Disfluency in relay and non-relay simultaneous interpreting: an initial exploration

Andrew K.F. Cheung

Hong Kong Polytechnic University

Shuxian Song

Qufu Normal University

This corpus-based study explores the effects of relay interpreting at meetings of the United Nations General Assembly by comparing features of disfluency between the outputs of relay and non-relay simultaneous interpreting (SI). The findings are as follows: (1) the output of relay interpreting is shorter and more dispersive than that of non-relay interpreting; (2) filled pauses are the most common type of disfluency; and (3) the relay SI output shows fewer lexical and phonetic E-repairs and more A-repairs for ambiguity, syntactic E-repairs, and D-repairs than the non-relay output. The results suggest that the use of relay vs. non-relay interpreting may affect interpreters' output.

Keywords: relay interpreting; disfluency; repairs; filled pauses; United Nations

1. Introduction

This study explores types of disfluency in the simultaneous interpreting (SI) output of professional interpreters with the United Nations (UN) interpreting from English (non-native ("B") language) into Chinese (native ("A") language). It taps into the effects of relay interpreting on SI output and explores the underlying causes of these effects by describing and comparing features of text and disfluency annotations in relay and non-relay interpreting corpora.

The importance of fluency, a complex and high-order linguistic phenomenon (Riggenbach, 1991), as a criterion for interpreting assessment has been recognized in a number of surveys asking users of interpreting services to comment on the aspects of interpreting output they most value (Bühler, 1986; Kurz, 1993; Moser, 1996; Pöchhacker, 2012). Although an interpreter's fluency ultimately provides no guarantee of the reliability of the interpretation, its contribution to the overall rendering of the target speech is far from negligible, making fluency an important feature of successful interpreting (Mead, 2000, 2005). Gósy (2007, 93) defines disfluencies as “phenomena that interrupt the flow of speech and do not add propositional content to an utterance.” These phenomena may help us to understand how interpreting output may reflect underlying cognitive processing (Shreve, Lacruz, & Angelone, 2011). Analysis of disfluencies can shed light on the effects of relay interpreting.

Relay interpreting is common in the UN and numerous other international organizations. Relay interpreting is the practice of interpreting from one language to another through a third language (Shlesinger, 2010). Speeches in any of the six official languages of the UN (Arabic, Chinese, English, French, Russian and Spanish) are interpreted simultaneously into the other five. Most UN interpreters interpret only into their “A” languages, but Chinese interpreters interpret from and into Chinese. Due to their training, Chinese interpreters tend to have English and/or French as their “B“ language(s). Normally, relay interpreting into Chinese is needed when a speaker uses Arabic, Russian, or Spanish; Chinese interpreters rely on either English or French renditions to interpret these speeches into Chinese.

Although relay interpreting is common, it has received little attention from researchers. Relay interpreting has features distinct from those of non-relay interpreting, such as the more severe time constraints resulting from a further delay of the input message and the double role of pivot interpreters as both speakers for relay takers and interpreters for interpreting listeners. These unique features may require extra effort from interpreters, whether pivots or relay takers, and additional attention from interpreting service users. Research on whether relay or non-relay interpreting presents more difficulties for interpreters is inconclusive (Mackintosh, 1983; Seleskovitch & Lederer, 1989a, 1989b; Setton & Dawrant, 2016). This study adopts a corpus-based approach using naturalistic data to analyze the features of disfluency, shedding light on the effects of relay interpreting. The study’s corpus annotation framework is based on previous taxonomies of disfluency and the premise that disfluencies predict interpreters’ cognitive processes.

2. Research background

2.1 Disfluencies as predictors

Speech disfluencies may reflect an increase in cognitive effort demanded by lexical or syntactic uncertainty, planning, or production problems (Shreve et al., 2011). Each type of disfluency can be linked to a certain stage in the speech production process, from conceptual planning and grammatical encoding to articulatory planning and monitoring (Bakti, 2009; Gósy, 2005; Levelt, 1983; Schnadt & Corley, 2006). For instance, the duration of a filled or unfilled pause may indicate planning processes (Dechert & Raupach, 1980), and self-repairs reflect the monitoring mechanism used to verify the correctness of ongoing motor activity and response output (Postma, 2000). The study of fluency in interpreting could inform professional development and training pedagogies because research findings could reflect the difficulties facing interpreters.

SI is a complex task combining language comprehension and production in real time, and these two processes compete for the interpreter's limited attentional resources (Chernov, 2004; Gile, 2009; Seeber, 2011). Errors, omissions, infelicities, and disfluencies are indicators of a deterioration in quality (Gile, 2011). Plevoets and Defrancq's (2016) findings suggest that lexical density and the delivery rate of the source speech lead to disfluencies in interpreting. Gile (2009) suggests that performance may be injured by cognitive, linguistic, or cultural factors such as speech density, source speech quality, signal vulnerability, cross-linguistic differences, or the speaker's individual speaking style.

2.2 Disfluency in simultaneous interpreting

Four types of disfluencies are generally observed: false starts, repetitions, repairs, and filled pauses. A false start is an utterance that is begun and then abandoned or reformulated in some way, and repetition occurs when previously produced speech is produced again; the latter device may be used to allow time for on-line planning (Foster, Tonkyn, & Wigglesworth, 2000). Disfluent repetition differs from repetition used for rhetorical effect. "Repetition" in this study refers only to disfluent repetition.

A repair is a self-correction carried out by a speaker when s/he identifies an error during or immediately after production and stops to reformulate the speech produced (Foster et al., 2000). Levelt (1983) distinguishes overt repairs (post-articulatory) from covert repairs (pre-articulatory). Overt repairs are divided into three major categories:

A-repairs, E-repairs, and D-repairs. A-repairs are performed to address potential ambiguity in context, the use of appropriate terminology, and coherence with previously used terms or expressions. E-repairs are repairs of errors, sub-divided into lexical, syntactic, and phonetic repairs. D-repairs involve the replacement of the current message with a different message, and originate from an error in the conceptualizer (Kormos, 1999). Covert repairs (C-repairs) are characterized by the use of an interruption and editing term(s), with no morphemes changed, added, or deleted, or by the repetition of one or more lexical items. C-repairs in this study's corpus are represented by filled pauses and repetitions**.**

Pauses are periods during which no acoustic signal occurs for at least 200-270 ms (Hargrove & McGarr, 1994). Pauses are categorized as filled or unfilled/silent pauses. Filled pauses may reflect brief attention to planning or retrieval, and are in most cases included in the divisions of disfluency. Filled pauses are categorized as fillers or the lengthening of syllables or words. Unfilled pauses refer to silent periods between vocalizations, and filled pauses to interruptions of speech flow by non-lexical sounds such as "ah," "er," or "erm" (Cenoz, 1998). As a sub-parameter of fluency, unfilled pauses are expected to have negative effects on the evaluation of fluency in SI (Pradas Macías, 2006). However, unfilled pauses have multiple roles in language production, and their functions may be grammatical, communicative, or hesitative (Simone, 1995). Whether a silent pause is disfluent is subjective, and standards for judgment may not be consistent across audiences or occasions. Therefore, unfilled pauses are not included in the current study of disfluency in SI output.

Most studies of disfluency in SI address the distribution of disfluency types. Generalization is difficult, because previous studies (Bakti, 2009; Bendazzoli, Sandrelli, & Russo, 2011; Cecot, 2001; Tissi, 2000) vary in their comparison of disfluency features in the source speech and interpreting output; their selection of variables affecting disfluency, such as input rate; their use of qualitative versus quantitative analysis; and their taxonomies of disfluency.

Pöchhacker (1995) explores fluency in terms of tempo, pauses, voiced hesitation, and false starts, providing a good example of quantitative analysis of the features of fluency. Tissi (2000) proposes a SI-specific functional taxonomy of disfluency and explores possible correlations between disfluencies in source and target texts. Bakti's

(2009) findings show that restarts and grammatical errors are the most frequently occurring error-type disfluencies, followed by grammatical errors and false word activation. Problems in coordinating lexical access may explain restart signals, and articulatory planning and grammatical errors may be triggered by difficulties in morphological and syntactical planning. Finally, high input speed may lead to either filled or unfilled pauses (Cecot 2001).

Focusing on repairs, a major type of SI disfluency, Petite (2005) analyzes three kinds of post-articulatory repair and one category of mid-articulatory repair. The data are drawn from the output of eight professional conference interpreters interpreting from B to A languages and recorded at four international conferences on topics of general interest. The interpreters expended processing resources on producing repairs, including non-obligatory ones, and tended to repair even when such repairs were not cost-effective. Quantitative analysis reveals that the majority of repairs in the corpus were output-generated with fewer E-repairs than A-repairs, D-repairs or mid-articulatory repairs.

2.3 Effects of relay interpreting

Relay interpreting has unique features. Relay interpreters are subject to the usual constraints on cognitive resources and time induced by SI. However, relay interpreting involves a greater delay than non-relay interpreting, because the relay interpreter has to wait for the rendition from the pivot interpreter before interpreting, as opposed to interpreting directly from the speaker. Some researchers are convinced that relay interpreting presents more difficulties for interpreters than direct interpreting. Due to the time lag between speaker and pivot interpreter, relay takers are subject to more severe time constraints than direct interpreters. In addition, the relay taker's inability to rely on the original speaker's prosodic features and lack of familiarity with the original language and source culture may cause difficulties in understanding (Seleskovitch & Lederer, 1989a, 1989b; Setton & Dawrant, 2016). The pressure on both the pivot and the relay taker is greater when synchronicity is more important (Shlesinger, 2010) increasing the probability that output quality may suffer.

However, other researchers argue that relay interpreting does not necessarily reduce quality. Mackintosh (1983) finds no significant difference in message loss between direct and relay interpreting. Quality of relay improves when relay takers are assisted by good pivot interpreters who make special effort to provide clear and coherent interpretations (Seleskovitch & Lederer, 1989a; Setton & Dawrant, 2016). Relay takers may also devote more attention to ensuring a fluent output because the information incoming from the pivot has already been processed and clarified. As a result, the quality of relay takers' renditions may not suffer.

Whether relay interpreting presents more difficulties for interpreters than direct interpreting is debatable. Studies of relay interpreting are few, and disfluency in relay output has not been extensively investigated. This study contributes to the understanding of relay interpreting quality through analysis of authentic UN data.

3. Methodology

3.1 Data source

The source of the data used in this research is the Chinese SI output of eight speeches selected randomly from the 70th session of the UN General Assembly. To compare the features of relay and non-relay interpreting, four speeches in English (by speakers from the U.K., the U.S., Australia, and New Zealand) are analyzed. All of these speeches were originally interpreted through direct SI. Three of the remaining four speeches are in Arabic (by speakers from Egypt, Saudi Arabia, and Syria) and the last is in Spanish (by a speaker from Argentina). Due to their language combinations, UN Chinese interpreters interpreting speeches in Arabic, Russian and Spanish normally need to relay from the English booth.

The corpus, a set of interpreting renditions collectively lasting for about 122 minutes, contains 22,457 Chinese characters, with 426 annotations of types of disfluency. The speeches given by speakers by Australia, the U.K., and New Zealand are given in full, as their total duration is no more than 20 minutes. For the other five speeches, only the first 16 minutes of the interpreting output, on average, are included. This helps to eliminate the effects of fatigue, as prolonged turns may have negative effects on the quality of an interpreter's output and the interpreter's attitude toward the task (Moser-Mercer, Künzli, & Korac, 1998).

3.2 Controlled variables

Variables that may affect the reliability of the data in the relay and non-relay interpreting outputs are controlled to make the sub-corpora comparable. These controlled variables comprise interpreter competence, nature of source speech, speaker's accent, and delivery speed. All of the interpreters under study are professional interpreters working for the UN under the same working conditions; they have similar levels of interpreting competence and are guided by the same professional norms. The source speeches are all formal, political speeches delivered by political leaders on the same occasion, and are therefore comparable in terms of content and style. In terms of speaker accent, the speakers in the non-relay group are all native speakers of English. In the relay group, interpreters serving the English booth are English "A" interpreters and have no strong foreign or regional accents.

In the corpus of this study, the average input rates for the non-relay group and the relay group (i.e. of the pivots' rendition) are 117 and 110 words per minute, respectively. Both are within the ideal range of input rates for SI, and neither shows significant variance.

3.3 Annotation

The source speeches and respective interpreting renditions are transcribed and synchronized using the software package ELAN 5.0.0, a professional tool for the creation of complex annotations on video and audio resources. ELAN can convert acoustic signals into an oscillogram to visualize sounds as a continuous wave pattern. Synchronization with the timeline of the oscillogram gives the annotation markers an accuracy of up to 1 ms.

Table 1. Annotation system for the interpreting output

<table>
<tr><td>Tier 1</td><td colspan="4">Texts</td></tr>
<tr><td rowspan="9">Tier 2</td><td rowspan="9">Disfluency Types</td><td colspan="3">False Starts</td></tr>
<tr><td colspan="3">Repetitions</td></tr>
<tr><td rowspan="6">Repairs</td><td rowspan="2">A-repairs</td><td>Ambiguity</td></tr>
<tr><td>Coherence</td></tr>
<tr><td rowspan="3">E-repairs</td><td>Lexical</td></tr>
<tr><td>Syntactic</td></tr>
<tr><td>Phonetic</td></tr>
<tr><td colspan="2">D-repairs</td></tr>
<tr><td>Filled Pauses</td><td colspan="2">Lengthening</td></tr>
</table>

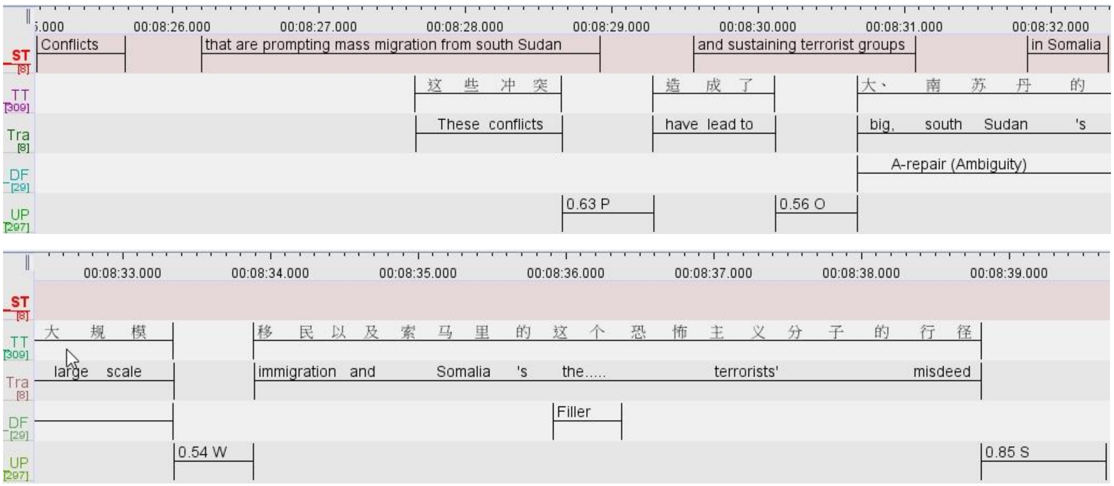
			Fillers
Tier 3	Unfilled Pauses		

Table 1 illustrates the system of annotation of the SI output. Three tiers of annotation are created, highlighting the texts of the interpreting renditions (Tier 1), types of disfluency (Tier 2), and unfilled pauses (Tier 3). For the source speeches, only the texts are annotated, because disfluencies in these speeches are rare. The tiers of the source speeches and their SI output are merged by the software to create synchronized versions. ELAN files of the speeches and corresponding interpreting output are merged when necessary to facilitate data analysis.

For the texts, an annotation refers to a run of words between two unfilled pauses (≥ 0.25 s). Unfilled pauses necessary for the segmentation of text annotations are marked in Tier 3. Following Cecot (2001), Duez (1982), Goldman-Eisler (1958), and Tissi (2000), the threshold for unfilled pauses in this study is set at 0.25 s.

Based on the taxonomies of disfluency in Cecot (2001), Gósy (2005), Levelt (1993), and Tissi (2000), the types of disfluency identified in this study include repetitions, false starts, repairs (A-repairs, E-repairs, and D-repairs), and filled pauses (lengthening and fillers). Figure 1 provides an example of annotation using ELAN.

Figure 1. An annotation example in Elan



Notes: Tier 1 (ST) is the transcription of source speech; Tier 2 (TT) is the transcription of interpreting output, synchronized with the time line; Tier 3 (Tra) is the English back translation of Tier 2; Tier 4 (DF) marks the types of disfluency; Tier 5 (UP) marks the duration and position of unfilled pauses.

4. Data analysis and discussion

ELAN offers statistical information on the frequency and duration of the annotations, synchronized with the timeline of the oscillogram. This section presents the corpus data from two perspectives, providing an overview of the annotation statistics and an analysis of disfluency in the SI output.

4.1 Overview of annotation statistics

Table 2 displays annotation statistics for the relay and non-relay SI output, providing an overview of disfluency frequency and duration in the texts of the corpus.

Table 2. Annotations of texts and disfluencies of SI output

Annotations	Type	Total No.	Average Duration	Median Duration	Annotation Duration Percentage	Mean Length of Run
Texts	Relay	1,357	1.59	1.31	62.92	7.21
	Non-relay	1,269	2.08	1.68	76.80	10.30
Disfluency	Relay	226	1.79	1.38	11.82	-
	Non-relay	200	1.78	1.47	10.38	-

Notes[②]:
Mean length of run: the total words count divided by the number of annotations
Median duration: the median duration of the annotations on that tier
Total annotation duration: the total duration of all annotations on that tier
Annotation duration percentage: the percentage of the total annotation

The total word count of the relay interpreting output is smaller than that of the non-relay interpreting output (9,756 vs. 14,757); more text annotations, i.e. runs of words between two unfilled pauses (≥ 0.25 s), are observed in the relay output than in the non-relay output (1,357 vs. 1,269); and both the average duration and the median duration of text annotations are shorter in the relay output. The relay interpreting output exhibits shorter runs of words with a more dispersive pattern, as evidenced by the shorter mean length of runs in this output. This may indicate that more segmentations are found in the relay interpreting output, and that the relay interpreters followed the pivot interpreters' messages more closely than the non-relay interpreters followed the source speakers' messages (in most cases with a shorter ear to voice span, EVS). Relay interpreting inherently contains two time lags: the first between the source speech and the pivot interpreter's production, and the second between the rendition of the pivot interpreter and that of the relay interpreter. The relay interpreters investigated in this study may have focused on the need for synchronicity and sought to avoid lagging behind.

The percentage duration of text annotations in the relay interpreting output is much lower than that in the non-relay interpreting output (62.92% vs. 76.80%). The relay takers generally produced less output than their non-relay counterparts. This suggests that the pivot interpreters may have streamlined the information provided by eliminating redundant messages and restructuring complex sentences to give simpler ones.

In terms of disfluency, the average annotation duration of disfluency in the relay SI output is almost the same as that in the non-relay SI output, but the relay output has a slightly smaller median duration of disfluency. This suggests that relay and non-relay interpreters spend similar amounts of time on information processing when they encounter difficulties. Comparing the disfluency annotations in the relay and non-relay SI outputs reveals that disfluencies are more frequent in the relay interpreting output (226 vs. 200), and account for a larger proportion of the overall duration of this output (11.82% vs. 10.38%), although the overall word count of the relay output is smaller. This implies that the relay interpreters encountered more difficulties and uncertainties than their non-relay counterparts.

4.2 Analysis of disfluencies in SI output

4.2.1 Overview of disfluency distribution

The corpus of this study contains 426 occurrences of disfluency (repairs, filled pauses, repetitions, or false starts).

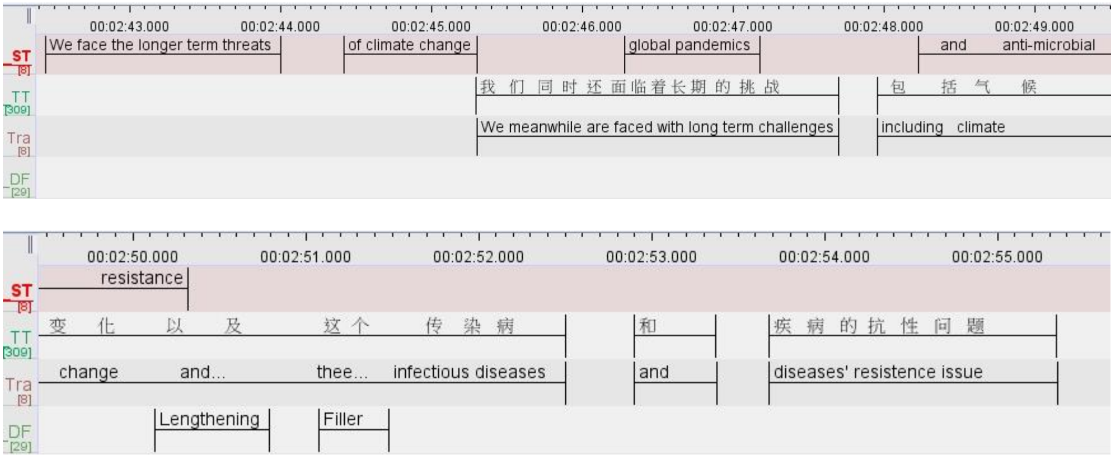
Table 3. Disfluency distribution in general (percentage)

Disfluency	Repairs	Filled Pauses	Repetition	False Start	Sum
Percentage	27%	59.39%	9.39%	4.23%	426

Filled pauses constitute the majority (59.39%) of the disfluent utterances, followed by repairs; however, the rate of repairs is much lower (27%), as shown in Table 3.

Although interpreters show idiosyncrasies, most filled pauses signify difficulties in production (Gile, 2009). Both filled and unfilled pauses are used to address lexical, morphological and planning difficulties, but filled pauses are a more common response to planning difficulties (Cenoz, 1998; Setton, 1999). The following case exemplifies this kind of processing difficulty.

Figure 2. An example of filled pauses (from the UK speech)



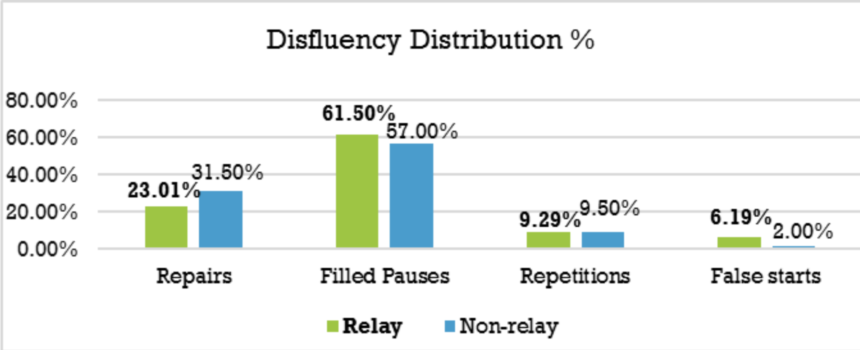
Notes: Tier 1 (ST) is the transcription of source speech; Tier 2 (TT) is the transcription of interpreting output, synchronized with the time line; Tier 3 (Tra) is the English back translation of Tier 2; Tier 4 (DF) marks the types of disfluency.

The first segment of the source speech (Tier 1), "we face the long term threats," is followed by the enumeration of three phrases, "climate change, global pandemics and anti-microbial resistance." Having interpreted (Tiers 2 and 3) "climate change," the interpreter prolongs the enunciation of "and" and the following filler, "thee," perhaps to gain more time to process the subsequent message, "global pandemics and anti-microbial resistance," which is lengthy and contains a medical term, requiring substantial processing effort.

4.2.2 Disfluency in relay and non-relay interpreting

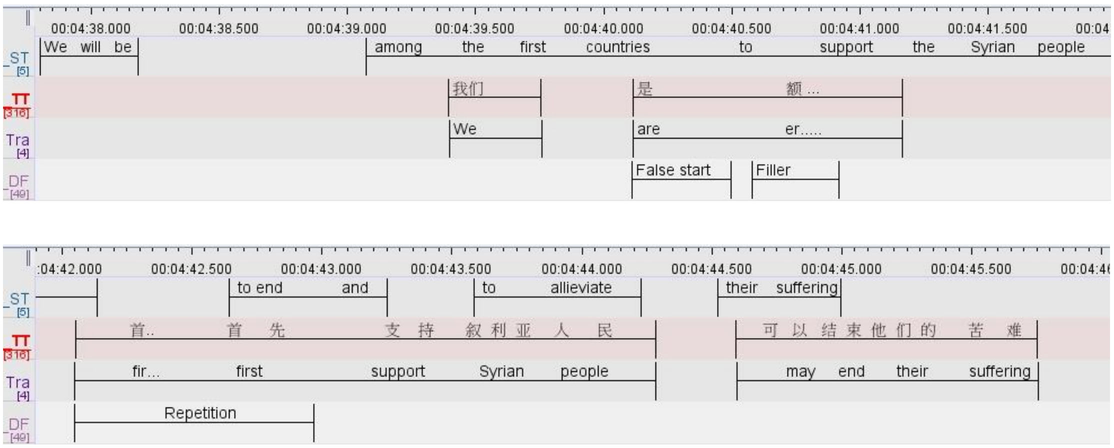
Chart 1 compares the distributions of disfluency types in the relay and non-relay output. Filled pauses (lengthening and fillers), which indicate hesitation and planning problems, make up the majority (61.5% and 57% respectively) of the disfluent utterances in both outputs, followed by repairs.

Chart 1. Disfluency distribution in relay and non-relay interpreting



Repairs make up a smaller proportion (23.01%) of the disfluencies in the relay interpreting output than in the non-relay output (31.50%), and both filled pauses and false starts occupy a slightly larger proportion of the disfluencies in the relay output than in the non-relay output, with a similar percentage for repetitions. The smaller proportion of repairs in the relay relative to the non-relay output may be due to the more severe time constraints faced by relay interpreters. The slightly higher proportions of filled pauses and false starts in the relay interpreting output imply that compared with non-relay interpreters, relay interpreters encounter more difficulties and uncertainties in the planning and message formulation stages. The following example illustrates difficulties of this kind.

Figure 3. An example of false start (from the Arabian speech)



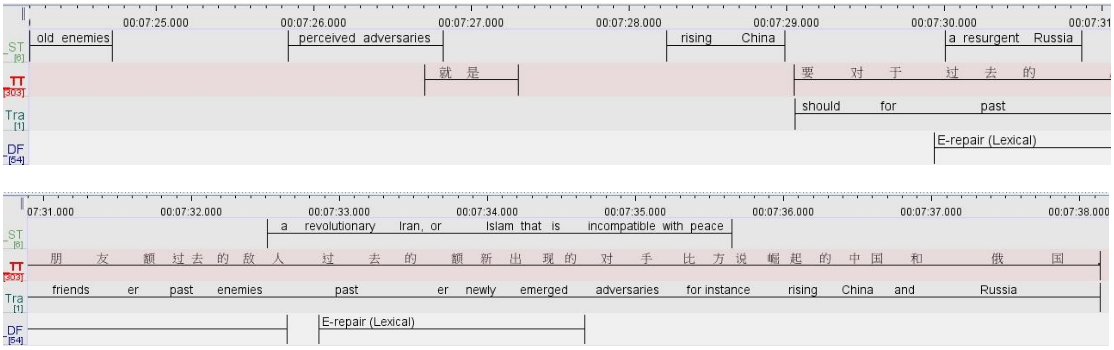
Notes: Tier 1 (ST) is the transcription of source speech; Tier 2 (TT) is the transcription of interpreting output, synchronized with the time line; Tier 3 (Tra) is the English back translation of Tier 2; Tier 4 (DF) marks the types of disfluency.

The Chinese rendition shown in Figure 3 begins with a false start, “we are…” As this could not serve as a grammatically correct opening to the following sentence, the interpreter began again with “first support Syrian people.” This false start is followed by a filler, “er,” and the repetition of a single Chinese character, which suggests that the interpreter experienced additional difficulty and uncertainty while formulating the output.

4.2.3 Probing into repairs

Repairs are the second most prolific type of disfluency in both the relay and the non-relay interpreting output, as shown in Chart 1. A repair typically consists of the original utterance (containing the reparandum), the editing phase (a period of hesitation), and the repair proper (Levelt, 1983). In this corpus, a repair annotation comprises the reparandum, the editing phase (silent pause or fillers), and the repair proper, as illustrated in the following example.

Figure 4. An example of E-repair (from the US speech)



Notes: Tier 1 (ST) is the transcription of source speech; Tier 2 (TT) is the transcription of interpreting output, synchronized with the time line; Tier 3 (Tra) is the English back translation of Tier 2; Tier 4 (DF) marks the types of disfluency.

The above example shows two lexical E-repairs within a single utterance. The first is undertaken to correct “past friends” to “past enemies,” and the second to correct “past” to “newly emerged.” Taking the first repair, “past friends (er) past enemies” as an example, this annotation of repair has three parts: the reparandum, “past friends”; the editing phrase, with the filler “er”; and the repair, “past enemies”.

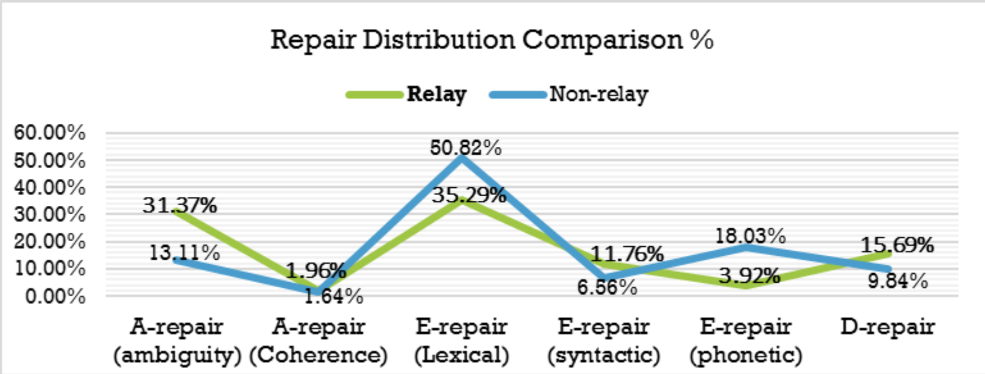
Table 4. Distribution of Repair Types

Repairs	A-repairs (ambiguity)	A-repairs (coherence)	D-repairs	E-repairs (Lexical)	E-repairs (syntactic)	E-repairs (phonetic)
	21.43%	*1.79%*		*43.75%*	*8.93%*	*11.61%*
	23.21%		12.5%	64.29%		

E-repairs account for the majority (64.29%) of the repairs in the corpus, occurring considerably more frequently than A-repairs and D-repairs combined. This distribution of repair types is consistent with Petite's (2005) finding that D-repairs make up the smallest proportion (12.5%) of relay interpreting output and occur when a different message is constructed to replace the current one. In contrast with natural language production, interpreting generates a message provided by an original speaker instead of the interpreter. However, the interpreter also needs to produce an accurate message. D-repairs in interpreting signify problems with the comprehension and formulation of the original message that require the interpreter to produce a new message to replace it, thus requiring additional cognitive effort. The focus of this paragraph is E-repairs.

In terms of the sub-classification of types of repair, most (21.43%) of the A-repairs in the corpus used in this research are undertaken to repair ambiguity; A-repairs for coherence occur rarely (1.79%). The A-repairs do not involve elaboration. Of the E-repairs observed, lexical E-repairs rank highest (43.75%) and syntactic E-repairs lowest (8.93%). A lexical E-repair normally occurs instantly after an interpreter has detected a lexical error, reflecting the interpreter's monitoring process. The large proportions of lexical E-repairs and A-repairs for ambiguity (no such subtype analysis for A-repairs in Table 4) provide evidence of interpreters' adherence to the norms of accuracy and clarity, which helps them to avoid making errors. Compared with other types of repair, syntactic E-repairs require more processing time and cognitive effort, which may explain why interpreters, who are under severe time and cognitive pressure, perform this kind of repair less frequently than other types. The relative lack of phonological interference between English and Chinese, two very different languages (the former an Indo-European language and the latter a Sino-Tibetan language), is believed to partially explain the small proportion (11.61%) of phonetic E-repairs.

Chart 2. Repair Distribution in relay and non-relay



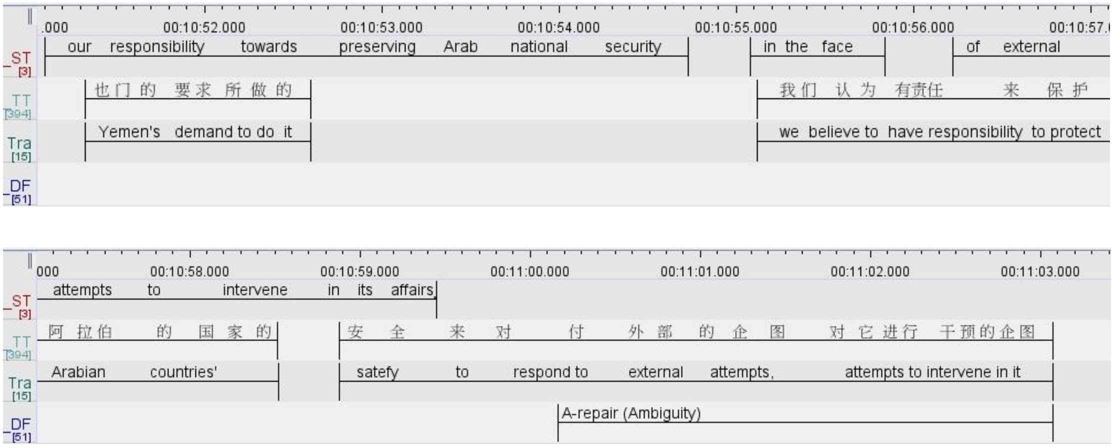
As illustrated in Chart 2, lexical E-repairs constitute the majority of all types of repair in both the relay and the non-relay interpreting output (50.82% and 35.29% respectively). The proportion of A-repairs for ambiguity in the relay interpreting output is more than double that in the non-relay interpreting output (31.37% vs. 13.11%), whereas A-repairs for coherence make up very similar proportions of the two corpora.

Lexical E-repairs account for a smaller proportion of the overall E-repairs in the relay interpreting output than in the output of the non-relay group (35.29% vs. 50.82%). A similar trend is observed for phonetic E-repairs, which occur rarely (3.92%) in the relay interpreting output—substantially less often than in the non-relay output (18.03%). Syntactic E-repairs occur more frequently in the relay interpreting output than in the non-relay output (11.76% vs. 6.56%). In addition, D-repairs constitute a larger proportion of the relay interpreting output (15.69%) than the non-relay renditions (9.84%).

The lower frequency of lexical and phonetic E-repairs in the relay interpreting output indicates that relay interpreters tend to make fewer errors than their non-relay counterparts at the syllabic and lexical levels. The message pre-processed by a pivot interpreter may have been streamlined relative to that delivered by the original speaker. In addition, relay takers perform more segmentations under more severe time constraints, so a shorter EVS may to some extent reduce interpreters' cognitive load, facilitating the production of syllabically and lexically accurate renditions.

The higher proportions of A-repairs for ambiguity, syntactic E-repairs, and D-repairs in the relay than the non-relay output imply that relay interpreters face more uncertainties than non-relay interpreters at the phrasal and syntactic levels. Relay interpreters seem to more closely follow the source messages produced by pivot interpreters than direct interpreters adhere to the messages of source speakers. More frequent segmentation, although alleviating interpreters' memory load, makes it more difficult for interpreters to obtain a holistic picture of the source message. The above finding may also result from difficulties such as relay interpreters' lack of access to features of the speaker's spontaneous discourse and relative lack of familiarity with the source language and culture.

Figure 5. An example of A-repair (from the Egyptian speech)



Notes: Tier1 (ST) is the transcription of source speech; Tier 2 (TT) is the transcription of interpreting output, synchronized with the time line; Tier 3 (Tra) is the English back translation of Tier 2; Tier 4 (DF) marks the types of disfluency.

The segments of the annotation above exemplify an A-repair for ambiguity. The relay interpreter originally produced the rendition "We believe to have responsibility to protect Arabian countries' safety to respond to external attempts," then repaired the reparandum "external attempts" for conceptual clarification, giving "attempts to intervene in it" (English back translation).

5. Conclusion

This study offers a systematic analysis of the texts of and types of disfluency in the SI output of relay and non-relay interpreters with the UN through investigation of a corpus of naturalistic data.

Annotation statistics show that the relay interpreting output is shorter than the non-relay output, and shows a dispersive pattern. It seems that relay interpreters produce less output containing more unfilled pauses. Analysis of disfluency types (repairs, filled pauses, repetition, and false starts) reveals that filled pauses account for the majority of the disfluencies observed, followed by repairs. Filled pauses and false starts account for a larger proportion and repairs constitute a smaller proportion of the relay interpreting output than the non-relay output, which may suggest that relay interpreters face more uncertainties than their non-relay counterparts in the message comprehension and planning stages, which are subject to more severe time constraints.

Probing into repair types in general, E-repairs constitute the majority of the repairs undertaken, followed by A-repairs and D-repairs. A large proportion of lexical E-repairs and A-repairs performed to correct ambiguity may reflect interpreters' adherence to the norms of accuracy and clarity. Comparison of repair types between the relay and non-relay SI outputs reveals that the relay output shows fewer lexical and phonetic E-repairs and more A-repairs addressing ambiguity, syntactic E-repairs, and D-repairs. This indicates that relay interpreters tend to make fewer errors than their direct counterparts at the syllabic and lexical levels, but face more uncertainties at the phrasal and syntactic levels, which may be the result of more frequent segmentation in relay vs. non-relay interpreting.

In sum, findings suggest that relay interpreting may affect the output of relay takers from multiple perspectives. It is important for interpreters and trainee interpreters to understand the process of relay interpreting and learn how to respond to the challenges it presents, both as a pivot and as a relay taker. Currently very few SI classes incorporate the relay element in the training of interpreters. It is understandable because most SI classes focus on one language pair. Training in relay is ideal at institutes that offer multiple language pairs. Relay is inevitable in the professional setting, interpreters should be aware of the needs of their colleagues and the constraints of relay, for instance the importance of being concise toward the end in order to release the occupied channel to minimize delays. Relay interpreting also deserves more weight in the development of professional interpreting standards. Research should also focus on relay to improve our understanding.

This corpus-based study of authentic data on disfluencies in relay and non-relay SI output helps to substantiate the findings of previous research on relay interpreting and disfluency in interpreting and provides valuable recommendations for improving the quality of relay interpreting. The data are representative and selected from a large source. Potential variables, particularly accent and input rate, are reasonably well controlled. However, the corpus used in this study is still fairly small. A larger corpus including multiple language pairs would generate more solid findings and facilitate more scientific statistical analysis. In addition, this study focuses on analyzing interpreter output. Considering the prosodic and semantic features of source speeches (or the output of pivot interpreters in cases of relay) may yield more multifaceted findings. Such features deserve further investigation.

Notes：

① https://tla.mpi.nl/tools/tla-tools/elan

② http://www.mpi.nl/corpus/html/elan/index.html

References

Bühler, Hildergund (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua, 5*(4), 231-235.

Bakti, Maria (2009). *Speech disfluencies in simultaneous interpretation*. Paper presented at the CETRA Research Seminar in Translation Studies 2008, Leuven.

Bendazzoli, Claudio, Sandrelli, Annalisa, & Russo, Mariachiara (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). London/New York: Continuum.

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

Cenoz, Jasone (1998). *Pauses and Communication Strategies in Second Language Speech*. Retrieved from https://eric.ed.gov/?id=ED426630

Chernov, Ghelly V. (2004). *Inference and Anticipation in Simultaneous Interpreting: A probability-prediction model*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Dechert, Hans Wilhelm, & Raupach, Manfred (Eds.). (1980). *Towards a cross-linguistic assessment of speech production*. Frankfurt: Peter Lang.

Duez, Danielle (1982). Silent and non-silent pauses in three speech styles. *Language and speech, 25*(1), 11-28.

Foster, Pauline, Tonkyn, Alan, & Wigglesworth, Gillian (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354-375.

Gósy, Mária (2005). *Pszicholingvisztika* . Budapest: Osiris Kiadó.

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

Gile, Daniel (2009). *Basic concepts and models for interpreter and translator training*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Gile, Daniel (2011). Errors, omissions and infelicities in broadcast interpreting:

Preliminary fndings from a case study. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and strategies of process research: Integrative approaches in translation studies* (pp. 201-218). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Goldman-Eisler, Frieda (1958). The predictability of words in context and the length of pauses in speech. *Language and speech, 1*(3), 226-231.

Hargrove, Patricia M., & McGarr, Nancy S. (1994). *Prosody management of communication disorders*. Sandiego and California: Singular Publishing Group Inc.

Kormos, Judit (1999). Monitoring and self-repair in L2. *Language learning, 49*(2), 303-342.

Kurz, Ingrid (1993). Conference interpretation: Expectations of different user groups. *The Interpreter's Newsletter*(5), 13-21.

Levelt, Willem (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104.

Mackintosh, Jennifer (1983). *Relay interpretation: An exploratory study.* (MA thesis), University of London.

Mead, Peter (2000). Control of pauses by trainee interpreters in their A and B languages. *The Interpreters' Newsletter, 10*, 89-102.

Mead, Peter (2005). Methodological issues in the study of interpreters' fluency. *The Interpreters' Newsletter, 13*, 39-63.

Mikkelson, Holly (1999). Relay interpreting: A solution for languages of limited diffusion? *The translator, 5*(2), 361-380.

Moser-Mercer, Barbara, Künzli, Alexander, & Korac, Marina (1998). Prolonged turns in interpreting: Effects on quality, physiological and psychological stress (Pilot study). *Interpreting, 3*(1), 47-64.

Moser, Peter (1996). Expectations of users of conference interpretation. *Interpreting, 1*(2), 145-178.

Pöchhacker, Franz (1995). "Clinton speaks German": A case study of live broadcast simultaneous interpreting. In M. Snell-Hornby, Z. Jettmarová, & K. Kaindl (Eds.), *Translation as intercultural communication: selected papers from the EST Congress, Prague 1995* (Vol. 20, pp. 207-216). Amsterdam/Philadelphia:

John Benjamins Publishing Company.

Pöchhacker, Franz (2012). *Interpreting quality: Global professional standards.* Paper presented at the Interpreting in the Age of Globalization: Proceedings of the 8th National Conference and International Forum on Interpreting, Beijing.

Petite, Christelle (2005). Evidence of repair mechanisms in simultaneous interpreting: A corpus-based analysis. *Interpreting, 7*(1), 27-49.

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

Postma, Albert (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*(2000), 97-131.

Pradas Macías, Macarena (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting, 8*(1), 25-43.

Riggenbach, Heidi (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14*(4), 423-441.

Schnadt, Michael J., & Corley, Martin (2006). *The influence of lexical, conceptual and planning based factors on disfluency production.* Paper presented at the Proceedings of the 28th Annual Meeting of the Cognitive Science Society, Mahwah.

Seeber, Kilian G. (2011). Cognitive load in simultaneous interpreting: Existing theories-new models. *Interpreting, 13*(2), 176-204.

Seleskovitch, Danica, & Lederer, Marianne (1989a). The problems of relay. In D. Seleskovitch & M. Lederer (Eds.), *A systematic approach to teaching interpretation* (pp. 173-192). Luxembourg: Didier.

Seleskovitch, Danica, & Lederer, Marianne (1989b). *A systematic aproach to teaching interpretation*. Paris: European Communities.

Setton, Robin (1999). *Simultaneous interpretation: A cognitive-pragmatic analysis*. Amsterdam/Philadelphia: John Benjamins Publishing.

Setton, Robin, & Dawrant, Andrew (2016). *Conference interpreting: A complete course*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Shlesinger, Miriam (2010). Relay interpreting. In Y. Gambier & L. v. Doorslaer (Eds.), *Handbook of Translation Studies* (Vol. 1, pp. 276-278).

Amsterdam/Philadelphia: John Benjamins Publishing Company.
Shreve, Gregory M., Lacruz, Isabel, & Angelone, Erik (2011). Sight translation and speech disfluency: Performance analysis as a window to cognitive translation process. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and strategies of process research: Integrative approaches in translation studies* (Vol. 93-120). Amsterdam/Philadelphia: John Benjamins Publishing Company.
Simone, Raffaele (1995). *Fondamenti di linguistica*. Bari: Laterza.
Tissi, Benedetta (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter, 10*, 103-127.

Résumé

La recherche, qui repose sur le corpus, mène une étude sur l'effet de l'interprétation en relais dans l'Assemblée générale des Nations Unies en comparant les caractéristiques de disfluence entre l' l'interprétation simultanée en relais et celle en non-relais. Voici les découvertes: (1) L'output de l'interprétation en relais est plus court et plus décentralisé que celui de l'interprétation en non-relais. (2) La pause sonore est le type le plus commun des disfluences. Et (3) En comparaison de l'output de l'interprétation simultanée en non-relais, il y a moins d' E-réparations lexicales et phonétiques, et plus d'A-réparations pour les équivoques, plus d' E-réparations syntatiques et plus de D-réparations dans l'output de l'interprétation en relais. Les résultats montrent que l' utilisation de l'interprétation en relais peut donner de l'influence à l'output des interprètes.

Mots-clés: interprétation en relais; disfluence; réparations; pause sonore; Nations Unies

Author's addresses

Andrew K.F. Cheung
Department of Chinese & Bilingual Studies
Hong Kong Polytechnic University
Kowloon

Shuxian Emma Song
School of Translation Studies
No. 80, Yantai Road,
276826 Rizhao, Shandong

Hong Kong
andrewkfcheung@gmail.com

China
emmashuxian@163.com