The International Journal for Translation & Interpreting Research

trans-int.org

Disfluencies and directionality in simultaneous interpreting: A corpus study comparing into-B and into-A interpretations from the European Parliament

Magdalena Bartłomiejczyk
University of Silesia in Katowice/University of Vienna
magdalena.bartlomiejczyk@us.edu.pl

Ewa Gumul
University of Silesia in Katowice/University of Las Palmas de Gran Canaria
ewa.gumul@us.edu.pl

DOI: 10.12807/ti.116201.2024.a03

Abstract: Although directionality in interpreting has attracted considerable scholarly interest, to date, there are no observational studies providing empirical evidence on the directionality effect in the performance of professional interpreters who regularly work in both directions. Our aim is to partially fill this gap by exploring the relationship between disfluencies and directionality in simultaneous interpreting. To this end, we compare A-B and B/C-A interpretations from plenary debates of the European Parliament. Using English and Polish subcorpora extracted from the EP-Poland corpus, we performed quantitative analyses focusing on three types of disfluencies: anomalous pauses, hesitation markers, and false starts. Our initial assumption was that, in line with the prevalent belief, cognitive load is lower for into-A interpreting. Accordingly, we hypothesised that into-B interpretations should exhibit significantly more disfluencies. However, we have found no directionality effects related to disfluencies. Therefore, our results do not support the advantage of into-A interpreting. This finding raises some doubts about the legitimacy of favouring interpreting into the native language to the extent it is done now in many institutions.

Keywords: directionality; disfluencies; pauses; simultaneous interpreting; European Parliament (EP)

1. Introduction

Simultaneous interpreters have to manage and coordinate concurrent source language comprehension and target language production, while often unable to predict the speaker's ultimate communicative intent. This produces significant cognitive load, which, however, normally fluctuates across an interpreting task. In our study, we will use the construct of cognitive load to refer to the "portion of an interpreter's limited cognitive capacity devoted to performing an interpreting task in a certain environment" (Chen, 2017, p. 643). Research has revealed symptoms of increased cognitive load, detectable either in the interpretation or in the interpreter's physical reactions (Chen, 2017). The former include disfluencies such as filled and silent pauses (e.g., Plevoets & Defrancq, 2018; Collard & Defrancq, 2020; Gumul, 2021). At the same time, disfluencies can also be perceived as presentation errors (e.g., Kurz & Färber, 2003; Bartłomiejczyk, 2010), i.e., symptoms of deteriorating interpreting quality.

International organizations generally favour interpreting into A (the native language) due to the widespread belief in the superiority of this direction (see, e.g., Graves et al., 2022). It tends to be seen as both easier for the interpreter and conducive to better quality (e.g., Donovan, 2004; Chmiel, 2016; Gumul, 2021). However, into-B interpreting often fulfils a genuine market need, and this is definitely the case for interpreting from languages of low diffusion, understood as ones "rarely learned by non-native speakers" (Whyatt & Pavlovič, 2021, p. 143), including Polish.

Since the beginning of this century, a plethora of empirical studies (e.g., Al-Salman & Al-Khanji, 2002; Kurz & Färber, 2003; Bartłomiejczyk, 2006; Chang & Schallert, 2007) have shown directionality effects for various language combinations, typically including English. In this vein, we would like to explore disfluencies in A-B vs. B/C-A interpreting by comparing output in each of the interpreting directions by European Parliament (EP) interpreters working between Polish and English. Native speakers of English only interpret from Polish into English, while some native speakers of Polish interpret both ways. Our quantitative analysis focuses on three types of disfluencies: anomalous pauses, hesitation markers, and false starts. For the formulation of our initial hypotheses, we follow the prevalent, though controversial, belief that cognitive load is lower for interpreting into A:

- Hypothesis 1: Polish-English interpretations by Poles (A-B) would exhibit significantly more disfluencies of various types than both English-Polish interpretations by Poles (B/C-A) and Polish-English interpretations by native speakers of English (C-A)
- Hypothesis 2: Interpretations into the native language, be it Polish or English, would feature comparable prevalence of disfluencies
- Hypothesis 3: Individual bidirectional interpreters would produce more disfluencies when working into their B than when working into A.

Our project has considerable practical importance for the setting under investigation, i.e., EU institutions. It is meant to provide hard data on one aspect of directionality that, among others, may guide the general human resources policies of the EU interpreting services and influence their outlook on into-B interpreting. Excessive disfluencies lower the quality of the interpreting product, particularly from the perspective of the audience (e.g., Shlesinger, 1994; Pradas Macías, 2006). If into-B interpreting is shown to produce target texts of markedly inferior quality, the need for it might be substantially reduced over time through appropriate recruitment and professional development policies. If, however, the existing concerns over the quality of into-B interpreting are found to be ungrounded, there may be no need to foster into-A interpreting to the extent it is encouraged today.

2. Directionality

Directionality in translation and interpreting is usually defined as working either into or out of one's native language, L1 or A (e.g., Tomczak & Whyatt, 2022) and this is the sense we refer to in this paper.[1] The directionality debate in interpreting studies (see, e.g., Gile, 2005; Bartłomiejczyk, 2015) relates mainly

[1] Rather confusingly, the same term is sometimes also used to contrast interpreting in either direction between two specific languages irrespective of their status as the interpreter's native or foreign language (e.g., Monti et al., 2005; Dayter, 2020).

to simultaneous interpreting. Over several decades, the debate has generated so many publications that their overview here has to be highly selective. At the same time, some important aspects of the directionality issue, such as the reception of authentic, professional into-B interpretations, have been largely overlooked, and hardly any conclusions have been accepted as universally convincing. Perhaps the only uncontested point is that the processes of interpreting into one's native language and vice versa somehow differ.

The dominant views on the superiority of into-A interpreting held by the "Western" side considerably influenced the practices at international organizations. However, these views (and opposing voices, too) were initially mostly based on personal convictions of trainers and theoreticians and, later on, on experiments which often departed very far from real field conditions; e.g., subjects performing "mental interpreting" without articulating the target text (Kurz, 1994) or providing target language counterparts of single words (De Bot, 2000). To the best of our knowledge, evidence from observation of authentic interpreting by professionals who regularly work in both the directions is non-existent. As directionality effects are sometimes difficult to disentangle from those produced by systemic differences between specific source and target languages (Gile, 2005), various language combinations would need to be examined to reach reliable overall conclusions.

The most negative views on simultaneous interpreting into B were held by the "Paris School" (e.g., Seleskovitch, 1968/1978). According to the authors associated with this school, into-B interpreting always suffered from extreme interference from the source language and was therefore intolerable. Although this position had slightly softened over the years to allow for into-B interpreting when no into-A interpreters can be employed (e.g., Seleskovitch & Lederer, 1989/1995; Donovan, 2004), it was still treated as a necessary evil, "like flying a plane with engines that are not only too small to start with but also throttled" (Dejean le Féal, 2003, p. 69). This attitude is also reflected in the popular term *retour*, which implies that into-A interpreting is the default option. By contrast, into-B interpreting was regarded highly and fostered in the Soviet Union and its sphere of influence. The proponents (e.g., Denissenko, 1989) emphasised the comprehension advantage in the interpreter's native language and minimized the weight of production problems such as accentedness or limited idiomacity.

As neatly outlined by Gile (2005), cognitive load is distributed differently depending on directionality, which can be explained in terms of his Effort Model of simultaneous interpreting. Presumably, the Listening and Analysis Effort will be lower for interpreters working from their A language, while the Production Effort will be lower for interpreters working into their A language. The directionality issue cannot be resolved once and for all, as it is not possible to determine which of the two Efforts is universally higher. Gile's observations have been instrumental in gaining at least limited acceptance for A-B interpreting even among its most vehement opponents, as he successfully challenged the Paris School's initial view that source language comprehension was practically effortless (Seleskovitch, 1968/1978).

The tendency in more recent studies is to explore selected directionality effects for a given language pair, e.g., English and Polish (Bartłomiejczyk, 2006 on strategic processing; Gumul, 2017a and 2017b on explicitation) while withholding categorical judgement on the relative superiority of one interpreting direction. Many authors (e.g., Wu & Liao, 2018) emphasise that directionality effects should have a bearing on interpreting training. The view that into-B interpreting is cognitively more demanding dominates nowadays (e.g., Chmiel, 2016; Gumul, 2021), but it is far from clear if and how this translates into product quality.

Interestingly, ample as it is, research dealing with directionality has rarely been concerned with any quality parameters of into-B vs. into-A interpretations, even though quality was always the major argument for rejecting into-B interpreting by the Paris School. Instead, postulated superiority of one of the directions has usually been extrapolated from interpreters' declared preferences (e.g., Donovan, 2004; Nicodemus & Emmoray, 2013) or proven cognitive advantages for the native language such as faster lexical access (e.g., Chmiel, 2016), higher anticipation potential (e.g., Kurz & Färber, 2003) or better memory of input (Opdenhoff, 2012). Arguably important quality parameters that have come under scrutiny in some experimental studies are completeness and accuracy (operationalized as propositional accuracy scores and/or error counts), with conflicting evidence emerging. Chang & Schallert (2007), for instance, report higher scores for B-A interpreting (from English into Chinese), whereas van Dijk et al. (2011) report higher scores for A-B interpreting (from Dutch into Sign Language of the Netherlands).

The European Union (EU) and the United Nations Organization (UNO) both favour interpreting into A, but still use A-B interpreting quite extensively. The UNO makes a clear distinction between booths that interpret only into their native language (English, French, Spanish and Russian), manned by two interpreters each; and booths that also interpret from their native language (Chinese and Arabic) into English and/or French, manned by three interpreters each (see, e.g., Ruiz Rosendo & Diur, 2022). The EU initially tried to rely exclusively on into-A interpreting, but found it increasingly difficult to find interpreters for some language combinations as the organisation became increasingly multilingual (see, e.g., Bartłomiejczyk & Stachowiak-Szymczak, 2022). The first major challenge came with the addition of Finnish in 1995, and an even greater one in 2004, when the biggest EU enlargement ever introduced nine new official languages. While no booth has been declared bidirectional for all purposes and occasions, into-B interpreting (predominantly into English) is used commonly for most present source languages (e.g., Polish, Hungarian, Estonian); and into-B interpretations serve as a basis for relay (indirect interpreting with the mediation of a 'pivot' language) for most of the other booths. At the same time, the EU very actively encourages its interpreters to acquire new C languages to avoid *retour* and relay. Consequently, some senior EU interpreters have as many as seven or eight C languages. This policy entails huge costs and time investments (intensive language courses, long stays abroad financed by the EU).

As bidirectional interpreters work for both the international institutions, the EU and UNO appear to be obvious sources of material for observational research on directionality. However, much to our surprise, we have not been able to identify any published studies that compare authentic A-B and B/C-A interpreting performed either in those settings or elsewhere. We are only aware of one on-going project (Chmiel et al., 2022) that intends to do this for the language pair English-Polish on the basis of PINC, a large bidirectional corpus compiled from EP material.

3. Disfluencies in interpreting

Scholarly interest in disfluencies in interpreting (also collectively referred to as flaws in delivery or non-fluencies) has not been nearly as widespread as that in directionality, still, the accumulated body of research seems more solid. To the best of our knowledge, no authors to date have investigated possible correlations of the two factors in simultaneous interpreting professionals (but see Mead 2000 for the consecutive mode and Lin et al. 2018 for simultaneous

interpreting of trainees). Various authors either focus on one disfluency type (e.g., Plevoets & Defrancq, 2018 on hesitation markers) or combine a wider array of these in a single study (e.g., Dayter, 2021), however, some confusion may arise from differences in nomenclature and definitions. In contrast to directionality, disfluencies have tended to be examined both under lab conditions (e.g., Wang & Li, 2015; Lin et al., 2018; Gumul, 2021) and in authentic material, including EP corpora such as EPIC (Bendazzoli et al., 2011), EPIC Ghent (e.g., Plevoets & Defrancq, 2016) and PINC (Chmiel et al., 2022). In this section, we would like to briefly discuss the existing literature dealing with disfluencies in simultaneous interpreting, focusing in particular on cognitive studies exploring the three types of disfluencies under analysis here.[2]

Disfluencies are one of the most prominent features that distinguish oral from written discourse. Spontaneous oral production, regardless of the register, the topic, the context or the speaker is bound to be marked by one type of disfluency or other (see e.g., Cecot, 2001). However, in interpreting, such interruptions in the smooth flow of speech might display different properties given the cognitive complexity of the task, as demonstrated by a number of studies on interpreting (e.g., Tissi, 2000; Hale, 2004; Chmiel et al., 2017; Gieshoff, 2021, etc.). Apart from the inherent difficulty of the interpreting task which lies in the complex multitasking, there are also certain text features that may aggravate fluency loss as they increase the cognitive load. Plevoets & Defrancq's (2016) results show that these are fast delivery rate, high lexical density, and the presence of numbers and long sentences. As emphasised by Plevoets & Defrancq (2016), the rise in disfluencies in interpreting should not only be associated with source text features, but also related to how interpreters manage their outputs and how their own production contributes to the cognitive load they experience. The causes of disfluencies as well as their frequency and patterns depend on their types. Silent pauses, filled pauses/hesitation markers, and false starts have been found to have different properties and to be conditioned by different factors.

Silent pauses, also referred to in literature as unfilled pauses (e.g., Han & An, 2021), are the breaches in fluency where no voice activity is recorded (in contrast to filled pauses) and the flow of speech is interrupted (Gieshoff, 2021). As observed by Igras-Cybulska et al. (2016), pausing behaviour in speech is highly idiosyncratic. The authors also note its dependence on the situational context and the cognitive complexity of the task. Silent pauses assume different functions in oral discourse depending on their length (see e.g., Bortfeld et al., 2001; Cecot, 2001). This property is also related to their perception by receivers. Whereas short pauses facilitate communication by marking speech structure and signalling syntactic boundaries in a sentence, longer pauses are associated with information loss (see e.g., Tissi, 2000; Pradas Macías, 2006; Rennert, 2010). Pradas Macías (2006) shows that pauses longer than 2 seconds are construed by the audience as omissions even if the rendition is complete. Therefore, this threshold value of 2 seconds is often adopted as defining anomalous pauses in studies investigating fluency in interpreting (e.g., Chmiel et al., 2017; Collard & Defrancq, 2019), while other studies take higher thresholds of 2.5 seconds (e.g., Cecot, 2001; Chmiel et al., 2022) or 3 seconds (e.g., Wang & Li, 2014).

Research on silent pauses in interpreting reveals a number of important features of interpreters' pausing behaviour. First of all, Tissi (2000) observed large individual variations. Her results also indicate that unfilled pauses tend to be longer in interpreted speech, while Cecot (2001) has noted their irregular

[2] Other, less widely researched disfluencies include mispronounced words (e.g., Bendazzoli et al. 2011), repetitions (e.g., Dayter, 2021) and intra-word prolonged vowels (e.g., Defrancq & Plevoets, 2018).

distribution within outputs of the same individuals. Studies on silent pauses in interpreting also reveal that interpreted discourse has fewer such pauses than non-interpreted speech. This tendency has been confirmed for natural speech (Tissi, 2000) and respeaking (Chmiel et al., 2017). Being aware of the quality criteria and clients' expectations, interpreters tend to avoid pauses by explicitating implied meanings or adding neutral phrases (Defrancq et al., 2015; Gumul, 2017b). Silent pauses in simultaneous interpreting may also be related to text processing strategies (Ahrens, 2005). For instance, prolonging EVS while waiting for more input may result in a long silent pause in some cases.

Another type of disfluency characteristic of oral discourse are hesitation markers, also referred to in literature as 'filled pauses' (Plevoets & Defrancq, 2016) or 'fillers' (Bortfeld et al., 2001). They are described as vocalized hesitations and non-lexical fillers in the form of meaningless strings of prolonged sounds (Gumul, 2021). The form of such items is language-dependent, for instance, in Polish they are typically prolonged vowels *yyy*, *eee* or the consonant *mmm* (Igras-Cybulska et al., 2016). Cecot's examples for Italian include: *eeh, mhm, mah, beh,* and *bah*, while Hale lists *eh*, *ah*, *em* and *am* for Spanish. In this study we shall use the term hesitation markers, as the term filled pauses appears to be broader and includes also inter-word vowel lengthening, repetitions and corrections (Gieshoff, 2021). Hesitation markers are believed to be induced by cognitive load and/or speech planning effort (Hale, 2004; Plevoets & Defrancq, 2016). Their use is also highly idiosyncratic and reflects one's individual way of speaking (Cecot, 2001). Research on hesitation markers in interpreting reveals higher frequency of such disfluencies in interpreted discourse than in non-interpreted speeches (Plevoets & Defrancq, 2016) or source speeches (Hale, 2004). Plevoets & Defrancq's (2016) corpus-based study provides evidence that denser distribution of this type of hesitations is mainly caused by two factors: the production effort imposed by lexical density of the target texts and the delivery rate of the source texts. Their later study (Plevoets & Defrancq, 2018) confirms that lexical density triggers more frequent hesitations, whereas source-text and target-text formulaicity helps interpreters to avoid this type of disfluency. Lin et al. (2018) found the correlation between directionality and the occurrence of hesitations. Their results show that students' interpreting outputs exhibit fewer hesitations when they interpret into their A language.

The last type of disfluency we analyse in our study are false starts. Typically, these are cases of word truncation. An interpreter either produces part of a word (usually the first syllable) and then the whole word (e.g., 'ba-baggage') or starts a word and then decides to articulate another one (e.g., 'ba-luggage') (Gilquin, 2008). False starts may also concern more extensive text segments. At the sentential level that would mean interrupting a sentence and venturing on a new one without completing or correcting the previous one (Rennert, 2010). False starts, especially the word-level ones, signal problems with lexical search during speech production (Gósy, 2007). Empirical research on false starts in interpreting demonstrates that such disfluencies can be found both in the output of interpreters and in the source speeches (Pöchhacker, 1995). Pöchhacker attributes them to speech production processes rather than solely to the constraints of simultaneous interpreting. By contrast, Gósy (2007) sees false starts in interpreting as resulting mainly from the constant need to cultivate split attention. According to the researcher, concurrent listening and speaking creates "noise" which triggers false starts.

Research on second language acquisition and learning also offers some insights on disfluencies that may be important for interpreting studies. For example, pausing behaviour has been shown to differ between native and non-native speakers (e.g., Tavakoli, 2011). Unsurprisingly, disfluencies are usually

more frequent in L2 than in L1 speech, but this effect becomes less pronounced as L2 proficiency level raises (e.g., Kosmala, 2021). The *retour* interpreters under investigation here are undoubtedly very proficient L2 speakers, however, they might still produce more disfluencies due to their non-native speaker status.

Importantly, disfluencies are just one aspect of interpreting quality among many. From the perspective of the product, a very fluent interpretation may, for example, contain sense errors or major omissions that impact the coherence and logic of the text and potentially compromise its perception by the audience. From the perspective of the process, while there is a wide consensus among researchers that disfluencies indicate processing problems, their absence is no proof that everything went smoothly (e.g., Gumul 2021, p. 70).

4. Our study

4.1 Material and method

The material for our analysis has been extracted from EP-Poland (see Bartłomiejczyk et al., 2022), a large bidirectional parallel corpus containing all Polish and English contributions to eleven plenary debates of the European Parliament (EP). The debates, held in the years 2016-2020, were selected because of their topic, i.e., the current developments in Poland related to the rule of law crisis and the resulting conflict between the Polish government and the EU. The main aim when compiling the corpus was to obtain material for discourse analytic explorations focusing on ideology, however, the topic also ensured relatively frequent use of Polish as a source language. Consequently, the share of English-Polish and Polish-English interpretations is fairly balanced across the corpus. While interpretations into Polish are exclusively provided by native speakers of Polish, interpretations into English are provided by both native and non-native speakers of English. For the needs of this analysis, three subcorpora were necessary: PL-A with interpretations from English into Polish by native speakers of Polish, EN-A with interpretations from Polish into English by native speakers of English, and EN-B with interpretations from Polish into English by native speakers of Polish.[3]

Individual interpreters were identified in the recordings semi-automatically by the timbre of their voice using the X-vector method (see Bartłomiejczyk & Rojczyk under review for details). Within the whole corpus, 36 individuals were identified, ten of whom interpret in both the directions. However, a threshold for inclusion into the analysis needed to be established so as to account for variable source language input. Considering that the EP speeches are very short (2 minutes 16 seconds on average across EP-Poland) and that disfluencies are likely to occur frequently, we settled on five minutes and at least two different speeches for each interpreting direction. Six bidirectional interpreters met the criteria to be included both in the PL-A and EN-B subcorpora. Additionally, 15 other interpreters were included in the PL-A subcorpus and three in the EN-B subcorpus, i.e., PL-A contains output from 21 interpreters, and EN-B from nine interpreters (see Table 1). The substantial difference in the number of included interpreters between the two subcorpora results from the fact that many Polish interpreters do not work into English and some who do provided too little output to qualify. Out of eleven interpreters in EP-Poland who only work into English, eight exceed the threshold as described

[3] A further subcorpus of English-Polish interpretations by native speakers of English would surely strengthen the analysis, but such interpreting is not practiced in the EP (or anywhere else, as far as we know).

above. However, a phonetic analysis of divergences from native pronunciation norms in which speech samples were inspected aurally and visually by a human assessor using spectrogram and waveform (Bartłomiejczyk & Rojczyk, under review) conclusively shows that only five of these are definitely native speakers of English. Consequently, the output of these five interpreters is included in our EN-A subcorpus (see Table 1).

Table 1: EN-B, PL-A and EN-A subcorpora

Interpreter	Length of interpretations PL-EN (EN-B)	Length of interpretations EN-PL (PL-A)	Length of interpretations PL-EN (EN-A)
Anita	11:11	08:10	-
Daria	06:33	12:10	-
Gustaw	16:15	11:50	-
Jerzy	08:08	10:05	-
Ryszard	11:07	24:30	-
Zygmunt	36:40	13:25	-
Ewa	07:06	-	-
Filip	09:43	-	-
Władysław	11:47	-	-
Amelia	-	06:31	-
Barbara	-	33:41	-
Bogna	-	10:16	-
Dominika	-	06:22	-
Edward	-	21:51	-
Elżbieta	-	06:38	-
Irena	-	12:20	-
Izabela	-	09:11	-
Julia	-	20:07	-
Katarzyna	-	18:21	-
Marcin	-	06:29	-
Mariusz	-	14:38	-
Marta	-	21:46	-
Olga	-	21:34	
Szymon	-	10:52	-
Daniel	-	-	10:13
Henry	-	-	19:28
Julian	-	-	58:44
Michael	-	-	67:52
Susan	-	-	10:46
TOTAL	01:58:30	05:00:07	02:47:03

For ease of reference, the interpreters were given fictitious names corresponding to their gender. The individual interpreters' contributions are not balanced, being a random result of source language distribution across the floor and work division in their respective booths during the debates included in EP-Poland. The total length of target texts in native Polish (five hours) is close to the combined total lengths of target texts in native and non-native English (about fifteen minutes less). EN-A is larger than EN-B, but this reflects the higher share of into-A interpreting across the English target texts in EP-Poland. We are not able to state whether this dominance is due to the general avoidance of *retour* during EP debates or whether these particular debates were serviced by many interpreters with Polish as their C language by way of exception, since frequent use of Polish was predicted on the basis of the topics under discussion.

As we are particularly interested in the performance of individual bidirectional interpreters (Anita, Daria, Gustaw, Jerzy, Ryszard and Zygmunt), their output as listed in Table 1 was further extracted to create smaller subcorpora EN-B/1 and PL-A/1. Their length is 01:29:54 and 01:19:30, respectively.

The disfluency phenomena in which we are interested, i.e., hesitation markers, false starts, and anomalous pauses, were annotated already at the stage of manual transcription of the interpretations to be included in the EP-Poland corpus. Each interpretation was transcribed on the basis of the recording retrieved from the EP website and afterwards verified by another person (see Bartłomiejczyk et al. 2022). Hesitation markers, coded uniformly as <@> irrespective of their actual sound, are non-lexical fillers mainly in the form of prolonged vowels. False starts are retraced and non-retraced truncations at the word level, coded with a hyphen following the interrupted word, e.g., <pol->. Finished words, even if followed by a repair, are not treated as false starts in our taxonomy. Anomalous pauses (coded as <--->) comprise only pauses exceeding three seconds, as such a high threshold should unambiguously point to non-strategic interruptions of the interpreter's speech flow, possibly indicating processing problems.

4.2 Results and discussion

As the subcorpora are of different lengths, the prevalence of each type of disfluency had to be calculated as a normalized frequency per 100 words of target language output. The overall results are presented in Table 2. The subcorpora need to be compared pairwise. To verify Hypothesis 1, we will compare EN-B with EN-A (non-native vs. native interpretations into English) and EN-B with PL-A (non-native vs. native interpretations by members of the Polish booth, native speakers of Polish). To verify Hypothesis 2, we will compare EN-A with PL-A (native interpretations into two different languages). Finally, we will compare EN-B/1 with PL-A/1 (non-native vs. native interpretations by the same individuals) with a view to verifying Hypothesis 3.

Table 2: Quantitative comparison of the three subcorpora

	EN-B	**PL-A**	**EN-A**
Words	16 044	39 256	19 954
Hesitation markers (total/**per 100 words**)	921/**5.74**/M=5.32/SD=2.72	1719/**4.38**/M=4.22/SD=2.32	424/**2.12**/M=2.91/SD=1.26
Anomalous pauses (total/**per 100 words**)	22/**0.14**/M=0.11/SD=0.08	62/**0.16**/M=0.15/SD=0.18	125/**0.63**/M=0.69/SD=0.18
False starts (total/**per 100 words**)	60/**0.37**/M=0.39/SD=0.22	286/**0.73**/M=0.74/SD=0.47	87/**0.44**/M=0.52/SD=0.29
Disfluencies (all three types) **per 100 words**	**6.25**/M=5.82/SD=2.75	**5.27**/M=5.12/SD=2.47	**3.19**/M=4.13/SD=1.66

We will start with a comparison between non-native and native English (EN-B vs. EN-A). At first glance, the overall results reported in Table 2 suggest that native English interpreters produce considerably fewer hesitation markers (2.12 vs. 5.74) and considerably more anomalous pauses (0.63 vs. 0.14). We will examine each type of disfluency separately, using the Mann-Whitney U test, to see if the differences are statistically significant. The normalized

frequency of a given disfluency in each interpreter (see Table 3) is treated as a separate observation. The two samples consist of nine (EN-B) and five (EN-A) interpreters, and there are different individuals in each sample.

Table 3: Disfluencies in non-native and native interpretations into English across individual interpreters (normalized frequencies per 100 words)

EN-B					**EN-A**				
Interpreter	Hesitation markers	Anomalous pauses	False starts	All disfluencies	Interpreter	Hesitation markers	Anomalous pauses	False starts	All disfluencies
Anita	7.91	0	0.24	8.15	Daniel	4.16	0.66	0.73	4.13
Daria	3.87	0.23	0.23	4.33	Henry	4.52	1.04	0.99	6.55
Gustaw	9.59	0.12	0.70	10.41	Julian	2.02	0.69	0.26	2.97
Jerzy	4.61	0	0.45	5.06	Michael	1.17	0.47	0.41	2.05
Ryszard	5.81	0.13	0.79	6.73	Susan	2.71	0.62	0.21	3.54
Zygmunt	5.54	0.23	0.17	5.94					
Ewa	1.72	0	0.40	2.12					
Filip	7.94	0.19	0.10	8.23					
Władysław	0.93	0.12	0.43	1.48					

For hesitation markers, $Z = 1.46$, $U = 34$ and $p = 0.147$; for anomalous pauses, $Z = -2.953$, $U = 0$ and $p = 0.003$; and for false starts, $Z = -0.646$, $U = 17$ and $p = 0.518$. This means that only the difference related to anomalous pauses is statistically significant (at $p < 0.05$), while the other two are not. It may seem surprising as regards hesitation markers, but, notably, the overall result of 2.12 was calculated per 100 words of the total length of the EN-A subcorpus, in which a large majority of the material comes from just two interpreters (Julian and Michael) who both happen to use hesitation markers very sparingly (2.02 and 0.17 per 100 words, respectively). As the selected test treats each interpreter equally, independently of the amount of his/her output, the difference between the two subcorpora decreases and fails to reach statistical significance.

The results are more favourable for the EN-A interpreters as regards hesitation markers (though the trend in the data does not translate into statistical significance) and more favourable for the EN-B interpreters as regards anomalous pauses (which is statistically significant and contradicts the expected directionality effect). Although the surplus pauses in the EN-A subcorpus do not cover the relative deficit in hesitation markers (with hesitation markers occurring much more frequently overall), there might be a negative correlation between these two factors, which requires further study. Pausing is a natural behaviour in interpreting, often strategic as the interpreter needs more input to decipher the speaker's meaning (e.g., Ahrens 2005; Wang & Li 2015). Pauses may be filled or silent, and the native English-speaking interpreters gravitate towards the latter more than their Polish colleagues. As the English and Polish booths appear to be separate communities of practice (Duflou 2016), each of them might have developed somewhat different norms. Perhaps the English booth places more emphasis than the Polish one on avoidance of hesitation markers as obvious indicators of processing problems, and, therefore, its individual members have learned to suppress any vocalization during pauses more effectively.

Additionally, we will consider all the disfluencies under analysis jointly, provisionally assuming that each type carries the same weight, although it is uncertain whether they are equally disturbing for the audience, as relevant research is lacking. Here, $Z = 1.041$, $U = 31$ and $p = 0.298$, which means that the overall difference between the two samples is not statistically significant. On the basis of the overall results presented in Table 2, the EN-B and PL-A

subcorpora appear more similar to each other than the previous pair. Still, the prevalence of false starts is decisively higher in PL-A (0.73 vs. 0.37), while the incidency of hesitation markers is slightly higher in EN-B (5.74 vs. 4.38). Once again, we will perform a separate test for each disfluency type and afterwards consider all of them jointly. The samples consist of nine (EN-B) and 21 (PL-A) interpreters. Inclusion in both the samples partly overlaps, with six interpreters included in both subcorpora. Still, the results are not paired (uneven samples), so we will also use the Mann-Whitney U test.

Table 4: Disfluencies in non-native and native interpretations by Poles across individual interpreters (normalized frequencies per 100 words)

EN-B					**PL-A**				
Interpreter	Hesitation markers	Anomalous pauses	False starts	All disfluencies	Interpreter	Hesitation markers	Anomalous pauses	False starts	All disfluencies
Anita	7.91	0	0.24	8.15	Anita	8.21	0	1.58	9.79
Daria	3.87	0.23	0.23	4.33	Daria	8.28	0.14	0.55	8.97
Gustaw	9.59	0.12	0.70	10.41	Gustaw	3.98	0.06	0.53	4.57
Jerzy	4.61	0	0.45	5.06	Jerzy	1.60	0	1.04	2.64
Ryszard	5.81	0.13	0.79	6.73	Ryszard	6.74	0.08	0.66	7.48
Zygmunt	5.54	0.23	0.17	5.94	Zygmunt	5.34	0.56	0.56	6.46
Ewa	1.72	0	0.40	2.12	Amelia	0.15	0	0.15	0.30
Filip	7.94	0.19	0.10	8.23	Barbara	5.78	0.22	0.55	6.55
Władysław	0.93	0.12	0.43	1.48	Bogna	2.24	0.09	0.75	3.08
					Dominika	7.00	0.12	1.57	8.69
					Edward	3.71	0.04	0.44	4.19
					Elżbieta	4.01	0	0	4.01
					Irena	1.28	0.64	0.56	2.48
					Izabela	1.57	0.43	0.26	2.26
					Julia	4.27	0.12	0.66	5.05
					Katarzyna	2.36	0.13	1.29	3.78
					Marcin	3.86	0.27	1.33	5.46
					Mariusz	4.15	0.07	1.67	5.89
					Marta	7.10	0	0.73	7.83
					Olga	5.49	0.37	0.12	5.98
					Szymon	1.63	0	0.59	2.22

For hesitation markers, $Z = 1.041$, $U = 118$ and $p = 0.298$, for anomalous pauses, $Z = -0.02$, $U = 93.5$ and $p = 0.982$; and for false starts, $Z = -1.947$, $U = 51$ and $p = 0.051$. This means that none of the differences is statistically significant. The samples are almost identical as far as the prevalence of anomalous pauses is concerned, and this is because both EN-B and PL-A interpreters hardly ever produce any such pauses. The difference in terms of false starts is very close to statistical significance, but fails to reach it. We will treat it as indicative of a possible trend that might be explained by systemic differences between Polish and English rather than by directionality effects. As a highly inflectional language, Polish may enforce more planning so as to use correctly inflected nouns, verbs or adjectives. Interpreters may therefore interrupt their production of certain forms once they realize that the sentence they have planned will be grammatically incorrect. In particular, retraced false starts may be hypothesized to have emerged from such processing problems, as the lexical choice apparently remains the same. This hypothesis, however, is impossible to verify on the basis of our material, as typically words are interrupted much earlier than the interpreter reaches their inflectional ending.

As no direct access to the planning process is possible, a more comprehensive examination of repairs of various types might shed more light on this issue.

When all the disfluencies are considered jointly, Z = 0.588, U = 108 and p = 0.556. Unsurprisingly, the overall difference between the two supcorpora is very far from statistical significance. The higher p value (0.556 vs. 0.298) indicates that our intuitive assumption to the effect that the EN-B and EN-A subcorpora are somewhat more divergent than the EN-B and PL-A subcorpora was correct.

Our next step is to compare the EN-A and PL-A subcorpora (see data from the right halves of Table 3 and Table 4, respectively). The sample sizes are rather disproportionate, as EN-A includes five interpreters, and PL-A as many as 21. The samples are independent and we will follow the same protocol as before.

For hesitation markers, Z = -0.911, U = 38, and p = 0.362; for anomalous pauses, Z = 3.208, U = 102, and p = 0.001; and for false starts, Z = -0.846, U = 39, and p = 0.397. When all the disfluencies are considered jointly, Z = -0.813, U = 39.5 and p = 0.416. By analogy with the comparison we made between EN-B and EN-A, only the difference in the prevalence of anomalous pauses is statistically significant, with the native-English interpreters producing more disfluencies of this type than their Polish colleagues. This supports our previous tentative conclusion that the effect may be due to differing booth norms rather than to directionality. The overall p value of 0.416 is in the middle between 0.556 obtained for the juxtaposition of EN-B vs. PL-A and 0.298 for the juxtaposition of EN-B vs. EN-A, which means that the two native language subcorpora are in fact less similar to each other than the two produced by the Polish booth.

Finally, we will scrutinize renditions by the six bidirectional interpreters and the disfluencies they produce in Polish-English and English-Polish interpreting, i.e., the subcorpora EN-B/1 and PL-A/1 (see the first six rows of Table 3, i.e., data for Anita, Daria, Gustaw, Jerzy, Ryszard, and Zygmunt). In this case, there are six interpreters in each sample, and they are all the same individuals (i.e., the samples are dependent and the results are paired).

Interestingly, there are no clear patterns emerging from this individual comparison at first glance. For instance, if we look at hesitation markers (far more frequent in each interpreter's output than either anomalous pauses or false starts), there are interpreters who produce hesitation markers similarly often in each direction (Zygmunt, Anita, and Ryszard, enumerated according to the degree of convergence between their A-B and B-A interpreting), and also interpreters who display a very clear predominance for one interpreting direction (B-A: Daria, A-B: Gustaw and Jerzy). To test our intuition that there is no directionality effect here, we have compared the cumulative results (for all disfluencies) as well as the results for each disfluency type separately across both interpreting directions using the Wilcoxon Signed-Rank Test that is appropriate for paired results obtained for dependent samples. The values obtained in this test (calculated jointly for all three types of disfluencies) are: Z = -4.215, W = 10, and p = 1.000, which indicates that the results for EN-B/1 and PL-A/1 show no difference. Similarly, no difference was recorded for hesitation markers (Z = -4.215, W = 11, p = 1.000) and the difference obtained for false starts is not statistically significant (Z = -1.418, W = 3, p = 0.156). The distribution of the differences in the frequencies of anomalous pauses produced by individual interpreters proved too asymmetrical to perform a reliable calculation of the Wilcoxon Signed-Rank Test. Considering that these frequencies were in fact extremely low (i.e., the interpreters produced hardly any anomalous pauses, some of them none), we decided to abstain from performing any other statistical tests.

To sum up, the findings led us to reject our Hypothesis 3. A possible competing hypothesis that disfluencies, and hesitation markers in particular, might be a feature of individual interpreting style (among other features that have been explored, for instance, by Kajzer-Wietrzny 2013 or Gumul & Bartłomiejczyk 2022) deserves further exploration in a separate study. Such a hypothesis was also advanced by Dayter (2021), who investigated disfluencies in interpreting between English and Russian.

Our intention was also to complement the quantitative analysis with some qualitative findings. However, despite scrutiny of the material, we did not notice any regularities in disfluencies patterns that might be attributable to directionality.

5. Conclusions

Our analysis focused on three types of disfluencies occurring in authentic material from plenary sessions of the EP. We compared three pairs of subcorpora containing A-B and B/C-A interpretations (between Polish and English). Considering the prevailing views in the directionality debate in favour of into-A interpreting, we formulated three hypotheses: 1) that Polish-English interpretations by Poles (A-B) would contain more disfluencies than both English-Polish interpretations by Poles (B/C-A) and Polish-English interpretations by native speakers of English (C-A); 2) that interpretations into the native language, be it Polish or English, would be similarly (dis)fluent; and 3) that individual bidirectional interpreters would produce more disfluencies when working into their non-native language.

As for Hypothesis 1, the only statistically significant difference relates to anomalous pauses, which are more prevalent in EN-A than in EN-B. This contradicts the expected directionality effect, i.e., potentially shows an advantage of into-B interpreting. Overall, our Hypothesis 1 has to be rejected. As for Hypothesis 2, again, a statistically significant difference was shown for anomalous pauses, more prevalent in EN-A than PL-A. As in this case we were trying to confirm the similarity of the subcorpora, finding this difference led us to reject Hypothesis 2. The fact that the overall convergence between EN-A and PL-A is actually smaller than between EN-B and PL-A also speaks against Hypothesis 2. Finally, the comparison between EN-B/1 and PL-A/1, designed to test differences related to interpreting in each direction by the same individuals, failed to show any statistically significant effects, which does not validate Hypothesis 3.

Overall, we have found no directionality effects related to disfluencies that could render support to a conclusion that Polish-English interpretations by native speakers of English are characterized by higher quality and/or are performed with more ease than those provided by their Polish colleagues; or that Polish interpreters perform better when interpreting into their native language than vice versa. While into-A interpreting from multiple languages may be universally preferred over A-B interpreting for other reasons, such as avoidance of relay, or based on quality factors beyond the scope of our study, fluency understood as minimizing the prevalence of the three problem indicators under analysis has not been shown here to depend on interpreting into or out of the interpreter's native language.

Certainly, this study has its limitations, especially as regards the representativeness of the EN-A subcorpus, which includes output from only five members of the English booth. Rather than being a flaw in corpus design, this results from the relatively low popularity of Polish as a C language among native-English interpreters. A more varied sample might be obtained by

including another language in the analysis, preferably a more widely used one such as German or Spanish. However, this would require supplementing the Polish part of the material as well. We envisage that such an extension of the EP-Poland corpus may be possible in the future, as this would surely contribute to new insights into directionality, language-pair effects, inter-booth convergences as well as individual interpreting styles. As regards the Polish-English language pair in the context of the EP, we also have high hopes of the ongoing work of our colleagues from the PINC team (see Chmiel et al., 2022) and look forward to their findings.

Acknowledgement

This paper is an outcome of research activities co-funded by the Research Excellence Initiative of the University of Silesia in Katowice.

References

Ahrens, B. (2005). Prosodic phenomena in simultaneous interpreting: A conceptual approach and its practical application. *Interpreting* *7*(1), 51–76. https://doi.org/10.1075/intp.7.1.04ahr

Al-Salman, S. & Al-Khanji, R. (2002). The native language factor in simultaneous interpretation in an Arabic/English context. *Meta* 47 (4), 607–625.

Bartłomiejczyk, M. (2006). Strategies of simultaneous interpreting and directionality. *Interpreting* *8*(2), 149–174. https://doi.org/10.1075/intp.8.2.03bar

Bartłomiejczyk, M. (2010). Effects of short intensive practice on interpreter trainees' performance. In D. Gile, G. Hansen & N. K. Pokorn (Eds.), *Why Translation Studies matters* (pp. 183–194). John Benjamins.

Bartłomiejczyk, M. (2015). Directionality. In F. Pöchhacker (Ed.), *Routledge encyclopedia of Interpreting Studies* (pp. 108–110). Routledge.

Bartłomiejczyk, M., Gumul, E. & Koržinek, D. (2022). EP Poland: Building a bilingual parallel corpus for interpreting research. *GEMA Online Journal of Language Studies* 22(1), 110–126. https://doi.org/10.17576/gema-2022-2201-06

Bartłomiejczyk, M. & Stachowiak-Szymczak, K. (2022). Modes of conference interpreting: simultaneous and consecutive. In Albl-Mikasa, M. & Tiselius, E. (Eds.), *The Routledge handbook of conference interpreting* (pp. 19-33). Routledge.

Bartłomiejczyk, M. & Rojczyk, A. (under review). How native-like do conference interpreters sound in L2? A phonetic analysis of retour interpretations into English in the European Parliament.

Bendazzoli, C., Sandrelli, A. & Russo, M. (2011). Disfluencies in simultaneous interpreting: A corpus-based analysis. In A. Kruger, K. Wallmach & J. Munday (Eds.), *Corpus-based Translation Studies: Research and applications* (pp. 282–306). Continuum.

Bortfeld, H., Leon, S., Bloom, J., Schober, M. & Brennan, S. (2001). Disfluency rates in conversation: Effect of age, relationship, topic, role, and gender. *Language and Speech* *44*(2), 123–147. https://doi.org/10.1177/00238309010440020101

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

Chang, C. & Schallert, D. (2007). The impact of directionality on Chinese/English simultaneous interpreting. *Interpreting* *9*(2): 137–176. https://doi.org/10.1075/intp.9.2.02cha

Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives* *25*(4), 640–657. https://doi.org/10.1080/0907676X.2016.1278026

Chmiel, A. (2016). Directionality and context effects in word translation tasks performed by conference interpreters. *Poznan Studies in Contemporary Linguistics* *52*(2): 269–295. https://doi.org/10.1515/psicl-2016-0010

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

Chmiel, A., Koržinek, D., Kajzer-Wietrzny, M., Janikowski, P., Jakubowski, D. & Polakowska, D. (2022). Fluency parameters in the Polish Interpreting Corpus (PINC). In M. Kajzer-Wietrzny, A. Ferraresi, I. Ivaska & S. Bernardini (Eds.), *Mediated discourse at the European Parliament: Empirical investigations* (pp. 63–91). Language Science Press. https://doi.org/10.5281/zenodo.6977042

Collard, C., & Defrancq, B. (2020). Disfluencies in simultaneous interpreting: A corpus-based study with special reference to sex. In L. Vandevoorde, J. Daems, & B. Defrancq (Eds.), *New empirical perspectives on translation and interpreting* (pp. 264–299). Routledge.

Dayter, D. (2020). Strategies in a corpus of simultaneous interpreting. Effects of directionality, phraseological richness, and position in speech event. *Meta 65*(3), 594–617. https://doi.org/10.7202/1077405ar

Dayter, D. (2021). Variation in non-fluencies in a corpus of simultaneous interpreting vs. non-interpreted English. *Perspectives 29*(4), 489–506. https://doi.org/10.1080/0907676X.2020.1718170

De Bot, K. (2000). Simultaneous interpreting as language production. In B. Englund Dimitrova & K. Hyltenstam (Eds.), *Language processing and simultaneous interpreting: Interdisciplinary perspectives* (pp. 65–88). John Benjamins,

Defrancq, B., Plevoets, K. & Magnifico, C. (2015). Connective markers in interpreting and translation: where do they come from. In J. Romero Trillo (Ed.), *Corpus pragmatics in translation and contrastive studies, 3* (pp. 195–222). Springer.

Defrancq, B. & Plevoets, K. (2018). Over-uh-load, filled pauses in compounds as a signal of cognitive load. In C. Bendazzoli, M. Russo, & B. Defrancq (Eds.), *Making way in corpus-based interpreting studies* (pp. 43–64). Springer. https://doi.org/10.1007/978-981-10-6199-8_3

Déjean le Féal, K. (2003). Impact of the international status of the interpreting student's mother tongues on training. *Forum 1*(1), 53–76. 1 https://doi.org/0.1075/forum.1.1.04dej

Denissenko, J. (1989). Communicative and interpretative linguistics. In L. Gran & J. Dodds (Eds.), *The theoretical and practical aspects of teaching conference interpretation* (pp. 155–158). Campanotto.

Donovan, C. (2004). European Masters Project Group: Teaching simultaneous interpretation into a B language. *Interpreting 6*(2), 205–216.

Duflou, V. (2016). *Be(com)ing a conference interpreter: An ethnography of EU interpreters as a professional community.* John Benjamins.

Gieshoff, A.C. (2021). The impact of visible lip movements on silent pauses in simultaneous interpreting. *Interpreting 23*(2), 168–191. https://doi.org/10.1075/intp.00061.gie

Gile, D. (2005). Directionality in conference interpreting: A cognitive view. In R. Godijns, & M. Hinderdael, (Eds.), *Directionality in interpreting: The "retour" or the native?* (pp. 9–26). Communication & Cognition.

Guilquin, G. (2008). Hesitation markers among EFL learners: Pragmatic deficiency or difference? In J. Romero Trillo (Ed.), *Pragmatics and corpus linguistics: A Mutualistic entente* (pp. 119–149). Mouton de Gruyter.

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

Graves, A., Pascual Olaguíbel, M. & Peason, C. (2022). Conference interpreting in the European Union Institutions. In M. Albl-Mikasa, & E. Tiselius (Eds.), *The Routledge handbook of conference interpreting* (pp. 104–114). Routledge.

Gumul, E. (2017a). Explicitation and directionality in simultaneous interpreting. *Linguistica Silesiana 38*, 311–329.

Gumul, E. (2017b). *Explicitation in simultaneous interpreting. A study into explicitating behaviour of trainee interpreters.* Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Gumul, E. (2021). Explicitation and cognitive load in simultaneous interpreting: Product- and process-oriented analysis of trainee interpreters' outputs. *Interpreting 23*(1), 45–75. https://doi.org/10.1075/intp.00051.gum

Gumul, E. and Bartłomiejczyk, M. (2022). Interpreters' explicitating styles. A corpus study of material from the European Parliament. *Interpreting 24*(2), 163–191. https://doi.org/10.1075/intp.00081.gum
Hale, S. (2004). *The discourse of court interpreting.* John Benjamins.
Han, Chao and An, K. (2021). Using unfilled pauses to measure (dis)fluency in English-Chinese consecutive interpreting: in search of an optimal pause threshold(s). *Perspectives* 29(6), 917-933. https://doi.org/10.1080/0907676X.2020.1852293
Igras-Cybulska, M., Ziółko, B., Żelasko, P. & Witkowski, M. (2016). Structure of pauses in speech in the context of speaker verification and classification of speech type. *Journal on Audio, Speech, and Music Processing 18*. https://link.springer.com/article/10.1186/s13636-016-0096-7
Kajzer-Wietrzny, M. (2013). Idiosyncratic features of interpreting style. *New Voices in Translation Studies 9*, 38–52.
Kosmala, L. (2021). On the specificities of L1 and L2 (dis)fluencies and the interactional multimodal strategies of L2 speakers in tandem interactions. *Journal of Monolingual and Bilingual Speech 3*(1): 69–101. https://doi.org/10.1558/jmbs.15676
Kurz, I. (1994). A look into the 'black box' – EEG probability mapping during mental simultaneous interpreting. In M. Snell-Hornby, F. Pöchhacker, & K. Kaindl (Eds.), *Translation studies: An interdiscipline* (pp. 199-207). John Benjamins.
Kurz, I. & Färber, B. (2003). Anticipation in German–English simultaneous interpreting. *Forum 1*(2), 123–150. https://doi.org/10.1075/forum.1.2.06kur
Lin, Y., Lv, Q. & Liang, J. (2018). Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting. *Frontiers in Psychology* 9. https://doi.org/10.3389/fpsyg.2018.01543
Mead, P. (2000). Control of pauses by trainee interpreters in their A and B Languages. *The Interpreters' Newsletter 10*, 89–102.
Monti, C., Bendazzoli, C., Sandrelli, A. & Russo, M. (2005). Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus). *Meta*, *50*(4). https://doi.org/10.7202/019850ar
Plevoets, K. & Defrancq, B. (2016). The effect of informational load on disfluencies in interpreting. A corpus-based regression analysis. *Translation and Interpreting Studies 11*(2), 202 – 224. https://doi.org/10.1075/tis.11.2.04ple
Plevoets, K. & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency *uh(m)*. *Interpreting 20*(1), 1–32. https://doi.org/10.1075/intp.00001.ple
Pöchhacker, F. (1995). Slips and shifts in simultaneous interpreting. In J. Tommola (Ed.), *Topics in interpreting research* (73-90). The University of Turku, Centre for Translation and Interpreting.
Nicodemus, B. & Emmorey, K. (2013). Direction asymmetries in spoken and signed language interpreting. *Bilingualism: Language and Cognition 16*(3): 624–636. https://doi.org/10.1017/S1366728912000521
Opdenhoff, J-H. (2012). Directionality and working memory in conference interpreting – an experimental study. In B. Harris, B. Mayor, M. Jesus, J. Ivars & M. Amparo (Eds.), *Interpreting Brian Harris – Recent developments in translatology* (pp. 161–171). Peter Lang.
Pradas Macías, E. M. (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting* 8 (1), 25–43. https://doi.org/10.1075/intp.8.1.03pra
Rennert, S. (2010). The impact of fluency on the subjective assessment of interpreting quality. *The Interpreters' Newsletter 15*, 101–115.
Ruiz Rosendo, L., & Diur, M. (2022). Conference interpreting at the United Nations. In M. Albl-Mikasa, & E. Tiselius (Eds.), *The Routledge handbook of conference interpreting* (pp. 115-125). Routledge.
Seleskovitch, D. (1968/1978). *Interpreting for international conferences: Problems of language and communication* (Trans. Stephanie Dailey & Eric N. McMillan). Pen & Booth.
Seleskovitch, D. & Lederer, M. (1989/1995). *A systematic approach to teaching interpretation* (Trans. J. Harmer). RID Publications.
Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap:*

Empirical research in simultaneous interpretation (pp. 225–236). John Benjamins,

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal 65*, 71–79.

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

Tomczak, E., & Whyatt, B. (2022). Directionality and lexical selection in professional translators: Evidence from verbal fluency and translation tasks. *Translation & Interpreting 14*(2), 120–136. https://doi.org/10.12807/ti.114202.2022.a08

Van Dijk, R., Boers, E., Christoffels, I., & Hermans, D. (2011). Directionality effects in simultaneous language interpreting: The case of sign language interpreters in the Netherlands. *American Annals of the Deaf 156*(1), 47–55. http://www.jstor.org/stable/26235125

Wang, B. & Li, T. (2014). An empirical study of pauses in Chinese-English simultaneous interpreting. *Perspectives 23*(1), 124–142. https://doi.org/10.1080/0907676X.2014.948885

Whyatt, B. & Pavlovič, N. (2021). Translating languages of low diffusion: Current and future avenues. *The Interpreter and Translator Trainer* 15(2), 141–153. https://doi.org/10.1080/1750399X.2021.1917172

Wu, Y. & Liao, P. (2018). Re-conceptualising interpreting strategies for teaching interpretation into a B language. *The Interpreter and Translator Trainer 12*(2), 188–206. https://doi.org/10.1080/1750399X.2018.1451952