's style the best available starting point.

The definition of style adopted in this study purposefully disregards the problem of subconscious or motivated choice as well as the linguistic features of the source text and the original speaker. At this point, it is difficult to establish which of the linguistic features result from a conscious decision of the interpreter, which of them were initially conscious but became automatic over time, and which ones depend purely on subconscious processes. It is believed that the presence of such features in discourse is most probably not subject to a conscious choice, but proving it remains beyond the scope of this analysis.

Similarly, the present study does not involve a comparative analysis of the original and interpreted speeches. Although such analysis could undoubtedly be very revealing, the purpose of this paper is to examine whether the linguistic features serving as style markers are consistently present in the performance of individual interpreters when they provide simultaneous interpretation of speeches delivered by different speakers in two different languages. Such consistency in the overall tendency to repetitiveness, informativeness, and lexical sophistication in interpreted discourse is treated here as an indication of style. As, due

to the oral nature of interpreting, one cannot compare features like the interpreters' use of italics, different features had to be selected for the present analysis. The stylometric approach adopted in this paper calls for features that are quantifiable and could serve as style markers. A comparison of interpreters' lexical patterns, which can be operationalized with a set of measurable quantitative parameters, appears to be an objective criterion. Such a choice will not only help to pinpoint differences between individual interpreters, but will also allow to determine the overall characteristics of their interpreted discourse, e.g. the pattern based on lexical density (i.e. the proportion of lexical words to all words in a speech) may be indicative of a greater overall informativeness of the utterance, as lexical words are responsible for carrying information.

The methodology for a similar investigation has already been developed and successfully applied by Laviosa (1998), who examines the core patterns of lexical use in translated and non-translated texts using a set of different discriminators. Focusing on the regularity of specific linguistic features, though not so much on their function, Laviosa's analysis identifies what she refers to as "core patterns of lexical use" characteristic of translational texts in English:

> (a) translated texts have relatively lower percentage of content words versus grammatical words (i.e. their lexical density is lower),
> (b) the proportion of high frequency words versus lower frequency words is relatively higher in translated texts,
> (c) the list head of a corpus of translated text accounts for a larger area of the corpus (i.e. the most frequent words are repeated more often),
> (d) the list head of translated texts contains fewer lemmas
>
> (Laviosa 1998:563).

The first feature indicates that a translated text typically consist of fewer content words that carry information than a non-translated text of the same length (content words are more 'diluted' with function words in translated texts, which decreases their informativeness). The second feature suggests that translations contain a higher number of words most frequently used in English, which probably also equals the use of most standard expressions, thus revealing lower lexical sophistication. The last two features indicate that the range of vocabulary is narrower in translations since the same words (or word forms) are repeated more frequently, i.e. translated texts are characterized by greater lexical repetitiveness.

Three of these parameters (namely a, b and c) will be applied in the present study to judge the interpreters' tendency to informativeness, lexical sophistication, and repetitiveness respectively. The method adopted by Laviosa in her seminal paper (1998) will be treated as a point of reference in determining whether individual interpreter style can be revealed in any of the three textual features. To achieve this goal, I will examine whether the oral output of two interpreters analysed in this study differ with respect to the three parameters and whether any of these features may be considered distinctive of individual interpreting style.

Data

The three parameters of repetitiveness, informativeness, and lexical sophistication have been tested on a dataset carefully selected for this study, consisting of six small sampled corpora of authentic interpreted and non-interpreted discourse delivered by two professional (EU accredited) conference interpreters: Interpreter One (I1) and Interpreter Two (I2), whose details are presented in Table 1.

Table 1: Corpora of interpreted and non-interpreted discourse delivered by I1 and I2.

Corpora / Interpreters	I1	I2
Interpreting corpora	French into English	
	I1_SI_FR_EN 720 words	I2_SI_FR_EN 720 words
	German into English	
	I1_SI_DE_EN 720 words	I2_SI_DE_EN 720 words
Reference corpora of non-interpreted discourse in English	I1_ORG_SP_EN 720 words	I2_ORG_SP_EN 720 words

The interpretations have been delivered by both interpreters from French and German into English at plenary sessions of the European Parliament, and their recordings are available on the European Parliament website (Media Library of the European Parliament, 2011). The source speakers (four French and German MEPs) were different for both interpreters, but the source speeches were delivered in the same setting (several Plenary sessions of the European Parliament) and in the same manner (they were all read out by the source speakers). Each interpreting corpus consists of samples of four interpretations, delivered by I1 and I2; the samples include 60-word fragments from the beginning, ending and middle sections of the transcript. Thus, each sample is 180 words long and each corpus counts 720 words. Such a selection method allows to minimize the impact of source speaker style and to make the style of a particular interpreter more conspicuous.

Given that the parameters used are particularly susceptible to corpus size, it was decided that the most reliable results could be obtained from corpora which were even in size and consisted of a collection of samples. As the style of the original author might influence the linguistic features of the target text or speech (Rybicki 2012:246), it was decided that each investigated interpreting corpus should not consist of long speeches delivered by one speaker, but rather a set of interpreting samples, whereby the source speakers were different.

The interpreted performances have been compared to a corpus of non-interpreted discourse produced by the same interpreters. These corpora of non-interpreted discourse include samples of transcribed recordings of the interpreters talking (without interruption) on subjects related to the European Union. Both interpreters are native speakers of English with a diploma in higher education, have interpreted for the European institutions for at least fifteen years, work for the European Parliament on a regular basis, and have both Romance and Germanic languages in their combination. Both are male and have lived outside Britain for many years at the time of recording. They were recorded for approximately an hour. At the beginning, they were asked warm-up questions about their education and their interpreting career, which were then followed by a set of questions that prompted them to elaborate on the European Union. The idea behind the questions regarding the European Union was to keep the interpreters talking on topics that they usually interpret. The answers to the warm-up questions are not included in the analysed corpora, but two replies may be important for the interpretation of the data presented in this paper. The answers have been summarised in

Table 2 below.

Table 2: Summary of the answers given by interpreters in the warm-up session.

Questions:	I1	I2
Can you see any difference while interpreting from different languages?	There is no difference with respect to the structure of the original language, but one feels less confident in interpretation from languages they acquire last.	In difficult cases one can transcode word for word from French into English, and the interpretation will be still comprehensible (which is impossible in interpretation from German).
Does the language you speak differ from the language of your interpretation?	The language of interpretation is determined by the style of the source speaker, which the interpreter is obliged to follow.	Interpretation involves higher register, formal language and greater use of technical terminology.

The slightly different approaches to interpreting displayed by the answers above seem to, at least partially, reflect the differences in the way they interpret. One could expect that linguistic features of I1's interpreted discourse would be far less consistent if indeed his priority was to mimic the style of the source speakers. In case of I2, one may expect that if transcoding from French is occasionally produced, the language of this interpretation would be less repetitive and could perhaps use more unconventional lexis. It would also be expected that the latter's interpretation would be more formal.

Methodology

This paper seeks to examine whether the oral output of I1 and I2 differs with respect to three parameters: repetitiveness, informativeness, and lexical sophistication (adapted from Laviosa 1998). Repetitiveness will be measured here by the percentage of the investigated corpus covered by its list head, i.e. the one hundred most frequently occurring words in the analysed corpus, calculated for the most part according to the guidelines followed by Laviosa in the study cited above (1998), albeit with one exception. Although she eventually decided to extend the list head to the 108 most frequent words in the analyzed corpora, the list in the present study is limited to 100 words. This decision has been taken based on Laviosa's initial guideline, i.e. the presumption made by Lo Cascio (1994) that the greatest frequency jump in a corpus can be observed beyond the hundred most frequent words. Such a choice facilitates the comparison of the corpora in the present study, which are very limited in size.

Informativeness will be indicated by lexical density, i.e. the proportion of lexical words to the running words. Finally, the proportion of high frequency words (see Appendix 2) in a corpus will point to the degree of lexical sophistication. It must be noted that, in the present study, lexical sophistication is calculated based on the list of the 200 most frequent words generated by Stubbs (1996, Laviosa's personal communication, see Appendix 2) from the Cobuild Bank of English, and supplied by Laviosa in her PhD dissertation (1998). Due to the small size of the analyzed corpora, such an alternative would generate more valid results than the application of the list suggested by Laufer and Nation (1995), which is most frequently used in similar explorations of larger corpora.

Calculating the percentage of the corpus covered by list head is particularly problematic. The problem lies in the fact that the bigger the corpus, the greater is the chance for the hundred most frequent words to be repeated. The corpora used in this study are exceptionally small. At this point, any potential differences between the analysed corpora in the number of running

words could affect the number of function words and high frequency words in each corpus and therefore impact the validity of the list head analysis. That is why, for the calculation of the three parameters, all respective corpora had to be adjusted to the even size of 720 words. Such a solution results in a significant loss of data (as certain transcripts had to be excluded from the analysis), but contributes to its balance, enhances control, and allows the researcher to generate comparable results. It is worth remembering, however, that a 720 word-long text (transcript) is not a representative sample, and therefore all results should be treated as indicative only.

Measure of statistical significance

Corpus-based analytical methods are used to point to similarities and dissimilarities between the examined corpora. The results obtained may, in turn, reveal certain trends and patterns, which however need to be tested for statistical significance. Oaks (1998:24) suggests that chi-square (χ^2) test is the most appropriate non-parametric test for such estimations. Chi-square value allows to establish the p-value that points to the degree of probability that the generalisations inferred from the results yielded from the sample are wrong. For the purpose of this study, it is assumed that the threshold of tolerance of an error be lower than 0.05 ($p<0.05$), which is consistent with the standards applied in many fields of research (Oakes 1998:9).

In the following analyses, interpreting corpora shall be compared to reference corpora of non-interpreted speeches produced in English. As a rule, a chi-square test will be run to verify the statistical significance of the difference between the values characteristic of the original speeches and the interpreting corpora and the χ^2 and p-values will be presented in tables just below raw frequencies or percentage values. All χ^2 tests will be computed on raw counts only.

Analysis

This section presents the results of the comparative analysis carried out between the non-interpreted discourse of each interpreter and their interpreting performance. In the case of interpreted discourse, the analysis covers the target speech (i.e. interpretation) only, as each of the interpreting corpora consists of carefully selected samples (fragments) of interpretations of speeches delivered originally by four different source speakers.

Repetitiveness

The first feature under examination is the tendency of each interpreter to repetitiveness in three contexts: interpreting from French, interpreting from German and a non-interpreting situation. An analysis of list heads allows the researcher to compare corpora of transcripts of speech produced in such situations with respect to lexical repetitiveness. Data in Table 3 and Table 5 regard I1 and I2 respectively. The percentage values stand for the proportion of the total sum of the one hundred most frequent words (based on frequency lists) in the analysed corpus. The higher the values, the bigger part of the corpus the list head represents and the more repetitive the language in the analysed corpus. Consequently, a higher value points to a greater degree of repetitiveness, meaning that the interpreter used the one hundred most frequent words in the corpus far more frequently than the remaining words in that corpus.

Table 3: Proportion of the corpus covered by list head analysed on 720 word corpora of interpreter I1's non-interpreted and interpreted discourse consisting of transcript samples.

	I1_ORG_SP_EN	I1_SI_FR_EN	I1_SI_DE_EN
% Text covered	68%	66%	73%

by list head			

Table 4: Repetitiveness of I1's interpreted and non-interpreted discourse: Results of Chi-square tests calculated on raw counts.

	I1_ORG_SP_EN vs. I1_SI_FR_EN	I1_ORG_SP_EN vs. I1_SI_DE_EN
χ^2	1.02	3.21
p-value	p>0.05	p>0.05

The set of data presented in

Table 3 seems to indicate that the language of interpreter I1 is most repetitive when he interprets from German, less so while he speaks in his own language, and the least when he delivers an interpretation from French. Thus, no homogenous pattern valid for all I1's interpretations is revealed. As for I1's interpreting style, the results obtained through the analysis of the two interpreting corpora do not differ significantly from the outcome calculated for the subcorpus of I1's spoken discourse. The question is whether the observed tendencies would hold true for a bigger and more representative corpora of I1's interpretations.

From the statistical analysis of the present small samples (see Table 4), it can be inferred that I1's speaking style and interpreting style do not differ significantly with respect to repetitiveness. This suggests that I1's speaking style has a great impact on his interpreting, and that, in terms of repetitiveness, he has not developed a uniform interpreting tendency that would be consistent in the interpretations from each source language and significantly different from his non-interpreted discourse. On the contrary, the interpretations from French and from German can be placed on the opposite ends of the continuum, with I1's non-interpreted discourse in-between. While both seem to be similar to the non-interpreted speech from the statistical perspective, they do deviate significantly from each other ($\chi^2 = 7.83$, $p<0.01$); this seems to indicate that I1's tendency to be more repetitive may be dependent on the nature of the source text. This is also in line withI1's statement that the language of his interpretation is dictated by the style of the source speaker. It would follow that the tendency to repetitiveness would be inconsistent across his interpretations. It seems also that the comment made by the other interpreter (I2) on transcoding the original message from French might actually apply to I1's interpreting fashion. If indeed he is so strongly guided by the style of the source speaker, then naturally his interpretation from French might include more immediate English equivalents, thus contributing to overall lower repetitiveness of his language in such situation.

A completely different tendency can be observed in the case of I2, whose interpreting style is significantly different from his speaking style with respect to repetitiveness. Regardless of the source language, I2 uses a significantly less repetitive language in simultaneous interpreting than in non-interpreted discourse. His interpretations from French and German are highly homogenous ($\chi^2 = 0.11$, $p>0.05$), which implies that his interpretation is either 'resistant' to the source language and the source text or source speaker influence, or that this influence affects him always in a very similar way. This is actually surprising, considering that I2 admitted (in personal communication) that in case of difficulties it may be possible to deliver a comprehensible interpretation by transcoding the message from French word-for-word. It implies that either he does not, in fact, resort to this option very often, or that it does not significantly affect the tendency to repetitiveness in his interpretation.

Table 5: Proportion of the corpus covered by list head analysed on 720 word corpora of interpreter I2's non-interpreted and interpreted discourse consisting of transcript samples.

	I2_ORG_SP_EN	I2_SI_FR_EN	I2_SI_DE_EN
% Text covered by list head	74%	67%	66%

Table 6: Repetitiveness of I2's interpreted and non-interpreted discourse: Results of Chi-square tests calculated on raw counts.

	I2_ORG_SP_EN vs. I2_SI_FR_EN	I2_ORG_SP_EN vs. I2_SI_DE_EN
χ^2	9.06	11.18
p-value	$p<0.01$	$p<0.001$

It is also interesting to compare interpreters to one another. Table 7 shows statistical differences with respect to the tendency to repetitiveness between each interpreter's output.

Table 7: Repetitiveness of I1 vs. I2: Results of Chi-square tests calculated on raw counts.

	I1_ORG_SP_EN vs. I2_ORG_SP_EN	I1_SI_FR_EN vs. I2_SI_FR_EN	I1_SI_DE_EN vs. I2_SI_DE_EN
χ^2	6.28	0.028	5.28
p-value	$p<0.025$	$p>0.05$	$p<0.025$

This comparison reveals that while in non-interpreted discourse, I2 manifests a greater tendency to repetition than I1, there are no significant differences in the way they interpret from French, and that I1's language when interpreting from German is more repetitive. As a result, regardless of their natural tendency to repetitiveness manifested in non-interpreted discourse (in which they differ), both interpreters are equally repetitive when working from French into English. This finding is in line with the pattern of repetitiveness in interpretations from French observed in a larger interpreting corpus (Kajzer-Wietrzny 2012), which may imply that the tendency to repetitiveness, or the lack of it, is heavily contingent on the source language, and subject to interference. At the same time, the statistically significant difference between I1's and I2's interpretation from German suggests that there is a high chance that one's interpreting style might be consistent enough to be less subject to the influence of the source language.

Informativeness

The next feature parameter tested in search of consistent traits of interpreting style is the level of informativeness, operationalized as the measure of lexical density. Lexical density is expressed as the percentage of content words – words that carry information in a text – to the total number of grammatical words in the corpus (Laviosa 1998:563). Lower lexical density suggests that the text is less informative, as information load is diluted with more function words. In other words, of two texts equal in length, the one containing more lexical words would be considered more informative.

The analysis of lexical density in

Table 8 expressed in percentage suggests that I1's interpretations from French are slightly less informative than interpretations from German and his non-interpreted discourse. As shown in
Table 9, in statistical terms, I1's output is very homogeneous with respect to lexical density, i.e. there is no statistically significant difference between the three corpora.

Table 8: Lexical density analyzed on 720 word corpora of interpreter I1's non-interpreted and interpreted discourse consisting of transcript samples.

	I1_ORG_SP_EN	I1_SI_FR_EN	I1_SI_DE_EN
Lexical density	57%	56%	57%

Table 9: Lexical density of I1: Results of Chi-square tests calculated on raw counts.

	I1_ORG_SP_EN vs. I1_SI_FR_EN	I1_ORG_SP_EN vs. I1_SI_DE_EN
χ^2	0.10	0.04
p-value	p>0.05	p>0.05

Although the source speeches of the interpretations are not available for analysis, one vital characteristic is known: they were read out. The lexical density of written texts tends to be generally higher than the lexical density of spoken texts. Since the source speeches of all interpretations were read out by the source speakers, the results presented in Table 8 and Table 9 are very surprising. This means that all interpretations could have been influenced by the written nature of the source speeches, while I1's non-interpreted discourse was not affected by written discourse, as it was not read out or supported by notes. One would therefore naturally expect all interpretations to be more lexically dense than non-interpreted speeches produced by the interpreters.
Table 10 and
Table 11 point to a similar tendency in all the analyzed corpora of I2. Lexical density in I2's interpretations and non-interpreted discourse reveals the same uniformity of all analyzed corpora, just as in the case of I1.

Table 10: Lexical density analysed on 720 word corpora of interpreter I2's non-interpreted and interpreted discourse consisting of transcript samples.

	I2_ORG_SP_EN	I2_SI_FR_EN	I2_SI_DE_EN
Lexical density	56%	55%	57%

Table 11: Lexical density of I2: Results of Chi-square tests calculated on raw counts.

	I2_ORG_SP_EN vs. I2_SI_FR_EN	I2_ORG_SP_EN vs. I2_SI_DE_EN
χ^2	0.04	0.23
p-value	p>0.05	p>0.05

The most likely explanation of such findings is that if there are any differences between the analyzed corpora of I1 and I2, then possibly the sample subject to examination in this study is too small to show them. This, however, suggests that these differences are not very pronounced.

Lexical sophistication

Finally, the last parameter chosen, i.e. lexical sophistication, allows us to determine the range of vocabulary used in the corpora. A higher proportion of the most frequently occurring English words in the analyzed corpus indicates a narrower range of vocabulary, i.e. greater lexical simplification. As can be seen in
Table 12 and Table 13, I1's interpreted and non-interpreted discourse is, with the exception of the interpretations from French, rather homogeneous as regards the range of vocabulary used. The interpretations from French are significantly more sophisticated from the lexical point of view than not only the non-interpreted speeches, but also the interpretations from German (χ^2=6.58, p<0.025). Thus, I1's interpretations from French and German seem to be significantly different with respect to the repetitiveness and range of vocabulary: in both cases, the interpretations from French are more lexically complex. Such an outcome once again testifies to the fact that I1's interpretations from French are more subject to interference. Such interference would be likely to result in a broader range of vocabulary, since many English words of higher register derive from Latin, on which the French language is largely dependent. The similarity between the French and English lexical items at the surface level might prompt the interpreter to reach for the closest equivalents. It is also possible that I1's interpretations are influenced both by the source language and the style of original French speakers, which might be different from that of the German ones. This explanation seems plausible, as the interpreter expressly stated that the way he interprets depends on the style of the source speaker.

Table 12: The proportion of high frequency words analyzed on 720 word corpora of interpreter I1's non-interpreted and interpreted discourse consisting of transcript samples.

	I1_ORG_SP_EN	I1_SI_FR_EN	I1_SI_DE_EN
% High frequency words	63%	55%	62%

Table 13: Lexical sophistication of I1's interpreted and non-interpreted discourse: Results of Chi-square tests calculated on raw counts.

	I1_ORG_SP_EN vs. I1_SI_FR_EN	I1_ORG_SP_EN vs. I1_SI_DE_EN
χ^2	8.99	0.19
p-value	p<0.01	p>0.05

The comparison of the range of vocabulary used by I2 in interpretations and non-interpreted discourse proves that the corpora are very homogeneous (see

Table 14 and
Table 15). Not only is there no difference between the interpretations from French and German and non-interpreted discourse, but neither is there any between the interpretations from the two respective languages (χ^2=0.28, p>0.05). This suggests that, with respect to the range of vocabulary, I2's speaking style does not differ from his interpreting style, and his use of vocabulary is not easily affected by interference.

Table 14: The proportion of high frequency words analyzed on 720 word corpora of interpreter I2's non-interpreted and interpreted discourse consisting of transcript samples.

	I2_ORG_SP_EN	I2_SI_FR_EN	I2_SI_DE_EN
% High frequency words	62%	57%	58%

Table 15: Lexical sophistication of I1's interpreted and non-interpreted discourse: Results of Chi-square tests calculated on raw counts.

	I2_ORG_SP_EN vs. I2_SI_FR_EN	I2_ORG_SP_EN vs. I2_SI_DE_EN
χ^2	3.73	1.96
p-value	p>0.05	p>0.05

Interpreters I1 and I2 do not differ significantly with respect to the incidence of high frequency words, either in interpretations from French (χ^2 =1.77, p>0.05) or from German (χ^2 =3.13, p>0.05). Neither is the overall tendency contradictory: in both cases generally, the interpretations are slightly more lexically sophisticated, i.e. less simplified than the non-interpreted discourse. Yet, the differences do not earn statistical significance. The two interpreters, however, do differ in the homogeneity of the analyzed corpora: I1's interpretations from French into English deviate significantly from his non-interpreted English discourse, while in contrast there is little difference between I2's interpretations and non-interpreted speeches.

Limitations of the study

This study provides a valuable insight into the tendencies that can be observed in interpretations from different source languages and non-interpreted speech by individual interpreters. Admittedly, however, the analysis has been carried out on a compilation of data that may not be ideal in every aspect. The first limitation regards the size of the samples. Corpus-based methodology requires an abundance of text for a hypothesis to be properly tested and the results to be statistically significant**.** This criterion has not been met, therefore the results should be approached carefully. Ideally, the same study should be replicated on a larger dataset.

Additionally, the recordings of non-interpreted speeches were carried out for the sake of the experiment, i.e. in a less formal context. The interpreters were prompted by questions to talk on subjects related to the European Union. They were not interrupted in their responses, which resembled up to two-minute monologues. It is worth remembering that the mere fact that they were answering a question could have affected the structure of their speech. Such a constraint could also have an impact on the tendency to repetition. These aspects must be taken into account in the interpretation of the results. Finally, as in any study in the field of Translation and Interpreting Studies conducted entirely on comparable corpora, it is impossible to verify to what extent certain results could have been influenced by the source speech.

Discussion

The main goal of this study was to determine whether idiosyncratic differences could be revealed in one's tendency to repetitiveness, informativeness, and lexical sophistication, and whether a consistent presence of any of these features could be perceived as a trace of interpreting style distinguishable from others' and from one's own speaking style. The methodology presented above proved to be a good starting point for such explorations. As is apparent from the analysis, I1 does not seem to manifest any consistent tendencies that could be discernible in all his interpretations from two different languages. Apart from subtle differences, usually not statistically significant, I1's interpretations and non-interpreted discourse appear to be very similar. As a matter of fact, no statistically significant and consistent difference was observed in his interpretations, with the exception of I1's interpretations from French, which present a lower proportion of high frequency words are

therefore more lexically sophisticated. I1's interpretations are rather consistent with his speaking style in that they do not deviate significantly from the non-interpreted discourse (except for the abovementioned tendency to lexical sophistication in interpretations from French). However, there seems to be a statistically significant difference with respect to repetitiveness and lexical sophistication between I1's interpretations from French and German, which suggests a significant impact of the source language or the style of the source speakers, and a lack of homogeneity in this respect in his interpretations.

As for I2, the analysis shows that his performance is fairly consistent both in interpretations and non-interpreted discourse, and that his interpretations are never lexically simplified. The only parameter that clearly points to a difference between I2's non-interpreted and interpreted discourse is the analysis of list heads, which indicates that, regardless of the source language, I2 is significantly less repetitive while interpreting than while speaking. There are two plausible reasons for such results. It might be the case that I2 puts less effort into self-monitoring while speaking and more while interpreting, thus limiting himself only to the essence of the original message and not allowing for repetitions when time is at stake (during simultaneous interpretation). Additionally, the speeches interpreted by I2 were read out by the source speakers, and were thus presumably more structured and less repetitious than the non-interpreted discourse that he produced. And yet, under similar conditions I1 does not reveal such a pattern, which indicates that self-monitoring might have had a greater impact.

It has been ascertained that interpreters I1 and I2 are different notably in one aspect: while I1 does not exhibit any consistent patterns of repetitiveness, informativeness, or lexical sophistication that could distinguish all his interpretations from non-interpreted discourse, I2 does prove to be less repetitive in all his interpretations, regardless of the source language. Based on the purely impressionistic opinion of the author of the present paper, the two interpreters represent also the two styles observed by Van Besien and Meuleman (2008), which have been partially reflected in the linguistic data. While I1 produced an 'abundant' target speech characterized by very brisk pace, dynamic intonation and the presence of clarifications, I2 seemed to deliver a more 'lean' interpretation with neutral intonation, while at the same time avoiding redundant information. Although the former characteristics has not been fully reflected in the outcomes of the present analysis, the latter has been revealed in a lower tendency to repetitiveness.

Although one of the three investigated linguistic features can be seen as distinctive to one of the interpreters' interpreting style, it is nevertheless striking that, in the majority of the remaining comparisons, there is an overall similarity between interpretations and non-interpreted discourse. This means that in the presented case study, interpreters' speaking style has a great impact on their interpreting or vice versa. Such results may have a significant impact on training. The fact that it is possible to preserve one's speaking style in interpreting confirms that interpreter training should pay particular attention to the constant improvement of students' speaking style. It seems that a 'strong' and 'stable' speaking style may be more resistant to the linguistic influences of a foreign language or the source language speaker, and perhaps sound more natural.

All the limitations discussed in the previous section constitute a challenge for the upcoming research in corpus-based interpreting studies, and should be considered in future investigations of interpreter style. Nevertheless, it has been shown that corpus linguistics can be successfully applied not only in the investigation of translated texts and translator style, but also in the analysis of interpreting.

mkajzer@wa.emu.edu.pl

Acknowledgments

I would like to express my sincere gratitude to the interpreters, who consented to the use of their interpreted speech in my study and who devoted their time to the recordings of non-interpreted discourse made for the purpose of my research. I would also like to extend thanks to the Directorate General for Interpretation for allowing me to use transcripts of interpretations in my research.

This research was financed by the Polish Ministry of Education (grant No. N N104 053239).

References

Baker, Mona (2000) 'Towards a methodology for investigating the style of a literary translator', *Target*, 12(2): 241-266.

Bosseaux, Charlotte (2004) 'Translating point of view: A corpus-based study', Language Matters*: Studies in the Languages of Southern Africa: 35(1), Special Issue: Corpus-Based Translation Studies - Research and Applications*: 259-274

European Parliament (2011) Plenary Sessions, Available online at [http://www.europarl.europa.eu/ep-live/en/plenary/introduction].

Hermans, Theo (1996) 'The translator's voice in translated narrative', *Target*, 8: 23-48.

Kajzer-Wietrzny, Marta (2012) *Interpreting universals and interpreting style* (unpublished PhD dissertation).

Laviosa-Braithwaite, Sara (1996) *The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation* (unpublished PhD dissertation).

Laviosa, Sara (1998) 'Core patterns of lexical use in a comparable corpus of English narrative prose', *Meta*, 43(4): 557-570.

Media Library of the European Parliament (2011) Available online at [http://audiovisual.europarl.europa.eu] (accessed 4 September 2011).

McEnery, Tony – Michael Oakes (2012)'Authorship identification and computational stylometry', in: Dale et al (eds) *Handbook of Natural Language Processing*, New York: Dekker

Oakes, Michael (1998) *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Olohan, Maeve (2004) *Introducing Corpora in Translation Studies*, London and New York: Routledge.

O'Sullivan, Emer (2003) 'Narratology Meets Translation Studies, or The Voice of the Translator in Children's Literature', *Meta* 48 (1–2): 197–207.

Rybicki, Jan (2012) 'The Great Mystery of the (Almost) Invisible Translator: Stylometry in Translation', in Micheal Oakes and Meng Ji (eds) *Quantitative Methods in Corpus-Based Translation Studies*, Amsterdam: John Benjamins.

Saldanha, Gabriela (2010) 'Translator style: methodological considerations', *The Translator*, 17(1): 25-50.

Stubbs, Michael (1996) *Text and Corpus Analysis: Corpus-Assisted Studies of Language and Culture*, Oxford: Blackwell.

Van Besien, Fred and Chris Meuleman (2008) 'Style differences among simultaneous interpreters', *The Translator*, 14(1): 135-152.

Venuti, Lawrence (1995) *The Translator's Invisibility: A History of Translation*, London: Routledge.

Appendix 1: function words

The list of function words used in the calculation of lexical density.

a, above, across, after, against, all, along, alongside, although, amid, amidst, among, amongst, an, and, any, anybody, anything, anywhere, apropos, as, at, atop, because, before, behind, below, beneath, beside, besides, between, beyond, both, but, can, can't, cos, could, couldn't, dare, daren't, despite, doesn't, don't, during, each, either, every, everybody, everyone, everywhere, except, few, for, from, he, he'd, he'll, he's, her, hers, herself, him, himself, his, how, however, if, in, inside, into, it, it'd, it's, its, itself, many, may, mayn't, me, mhm, might, mine, minus, much, must, mustn't, my, myself, needn't, neither, never, nevertheless, no, no-one, nobody, non, nonetheless, no-one, nor, not, notwithstanding, of, off, on, or, ought, oughtn't, our, ours, ourselves, out, outside, over, per, plus, shall, shan't, she, she'd, she'll, she's, should, shouldn't, since, so, some, somebody, someone, than, that, that'd, that'll, that's, the, thee, their, theirs, them, themself, themselves, then, there, there'd, there's, there've, these, they, they'd, they'll, they're, they've, thine, this, those, thou, though, through, throughout, thy, till, to, toward, towards, uhuh, under, underneath, until, up, upon, us, via, we, we'd, we'll, we're, we've, what, what'd, what's, what've, whatever, when, whenever, where, wherever, which, whichever, while, whilst, who, whom, whose, why, will, with, within, without, won't, would, wouldn't, ye, yeah, yes, yet, you, you'd, you'll, you're, you've, your, yours, yourself, yourselves, I, I'd, I'll, I'm, I've.

Appendix 2: high frequency words

The list of the 200 most frequent words in English identified by Stubbs (1996: 36-37) and used by Laviosa (1996: 119) in her investigation of translational English.

a, about, after, again, against, all, also, always, an, and, another, any, are, around, as, at, away, back, be, because, been, before, being, between, both, but, by, came, can, children, come, could, course, day, did, didn't, do, does, don't, down, each, end, er, even, every, fact, far, few, find, first, for, from, get, go, going, good, got, great, had, has, have, he, her, here, him, his, home, house, how, i, i'm, if, in, into, is, it, its, it's, just, kind, know, last, left, life, like, little, long, look, looked, made, make, man, many, may, me, mean, men, might, more, most, mr, much, must, my, never, new, no, not, nothing, now, of, off, oh, old, on, once, one, only, or, other, our, out, over, own, part, people, perhaps, place, put, quite, rather, really, right, said, same, say, says, see, she, should, so, some, something, sort, still, such, take, than, that, that's, the, their, them, then, there, these, they, thing, things, think, this, those, though, thought, three, through, time, to, too, two, under, up, us, used, very, want, was, way, we, well, went, were, what, when, where, which, while, who, why, will, with, without, work, world, would, year, years, yes, you, your.