Fluency in rendering numbers in simultaneous interpreting

(Article begins on next page)

12 September 2024

Fluency in rendering numbers in simultaneous interpreting

Marta Kajzer-Wietrzny, Ilmari Ivaska and Adriano Ferraresi
Adam Mickiewicz University | University of Turku | University of Bologna

Abstract

There is general consensus among interpreting practitioners and scholars that numbers pose particular problems in simultaneous interpreting. Adopting the view that fluency disruptions in interpreters' renditions are signals of cognitive processing problems, the authors aim to isolate those contextual and textual factors which increase the likelihood of disfluencies when rendering numbers present in a source speech. In the reported study, we analyse data from the European Parliament Translation and Interpreting Corpus (EPTIC): we focus on target-text segments whose corresponding source segment contains a number and we find the best predictors of disfluencies by applying a generalized linear mixed model. Our approach is confirmatory and so the model accounts for factors that have been suggested in earlier studies as being associated with interpreting fluency. These factors include the nativeness of the original speaker, the type of number, the frequency of numbers in the same sentence, omission, language pair and whether the text was originally delivered impromptu or read out, and at what pace. The outcomes suggest that important predictors of disfluent renditions include omission, the frequency of numbers in a sentence and the type of number; these can be said to contribute to interpreters' cognitive load when they process numbers.

Key words: interpreting, fluency, cognitive effort, numbers, multilingual parallel corpus

1. Introduction

Whereas it is difficult to find research that points to numbers as a source of problems for native speakers in monolingual speech production, numbers are expressly referred to as 'problem triggers' in the context of interpreting (Plevoets & Defrancq 2016). Difficulties associated with interpreting numbers are reported by both practitioners and scholars of interpreting, and experimental studies suggest that their processing is particularly demanding from a cognitive point of view. This leads to a relatively large proportion of errors or omissions in their rendering (see e.g. Mazza 2001; Pellatt 2006).

Viewing disfluencies as indicative of cognitive difficulties, in this article we investigate the disfluencies that occur when simultaneous interpreters have to render the numbers present in a source speech. The study reported here is based on a review of the different factors that can affect the interpretation of numbers, including the language pair and speaker- and speech-related contextual factors (the nativeness of the speaker, whether the speech was delivered impromptu or read out, and at what pace); it also considered textual factors (type and frequency of numbers in the immediate co-text). Our aim was to assess whether these factors discussed

in previous studies significantly increase the likelihood of disfluencies and, if so, which of them. Therefore we aimed to identify those factors that are likely to make the processing of numbers most difficult. The dataset we drew on is a multilingual parallel corpus of plenary debates delivered at the European Parliament and their simultaneous interpretations.

2. Numbers as a problem trigger

Numbers have generally not been reported as being a difficulty in monolingual native communication (Plevoets & Defrancq 2016). But with interpreting, both practitioners and scholars widely acknowledge the difficulty of rendering numbers, especially in simultaneous interpretations (Pellatt 2006), and many experimental studies have been devoted to investigating the phenomenon (see Braun & Clarici 1996; Desmet et al. 2018; Korpal & Stachowiak-Szymczak 2018; Mazza 2001). Several reasons have been suggested to explain why numbers are particularly challenging, the most important being low predictability, low redundancy and high informational content (Mazza 2001; Pinochi 2010). Numbers are difficult to reconstruct based on context alone and therefore the usual anticipation and reformulation strategies frequently fail to guarantee their accurate rendering. The informational content to be stored in working memory and subsequently rendered in the target language is also usually multidimensional: numbers can have an arithmetic value (e.g. 7, 21), can be expressed as an order of magnitude or an approximate value (e.g. thousand, million), can be associated with a unit of measurement (e.g. metres, kilograms) or another referent (e.g. a measure of physical distance, time, etc.), and numbers can also be compared with other numbers having the same or different properties (Jones 2014).

Studies on numbers in simultaneous interpreting carried out to date confirm that numbers are difficult to process. Experiments involving both professional and trainee interpreters report inaccuracy rates in the rendering of numbers, including errors and omissions, of up to almost 50% (Braun & Clarici 1996; Desmet et al. 2018; Korpal 2016; Pinochi 2010). Interpreters become more accurate with numbers when they use notes (Frittella 2019; Mazza 2001) or technological support (Desmet et al. 2018; Frittella 2022). A study by Collard and Defrancq (2020) based on authentic, non-experimental data of French–Dutch interpretations from the European Parliament reports an error rate of about 18%. The authors attribute this relatively low error rate to good booth collaboration.

Among the factors which increase the burden posed by numbers on working memory (WM), the length of a numerical expression features prominently: longer numerical

expressions require more processing time and larger numbers are more likely to exert a greater effect on WM than small numbers (Pinochi 2010). Moreover, speakers may utter sequences of numerical expressions one after the other. Interpreting a series of numbers in the source speech is more difficult as the interpreter "must process, retain and translate more information per unit of time" (Mazza 2001: 90). According to Plevoets and Defrancq (2016), professional interpreters produce more disfluencies if the source text contains more numerals. If the target texts contain numerals, however, "the interpreter tends to produce fewer *uh(m)*'s" (Plevoets & Defrancq 2016: 218). In a later study, Plevoets and Defrancq (2018: 24) did not confirm the relationship between numbers and cognitive load but suggested that omission might play a role "in the reduction of cognitive load, particularly in the case of numbers".

Other factors have been suggested to increase the cognitive difficulty of interpreting numbers. Mazza (2001) reports that numbers with four or more digits pose the greatest difficulty, followed by decimals, ranges, small whole numbers and dates. Kajzer-Wietrzny et al. (2021) provided further evidence of this tendency; however, they showed that ranges are rendered most accurately. Language pair and interpreting direction can also play a role: Kajzer-Wietrzny et al. (2021) reported an up to 10% difference in accuracy rates depending on the language pair and directionality. Braun and Clarici (1996) observed more errors in interpreting from German into Italian than vice versa, hypothesising an effect related to the length of numerical expressions in the two languages. However, this was not observed in a study of the same language pair by Pinochi (2010), where comparable accuracy was observed in interpretations in both directions. Among speaker-related factors, the delivery rate also makes a difference. Korpal (2016) reported that interpreters render numbers less accurately when the delivery rate of the speech is faster. Similarly, Kajzer-Wietrzny et al. (2021) observed an effect of delivery rate on the number of errors: the effect was visible, in particular, in interpretations from Italian into English (but not in the opposite direction). Moreover, numbers in speeches delivered by non-native speakers were interpreted significantly less accurately than those delivered by native speakers.

But the same study (Kajzer-Wietrzny et al. 2021) showed that whereas errors in rendering numbers can be attributed to both contextual and textual factors, omissions (i.e., numbers not rendered at all) are explained best by speaker-related differences. They concluded that omissions and errors are fundamentally different phenomena. The tentative explanation that the authors provided is that omissions may stem from strategic choices related to the interpreters' management of their cognitive efforts. Errors, on the other hand, occur

subconsciously, that is, they are not introduced by interpreters on purpose and are more likely to be an effect of cognitive processing demands that are excessive.

The above overview indicates that the difficulty associated with interpreting numbers is not only a matter of interpreters' subjective perception but has been confirmed by experimental studies and studies based on data obtained in authentic interpreting situations. However, to date, no study has investigated disfluencies in the rendering of numbers in authentic interpreting contexts in more than one language combination.

3. Disfluencies as an indicator of cognitive effort

It is assumed that disfluencies – that is, "phenomena that interrupt the flow of speech and do not add propositional content to an utterance" (Gósy 2007: 93) – are associated with increased cognitive effort during speech production. The study by Gósy (2007) accounts for three types of disfluency: filled pauses, empty pauses (also called silent pauses or hesitations)[1] and false starts or truncated words. Filled pauses such as "ah", "ehm" and "eh" enable a speaker to plan their utterance (Fehringer & Fry 2007), but they may also hint at formulation problems. Empty pauses are pauses in speech production when no sounds are articulated. They may vary in length and, like filled pauses, are related to difficulties in planning speech online (Cecot 2001). Finally, "[f]alse starts are devices to correct what has been said immediately before" (Maclay & Osgood 1959; cited in Tissi 2000: 106).

In simultaneous interpreting, both filled and empty pauses and false starts are associated with processing difficulties. For example, empty pauses may signal comprehension difficulties and problems in finding equivalents and expressing the sense of the message in the target text (Chmiel et al. 2017). When compared to speaking, interpreting is characterized by more slips and shifts (Pöchhacker 1995). There are fewer empty pauses during interpreting, but they are longer (Tissi 2000). Chmiel et al. (2017: 1208) suggest that this may point to differences between the planning activity during interpreting and monolingual speech production. They also note that interpreters' pauses may be related to problems or be used strategically, for example, to gain time. The number of pauses seems to be sensitive to the level of interpreting

---

[1] Empty pauses may be either intentional or result from hesitations. When the speaker uses them intentionally, it is usually to organize information units (Shlesinger 1994) or to signal specific communicative intentions (Cecot 2001). Intentional pauses are not the main topic of this article; here we use the term "empty pauses" as this was the term chosen by the creators of the corpus that is analyzed in the present study.

expertise, while their length is associated with linguistic expertise and language direction (Piccaluga et al. 2005: 154).

Plevoets and Defrancq (2016) also reported that the number of filled pauses in interpreting increases when a source speech is delivered at a fast pace. Moreover, the authors observed significantly more filled pauses during interpreting than in independent speech production, which they associate with the increased lexical density of the target text. They also suggested that fluency might be dependent on the frequency of the occurrence of numerals in the source texts and the frequency of the occurrence of numerals in the target texts. In a later study, the same authors found "intra-word filled pauses" to occur significantly more frequently in interpretations than in non-interpreted speech (Defrancq & Plevoets 2018: 58).

Based on the reviewed literature, we can conclude that the lack of verbal fluency may hint at processing difficulties and that such difficulties are more likely to occur in interpreting when source texts contain numbers. By investigating fluency parameters, we consider it possible to estimate the conditions under which numbers are most demanding for interpreters, a problem which we deal with by using advanced methods of statistical modelling.

4. Research objectives

The aim of this article is to reveal whether the factors discussed in previous studies increase the likelihood of disfluent renditions of numbers in interpreting and, if they do, the extent to which this is the case. To this end, we analysed a parallel corpus of transcribed source speeches and their transcribed simultaneous interpretations. Specifically, we retrieved occurrences of numbers in source-text segments and analyzed the aligned renditions.

Here we investigate the following research question: When numbers are produced in a source speech, to what extent is (dis)fluency in interpreting related to the following variables:

- omissions of numbers by the interpreter;
- language pair;
- nativeness of the original speaker to the language spoken;
- speed of delivery of the original speech;
- mode of delivery of the original speech (i.e., whether the speech was read out or delivered impromptu);
- type of number;
- frequency of numbers in the same sentence?

Based on the literature reviewed in Sections 2 and 3, we hypothesize that all these variables may to varying extents have an impact on the likelihood of a disfluent rendition of numbers during interpreting.

5. Data and method

5.1 EPTIC and transcription guidelines

The reported analyses were performed on the European Parliament Translation and Interpreting Corpus (EPTIC; Ferraresi et al. 2019), which is a multilingual, intermodal parallel corpus composed of speeches delivered at the European Parliament (EP), their official interpretations and written translations. The corpus was compiled based on the materials available on the official website of the European Parliament, including verbatim reports and videos of the speeches with multilingual audio tracks. All the language pairs in EPTIC feature English in combination with another EU official language.

For the present study, transcripts of the source speeches and their interpreted versions in EPTIC were used, drawing on a subset of the English<>French, English<>Italian and Polish>English subcorpora. Only speeches featuring at least one number were selected from EPTIC. Table 1 shows the details of the specific set of transcripts included in the analysis.

Information regarding the nativeness of the original speaker was derived from the EPTIC corpus metadata (see Ferraresi et al. 2019), which in turn are based on EP members' biographical profiles made available by the EP website. Speed of delivery and mode of delivery[2] are also included among the EPTIC metadata and are based on manual scrutiny of the audio and video files of the speeches.

---

[2] Although telling the difference between impromptu delivery and reading aloud from a written script is not necessarily straightforward, the corpus creators report that in most cases it was possible to distinguish between the two modes based on both auditory and visual cues (e.g. intonation and whether the speaker was seen to read from a sheet of paper or other support; see Ferraresi et al. 2019).

Table 1. Metadata on the EPTIC transcripts included in the analysis

	en > fr	en > it	fr > en	it > en	pl > en	TOTAL
Words	17,795	15,157	18,710	9,608	4,434	**65,704**
Interpreted speeches	46	46	50	33	24	**199**
Speakers of original speeches[3]	30	30	33	23	22	**138**
Nativeness of the original speaker	yes: 21 no: 25	yes: 21 no: 25	yes: 48 no: 2	yes: 33 no: 0	yes: 24 no: 0	**yes: 147 no: 52**
Mode of delivery	read: 26 mixed: 12 impromptu: 8	read: 26 mixed: 12 impromptu: 8	read: 25 mixed: 15 impromptu: 10	read: 25 mixed: 6 impromptu: 2	read: 17 mixed: 3 impromptu: 4	**read: 119 mixed: 48 impromptu: 32**

Transcripts of EP speeches and interpretations are time-aligned with the corresponding videos. Subcorpora of individual language pairs are aligned to each other at the sentence level. The transcription conventions adopted in EPTIC aim to include orality traits (such as truncated words, repetitions, disfluencies and pauses) based on the audio or video recordings of the plenary sessions at the EP (Ferraresi et al. 2019). To facilitate alignment, utterances in EPTIC were transformed into sentences and the transcripts feature standard punctuation to re-create the sentence structure of the speech based on syntax and a speaker's intonation: for the sake of brevity, throughout the article we refer to these sentence-like utterances simply as "sentences".

As a consequence of this decision, "natural" empty pauses occurring at possible sentence boundaries are not annotated. But empty pauses which sound unnatural and interrupt the discourse flow are indicated with "…" and tagged as EPAUSE. Filled pauses, tagged as FPAUSE, are indicated with a single "ehm", regardless of their duration or repetition. Oral subcorpora of EPTIC are also annotated for truncated words (i.e., if the speaker does not pronounce the entire word, this is indicated with a hyphen, as in "propo-"), which receive the "DYSF" tag.

In the reported statistical analysis, we collated all the filled pauses (FPAUSE), empty pauses (EPAUSE) and disfluencies (DYSF) into one common category of disfluencies.

5.2 Annotation

Relying on the annotation of the corpus,[4] we extracted all the numbers in the original speeches in all of the source languages considered (English, French, Italian, Polish) and their immediate co-text, together with the entire aligned segments in all of the target texts (French and Italian for original English speeches; and English for original speeches in French, Italian and Polish).

---

[3] Some speakers delivered more than one speech included in our corpora.
[4] For corpus documentation and the language-specific tagsets, see https://corpora.dipintra.it/eptic/

The co-text considered comprised the seven words preceding and the seven words following the number. This span, which is slightly wider than the four or five words on either side suggested by Sinclair (2003), was established empirically to ensure that sufficient context was available for the annotator to understand what a particular number refers to, while not being "so large that a great deal of extraneous material is also collected" (2003: 174).

We manually checked all instances of numbers extracted in this way to make sure that they were counted only once (e.g. to avoid counting "1 million" separately as "1" and "million") and corrected a few cases of misaligned segments. Moreover, we made sure that, if a number was missing in the aligned interpreted versions, it did not appear in either the preceding or the following segment. These cases were annotated as omissions.

Each extracted number was annotated according to several variables, using a combination of ad hoc scripts written in Java and manual analysis. Table 2 shows the dependent and independent variables which were modelled in the statistical analysis (see Section 5.3).

Table 2. Variables considered in the statistical analysis

Variable	**Variable type and possible values[5]**	**Annotation**
Dependent variable		
Disfluency	Binary: yes/no	Automatic: EPTIC
Independent variables		
Language pair	Categorical: en–fr/en–it/fr–en/it–en/pl–en	Automatic: EPTIC
Nativeness of the original speaker	Binary: native/non-native	Automatic: EPTIC
Speed of the original speech	Numeric: words per min	Automatic: EPTIC
Mode of delivery	Categorical: read, mixed, impromptu	Automatic: EPTIC
Omission	Binary: yes/no	Manual
Type of number	Categorical: small whole numbers/date/four or more digits/part of name/range/decimal	Manual
Frequency of numbers in a single sentence	Binary: one number/more than one number	Manual

---

[5] In the case of categorical variables, the first value is the intercept (see Section 5.3).

A disfluency in interpreted speeches is represented by a disfluency tag (see Section 5.1) in the segments aligned to source segments containing numbers. Disfluencies in the preceding or following segment were considered only if the number occurred in either the preceding or the following segment. If the number was omitted altogether, we counted the disfluencies in the segment where the number should have been included.

Furthermore, we annotated each instance of numbers in the original speech according to language pair, nativeness of the original speaker, and speed and mode of delivery, based on the metadata derived from EPTIC (see Section 5.1). We finally coded manually the type of number and the frequency of numbers in the same sentence. We split the annotation task among all the authors and had two meetings (one before the annotation task and one after the first round of annotations) in which we created guidelines for annotation and discussed unclear cases so as to maximize the reliability of the annotation. The annotation categories with examples and clarifications of potentially ambiguous cases are presented in Table 3.

Table 3. Annotation categories with examples and clarifications for potentially ambiguous cases

Type of number and clarification	Example
Small whole numbers	
Numbers below 1,000.	*In the last ehm 7 days, there's been a programme* […] *They are today 300.*
Percentages are also considered small whole numbers if they do not include decimals or if they are not part of a range expression (see below).	[…] *but Parliament managed to secure 75 percent of the amendments* […]
If the word *one* performs a text-organizing function, it is treated as a number, whereas if it is used as a pronoun, it is not.	[…] *you address one concrete case* […]
Dates	
Numbers that occur as part of a date expression. In the case of multiple numbers in the same date expression, each one is counted and treated separately (e.g., "22nd" constitutes one numerical expression and "2010" another one).	[…] *manufacturer notified the German competent authorities on 22nd December 2010* [...]
Numbers with four or more digits	
Values above 999.	[...] *which is the fact that so many 38 million people rely on the construction sector* [...]
Numbers as part of name	
Numbers that act as part of a name (e.g. of an article or legislation), including dates that are parts of an established name.	[...] *on which the success of ehm Europe 2020 Strategy is hinged* [...] [...] *that any use of new Article 61 bis of the Penal Code would be* [...]

Numbers expressing range	
Numbers expressing a range are considered as such irrespective of the other categories they represent. Each number of the range expression is counted and treated separately (e.g. "2002" constitutes one numerical expression and "2008" another).	*Between **2002** and **2008**, the EU contributed 544 million euro* […] *Each year, **10** to **15 billion** dollars are lost to revenues* […]
Decimals	
Decimal numbers and fractions, including percentages that contain decimals. In the case of split fractions e.g. *one person in three*, 'one' and 'three' are counted and treated separately, i.e. "1" constitutes one numerical expression and "3" another one.	[…] *whole of the Pacific region is just **0.06** percent, and yet some areas have* […] […] *and imported into the EU constitutes almost **one-fifth** of the timber products on our* […] […] *importing Scottish haggis, yet **1** in **3** Americans claim Scottish ehm ancestry* […]

All the data analyzed in this article are made available as an open access repository hosted on the Open Science Framework.[6]

5.3 Statistical model

We used a generalized linear mixed model to analyze our data. This entailed modelling disfluencies in the interpretation as a function of omission, language pair, nativeness of the original speaker, speed of the original speech, mode of delivery, type of number and frequency of numbers in the same sentence (see Table 2). As the speed measured in words can be language-specific (e.g., depending on the languages' tendency to form compounds vs. phrases), we included an interaction between speed and directionality. Speaker-specific random intercepts and random slopes were also included[7] to indicate the effect of type of number and frequency of numbers in the same sentence, that is, those variables that relate directly to the difficulties specific to numbers.

As our statistical analysis is confirmatory, all the variables were motivated by earlier research and we did not follow any model selection procedures (for the rationale, see e.g. Barr et al. 2013; Whittingham et al. 2006; for an accessible and linguistically oriented discussion, see also Winter 2020: 274–280). Furthermore, we did not test the statistical significance of the overall model. Instead, the statistical significance of each variable and interaction included is calculated by comparing the full model to one without the respective variable or interaction with a likelihood ratio test using ANOVA.

---

[6] https://osf.io/nf5pz/
[7] Most of the texts in EPTIC are not annotated with the interpreters' ID, so only original speakers could be taken into account in the structure of random effects.

In addition, we evaluated the model regarding its prediction accuracy – that is, the degree to which it could predict the presence or absence of disfluencies in interpretations when numbers were present in the original speech. Values for the prediction accuracy can range between 0 and 1, where 0.5 is the chance level baseline. Hence, values above 0.5 indicate the model's ability to capture the phenomenon over and above chance.

Finally, we evaluated the possible collinearity of different variables in terms of the variance inflation factors. For each predicting variable, this metric measures the degree to which its values can be accounted for by other predicting variables. The more collinearity there is, the more difficult it becomes to interpret the model. Possible values range from 1 upwards, with a commonly used threshold for acceptable collinearity below 10 (e.g. Gries & Deshors 2014).

All our statistical analyses were conducted in R (R Core Team 2022) using the *lme4* package for the model fitting (Bates et al. 2015), the *MuMIn* package for pseudo-R^2 (Barton 2019)[,] and the *car* package (Fox & Weisberg 2011) for calculating the variance inflation factors.

6. Results

The fitted model does not capture the variance in the data very well ($R^2_{marginal}$=0.19, $R^2_{conditional}$=0.42), but the prediction accuracy of 0.71 is still well above chance and all the variance inflation factors are well below 10. Therefore, we felt confident to proceed with the analysis. Parameter estimates and the results of the ANOVA comparisons for the fixed effects can be found in Table 4. The fixed variables that did not reach the level of statistical significance include source-text mode of delivery, nativeness of the original speaker, speed of delivery of the source speech and interpreting direction. These variables are therefore not discussed in greater detail here. In what follows, we discuss the fixed effects, with statistically significant effects and the random structure included in the model.

**Table 4.** Parameter estimates for the fixed effects

Variable	Variable levels	Estimate	Std error	Z value	Likelihood ratio test	Confidence Interval 2.5%	97.5%
Intercept	–	–4.091	1.922	–2.129		–7.86	–0.32
Omission	No	Intercept	–	–	X(1)=43.126	–	–
	Yes	1.395	0.227	6.157	$p<.0001$	0.95	1.84
Speed of the original speech	Numeric	0.020	0.012	1.724	X(5)=6.0321 $p=.3031$	0.00	0.04
Language pair	en–fr	Intercept	–	–	X(8)=10.394	–	–
	en–it	3.454	1.780	1.941	$p=.2385$	–0.03	6.94
	fr–en	4.603	2.433	1.892		–0.17	9.37
	it–en	0.842	2.942	0.286		–4.92	6.61
	pl–en	2.401	3.457	0.677		–4.55	9.35
Mode of delivery	read	Intercept	–	–	X(2)=0.3541	–	–
	mixed	–0.113	0.240	–0.469	$p=.8377$	–0.58	0.36
	impromptu	–0.182	0.377	–0.484		–0.92	0.56
Nativeness of the original speaker	yes	Intercept	–	–	X(1)=0.0038	–	–
	no	–0.015	0.378	–0.041	$p=.9511$	–0.76	0.72
Frequency of numbers in the same sentence	one	Intercept	–	–	X(1)=10.431	–	–
	more than one	0.734	0.217	3.378	$p=.0012$	0.31	1.16
Type of number	Small whole number	Intercept	–	–	X(5)=14.823 $p=.0111$	–	–
	Part of name	0.911	0.448	2.032		0.03	1.79
	Date	–0.080	0.228	–0.349		–0.53	0.37
	Range	–2.373	0.953	–2.488		–4.24	–0.50
	Decimal	0.725	0.553	1.311		–0.36	1.81
	Four or more digits	0.602	0.338	1.780		–0.06	1.26
Speed of the original speech: Language pair	en–fr	Intercept	–	–	X(4)=5.4768 $p=.2418$	–	–
	en–it	–0.021	0.012	–1.737		–0.04	0.00
	fr–en	–0.027	0.014	–1.847		–0.06	0.00
	it–en	–0.003	0.019	–0.133		–0.04	0.03
	pl–en	–0.009	0.026	–0.336		–0-06	0.04

Omission. The omissions of numbers in interpretations have a statistically significant effect on interpretation fluency. As can be seen in Figure 1, the probability of disfluencies increases drastically in contexts where numbers occurring in the source text have been omitted in the target text. It can be hypothesized that omissions and disfluencies might signal processing problems.

13
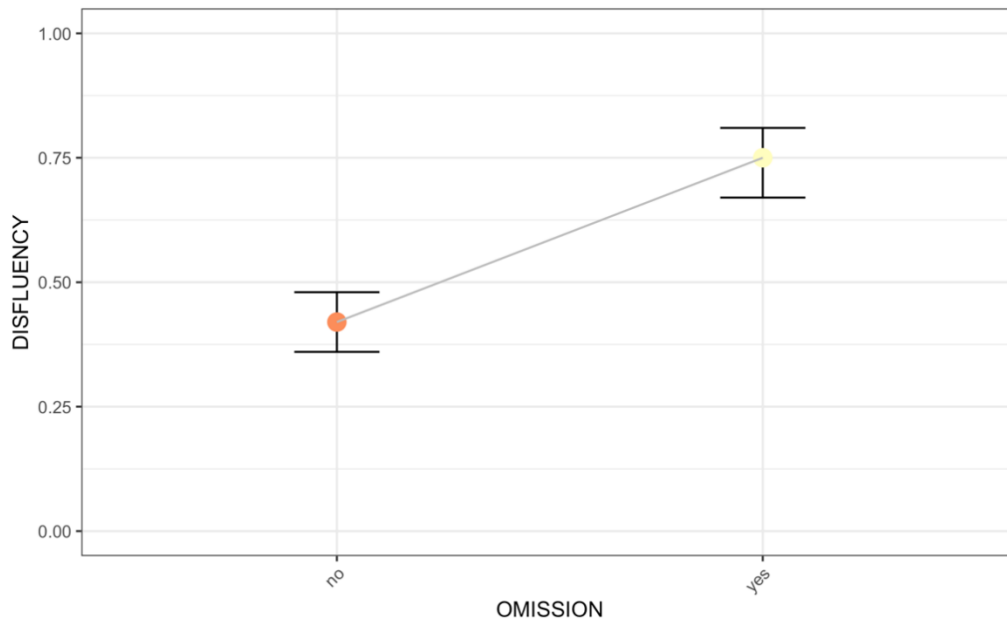
Figure 1. The effect[8] of omission on interpretation disfluencies

Type of number. The type of number, too, has a statistically significant effect on interpretation disfluencies. As Table 4 and Figure 2 show, numbers expressing a range are clearly less associated with disfluencies than any other types of number. Another noteworthy tendency is also observed for numbers used as part of a name: when interpreting numbers that are a part of a name, interpreters are likely to be less fluent.
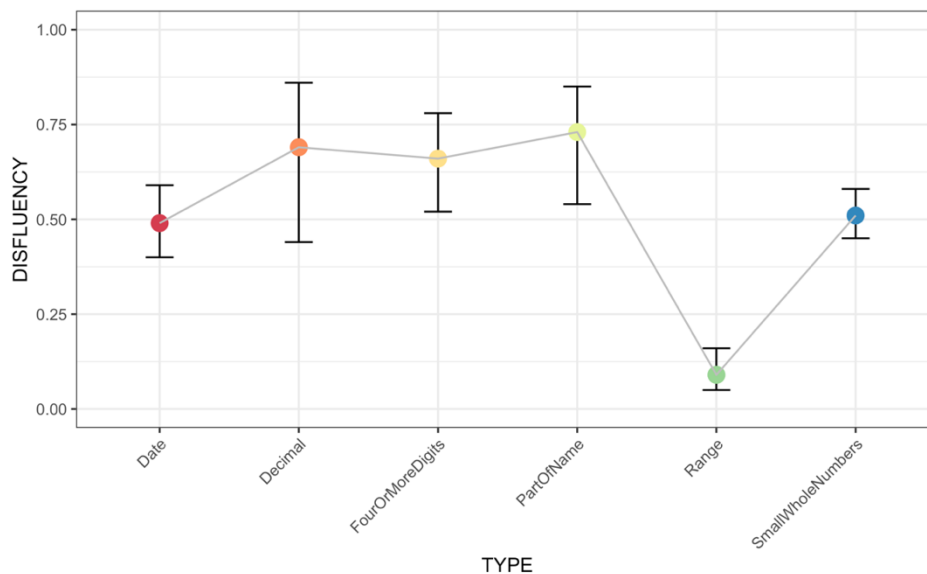
Figure 2. The effect of the type of number on interpretation disfluencies

---

[8] The term "effect" is commonly used when describing the predicting the role of variables in regression analyses. It is important to note, however, that the term does not necessarily imply a causal relationship.

Frequency of numbers. The frequency of numbers in the same sentence has a statistically significant effect on interpretation fluency. As indicated by the estimates in Table 4 and Figure 3, while interpreting utterances with more than one number, interpreters are more likely to produce disfluencies.
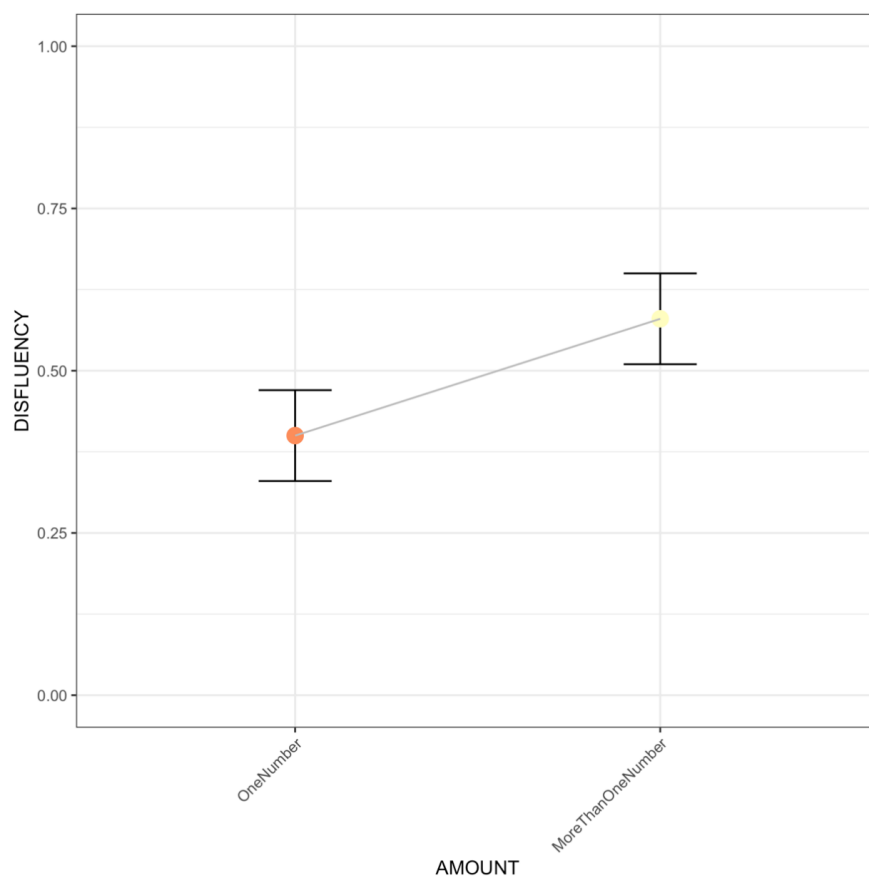
Figure 3. The effect of the frequency of numbers per sentence on interpretation disfluencies

Speaker. As mentioned in Section 5.3, we specified random intercepts to account for speaker-specific behavior regarding numbers, and also by-speaker random slopes for the type of number and the frequency of numbers per sentence. None of these yields a statistically significant effect on the interpretation fluency in contexts where numbers are present in the source text. Interestingly, however, the random structure considerably affects the variance explained by the model. Most notably, leaving out the random slopes for the type of number makes the $R^2_{conditional}$ drop to 0.29 (full model = 0.42), effectively meaning that there is indeed a considerable by-speaker variation in the way the type of number affects interpretation fluency. A closer look at the variance and the standard deviations of the random effects (see Appendix)

reveals that this is chiefly due to numbers expressing ranges: the by-speaker variance related to interpretation fluency is more than 10 times higher in ranges than in any other type of number, the frequency of numbers or the intercept. In other words, whereas ranges are generally interpreted more fluently than other types of number, there is a wide variance between different speakers. Note, however, that the different types of number are often also highly correlated with different speakers and that the observed by-speaker differences are therefore in some cases closely related to the overall occurrences of the different types of number across speakers rather than to differences in the disfluencies related to them.

7. Discussion

Summing up the results presented in section 6, we found generally that some factors which significantly increase the likelihood of disfluencies in interpreters' output are associated with the type of number and the frequency of numbers observed in the sentence-like source utterances. Specifically, the presence of more than one number in a source utterance causes a significant increase in interpreters' disfluencies. Such a result is consistent with the idea that a series of numbers might place an additional burden on working memory (Jones 2014).

Numbers that are part of names (e.g. "Article 18 [of the Declaration of Human Rights]") were also associated with more disfluencies, whereas the opposite trend was observed with numbers that are part of ranges (e.g. "between 100,000 and 200,000 tonnes of feed"). We hypothesize that both effects are related to the level of predictability of the two types of number and therefore to the extent to which interpreters can rely on anticipation and reformulation strategies when rendering them. Like numbers, names are characterized by both low predictability and high informational content and are widely acknowledged as problem triggers in simultaneous interpreting. A combination of names and numbers is therefore likely to have an additive effect on the cognitive effort required of interpreters in the processing and production phase, hence the higher rate of disfluencies in their rendition. In contrast, the lower rate of disfluencies associated with range expressions could be related to their greater predictability. Even though we did not control for the difference between first and subsequent numbers in expressions of a range, it could be hypothesized that the number initiating the range leads interpreters to switch to a 'literal' type of hearing (Pinochi 2010; Seleskovitch 1975) which facilitates the processing of successive numbers. We can only speculate that this might also explain why numbers included in ranges do not follow the trend pertaining to multiple

unrelated numbers occurring in the same source sentence, which results in an increased likelihood of disfluencies. As one of the reviewers of this article pointed out, it is also possible that since ranges repeat the same category of numbers, which in many cases also have the same order of magnitude, they are more redundant than isolated occurrences of numbers and therefore offer a higher potential for context-dependent reconstruction.

Finally, omission was found to be negatively correlated with interpreting fluency: the more numbers interpreters omit, the more likely they are to also pause or hesitate in their output. Plevoets and Defrancq (2018: 23) suggested that omission might play a role "in the reduction of cognitive load, particularly in the case of numbers". Owing to the design of that study, which did not involve item-by-item scrutiny of source-text numbers and their renditions in target texts, we could only hypothesize that "some source text numbers which were not interpreted, nevertheless generated disfluencies" in the interpreted output (Plevoets & Defrancq 2018: 23). The evidence presented in this study lends support to such a hypothesis, confirming that omissions might be a coping strategy adopted by interpreters when numbers in the source text add to the cognitive effort required in their task.

8. Conclusion

This article presented a corpus-based investigation of the effect of numbers in source texts on the fluency of interpretations. Unlike previous studies, our dataset is based on interpreters' performances in authentic professional settings and involves several combinations of source and target languages, including English, French, Italian and Polish. We operationalized fluency as the absence of empty and filled pauses and hesitations and considered several theoretically motivated factors that could affect interpretation fluency when a number is uttered in source speeches.

A statistical analysis based on a generalized linear mixed model revealed that only some of the factors considered have a significant effect on fluency. Numbers that are uttered as part of range lead to fewer disfluencies than any other type of number, whereas the presence of several numbers in source sentences and/or the use of numbers as part of proper names increases the likelihood of disfluencies. We interpreted these results as being related to the degree of predictability and redundancy of numbers, which may increase or decrease the burden imposed on interpreters' working memory and therefore the (dis)fluency of their renditions. A significant effect was also observed with omissions: the more numbers are omitted, the more disfluencies are observed in interpreted speeches, possibly pointing to the role played by

omissions in reducing cognitive load. The remaining factors which were taken into account did not yield statistically significant results. These included whether the source text was read or delivered impromptu, the speed of the speech, whether it was uttered by a native or a non-native speaker, and the language combinations involved.

In interpreting these results, one should not exclude the possibility that data scarcity might play a crucial role. Compared to other more widely studied lexical or syntactic features, numbers are a rather rare phenomenon, with occurrences in the range of 6.6 to 8.6 per thousand words in the sub-corpora of the source texts at our disposal. One might consider drawing on other types of speeches (e.g. at financial or demographic institutions), where numbers are likely to be uttered more frequently.

Whether one opts for differentiating text sources or enlarging existing resources, we believe that further corpus-based explorations of the ways in which numbers influence interpreters' cognitive effort are worth pursuing. Not only do such studies contribute to a deeper understanding of the challenges that underlie simultaneous interpreting; they also translate into practical insights which might be useful to interpreters and interpreter trainers who need to direct their attention to these problematic areas.

Funding information

References

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language 68* (3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barton, K. (2019). *MuMIn: Multi-model Inference.* https://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67* (1), 1–48. https://doi.org/10.18637/jss.v067.i01

Braun, S. & Clarici, A. (1996). Inaccuracy for numerals in simultaneous interpretation: Neurolinguistic and neuropsychological perspectives. *The Interpreters' Newsletter 7*, 85–102.

Cecot, M. (2001). Pauses in simultaneous interpretation: A contrastive analysis of professional interpreters' performances. *The Interpreters' Newsletter 11*, 63–85.

Chmiel, A., Szarkowska, A., Koržinek, D., Lijewska, A., Dutka, Ł., Brocki, Ł. & Marasek, K. (2017). Ear–voice span and pauses in intra- and interlingual respeaking: An exploratory study into temporal aspects of the respeaking process. *Applied Psycholinguistics 38* (5), 1201–1227. https://doi.org/10.1017/S0142716417000108

Collard, C. & Defrancq, B. (2020). Disfluencies in simultaneous interpreting, a corpus-based study with special reference to sex. In J. Daems, B. Defrancq & L. Vandevoorde, (Eds.), *New empirical perspectives on translation and interpreting*. Abingdon: Routledge, 264–299. https://doi.org/10.4324/9780429030376-12

Defrancq, B. & Plevoets, K. (2018). Over-uh-load, filled pauses in compounds as a signal of cognitive load. In M. Russo, C. Bendazzoli & B. Defrancq (Eds.), *Making way in corpus-based interpreting studies*. Singapore: Springer, 43–64.

Desmet, B., Vandierendonck, M. & Defrancq, B. (2018). Simultaneous interpretation of numbers and the impact of technological support. In C. Fantinuoli (Ed.), *Interpreting and technology*. Berlin: Language Science Press, 13–27. http://dx.doi.org/10.5281/zenodo.1493281

Fehringer, C. & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica 41* (1–2), 37–72.

Ferraresi, A., Bernardini, S., Milicevic Petrovic, M. & Lefer, M.-A. (2019). 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

Fox, J. & Weisberg, S. (2011). *An R companion to applied regression.* (Second edition). Thousand Oaks, CA: Sage.

Frittella, F. M. (2019). "70.6 Billion World Citizens": Investigating the difficulty of interpreting numbers. *Translation and Interpreting 11* (1)*,* 79–99.

Frittella, F. M. (2022). CAI tool-supported SI of numbers: A theoretical and methodological contribution. *International Journal of Interpreter Education 14* (1), Article 5. https://doi.org/10.34068/ijie.14.01.05

Gósy, M. (2007). Disfluencies and self-monitoring. *Govor 24* (2), 91–110.

Gries, S. T. & Deshors, S. C. (2014). Using regressions to explore deviations between corpus data and a standard/target: Two suggestions. *Corpora 9* (1), 109–136.

Jones, R. (2014). *Conference interpreting explained*. London/New York: Routledge.

Kajzer-Wietrzny, M., Ivaska, I. & Ferraresi, A. (2021). 'Lost' in interpreting and 'found' in translation: Using an intermodal, multidirectional parallel corpus to investigate the rendition of numbers. *Perspectives 29* (4), 469–488. https://doi.org/10.1080/0907676X.2020.1860097

Korpal, P. (2016). *Linguistic and psychological indicators of stress in simultaneous interpreting*. PhD dissertation, Adam Mickiewicz University.

Korpal, P. & Stachowiak-Szymczak, K. (2018). The whole picture: Processing of numbers and their context in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics 54* (3), 335–354.

Maclay, H. & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word 15* (1), 19–44.

Mazza, C. (2001). Numbers in simultaneous interpretation. *The Interpreters' Newsletter 11*, 87–104.

Pellatt, V. (2006). The trouble with numbers: How linguistic, arithmetical and contextual complexity affect the interpretation of numbers. In M. Chai & J. Zhang (Eds.), *Professionalization in interpreting: International experience and development in China.* Shanghai: Foreign Language Education Press, 350–365.

Piccaluga, M., Nespoulous, J.-L. & Harmegnies, B. (2005). Disfluencies as a window on cognitive processing. An analysis of silent pauses in simultaneous interpreting. Paper presented at DiSS'05, Disfluency in Spontaneous Speech Workshop, Aix-en-Provence, France.

Pinochi, D. (2010). Simultaneous interpretation of numbers: Comparing German and English to Italian. An experimental study. *The Interpreters' Newsletter 14*, 33–57.

Plevoets, K. & Defrancq, B. (2016). The effect of informational load on disfluencies in interpreting. *Translation and Interpreting Studies 11* (2), 202–224.

Plevoets, K. & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament. *Interpreting 20* (1), 1–28.

Pöchhacker, F. (1995). Slips and shifts in simultaneous interpreting. In J. Tommola (Ed.), *Topics in interpreting research*. Turku: University of Turku, Centre for Translation and Interpreting, 73–90.

R Core Team (2022). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Seleskovitch, D. (1975). *Langage, langues et mémoire: Étude de la prise de notes en interprétation consécutive*. Paris: Lettres modernes.

Shlesinger, M. (1994). Intonation in the production of and perception of simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation*. Amsterdam: John Benjamins, 225–236.

Sinclair, J. (2003). *Reading concordances: An introduction*. London: Pearson Education.

Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter 10*, 103–127.

Whittingham, M. J., Stephens, P. A., Bradbury, R. B. & Freckleton, R. P. (2006). Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology 75* (5), 1182–1189. https://doi.org/10.1111/j.1365-2656.2006.01141.x

Winter, B. (2020). *Statistics for linguists: An introduction using R*. New York: Routledge.

Appendix

Model summary

```
 Family: binomial  ( logit )
Formula: DISFLUENCY ~ OMISSION + ST_SPEED * DATA + ST_DELIVERY + ST_NATIVENESS +
    AMOUNT + TYPE + (1 + AMOUNT + TYPE | SPEAKER)
   Data: dat_for_fluency_plots
Control: glmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  1015.8   1238.5   -459.9    919.8      716

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.40330 -0.74227 -0.09857  0.75481  2.77790

Random effects:
 Groups   Name                     Variance Std.Dev. Corr
 SPEAKER  (Intercept)              0.3118   0.5583
          AMOUNTMoreThanOneNumber  0.6489   0.8056   -0.43
          TYPEPartOfName           0.3882   0.6231   -0.74  0.92
          TYPEDate                 0.1177   0.3431    0.19  0.80  0.52
          TYPERange                8.6491   2.9409    0.99 -0.27 -0.62  0.35
          TYPEDecimal              0.1349   0.3673    0.27 -0.99 -0.85 -0.89  0.11
          TYPEFourOrMoreDigits     0.1119   0.3345    0.99 -0.57 -0.84  0.03  0.95  0.42
Number of obs: 764, groups:  SPEAKER, 108

Fixed effects:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)              -4.091237   1.922013  -2.129 0.033286 *
OMISSIONyes               1.395286   0.226611   6.157 7.4e-10 ***
ST_SPEED                  0.020446   0.011860   1.724 0.084732 .
DATAen_to_it              3.454360   1.779708   1.941 0.052262 .
DATAfr_to_en              4.603045   2.433152   1.892 0.058517 .
DATAit_to_en              0.841779   2.941839   0.286 0.774771
DATApl_to_en              2.401227   3.546662   0.677 0.498381
ST_DELIVERYmixed         -0.112882   0.240473  -0.469 0.638771
ST_DELIVERYimpromptu     -0.182405   0.376985  -0.484 0.628492
ST_NATIVENESSn           -0.015326   0.377691  -0.041 0.967633
AMOUNTMoreThanOneNumber   0.734412   0.217381   3.378 0.000729 ***
TYPEPartOfName            0.910637   0.448187   2.032 0.042171 *
TYPEDate                 -0.079545   0.227987  -0.349 0.727164
TYPERange                -2.373225   0.953792  -2.488 0.012839 *
TYPEDecimal               0.724747   0.552762   1.311 0.189811
TYPEFourOrMoreDigits      0.602004   0.338136   1.780 0.075017 .
ST_SPEED:DATAen_to_it    -0.020560   0.011835  -1.737 0.082351 .
ST_SPEED:DATAfr_to_en    -0.026692   0.014453  -1.847 0.064767 .
ST_SPEED:DATAit_to_en    -0.002526   0.019037  -0.133 0.894460
ST_SPEED:DATApl_to_en    -0.008854   0.026370  -0.336 0.737037
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
      R2m       R2c
0.1876567 0.4242952
```

Address for correspondence
Marta Kajzer-Wietrzny
Faculty of English
Adam Mickiewicz University
Grunwaldzka 6
60-780 Poznań
Poland

kajzer@amu.edu.pl

Biographical notes

Marta Kajzer-Wietrzny is an assistant professor in the Department of Translation Studies at the Faculty of English, Adam Mickiewicz University in Poznań. Following her PhD dissertation on *Interpreting universals and interpreting style* (2012) she continued with empirical investigations of interpreted, translated and non-native language use, e.g. within the recent TRINFO project carried out in part during an extended research stay at the University of Bologna.
ORCID: 0000-0002-4441-979X

Ilmari Ivaska is an assistant professor in the Department of Finnish and Finno-Ugric Languages of the University of Turku. After getting his doctorate in Turku, Ivaska has worked as post-doctoral research fellow at the University of Bologna and visiting lecturer at the University of Washington. Ivaska is an applied linguist with research interests revolving around corpus methodologies as well as quantitative contrastive and crosslinguistic research designs, especially in the context of non-native and mediated language varieties.
ORCID: 0000-0001-7366-7111

Adriano Ferraresi is Associate Professor in the Department of Translation and Interpreting of the University of Bologna, where he is also a member of the Board of the Doctorate in Translation, Interpreting and Intercultural Studies. He holds a doctorate in English for Special Purposes from the University of Naples "Federico II". His main research interests lie in the field of construction of linguistic resources and in corpus-based investigations of phraseology and terminology in native, translated and lingua-franca varieties of English.
ORCID: 0000-0002-6957-0605