Available online at www.sciencedirect.com

ScienceDirect

Lingua 291 (2023) 103562

Lingua

www.elsevier.com/locate/lingua

A corpus-based analysis of the effect of syntactic complexity on disfluency in consecutive interpreting

Mingxia Shen [a,b], Yumeng Lin [a], Qianxi Lv [c], Junying Liang [a,*]

[a] Department of Linguistics, Zhejiang University, Hangzhou, PR China
[b] School of Foreign Languages, Hangzhou City University, Hangzhou, PR China
[c] School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, PR China

Received 13 July 2022; revised 1 June 2023; accepted in revised form 1 June 2023; available online 19 June 2023

Abstract

This study investigates the effect of input and output syntactic complexity on disfluency based on the corpus of press conference interpreting. In line with widespread practice in quantitative linguistics, mean dependency distance has been taken as the metric here for quantifying syntactic complexity. As the occurrence of disfluency is count data, this study uses the Poisson regression model to evaluate the effect of input and output syntactic complexity on disfluency occurrences, and on the variation of such occurrences with different types of reformulation methods. Our results show that occurrence of disfluency can be predicted by both the input and output syntactic complexity as quantified by the mean dependency distance, and that when the reformulation method of “divide” is used, the occurrence of disfluency does not increase significantly in the output even when the output sentence has a higher mean dependency distance. The findings reveal how input and output syntactic complexity predicts disfluency, and how reformulation methods interact with syntactic complexity to moderate cognitive load.

Keywords: Disfluency in consecutive interpreting; Syntactic complexity; Dependency distance; Reformulation; Corpus-based analysis

1. INTRODUCTION

“Fluency” is a complex concept for which there is no consensus about how it should be defined and measured (Guillot, 1999). Conversely, many researchers have argued that disfluency is objective and measurable to a certain extent and researchers in interpreting studies agree that characteristics such as pauses, hesitations, repetitions, and corrections are indicators of interpreting disfluency (Tissi, 2000; Cecot, 2001; Mead, 2005; Rennert, 2010; Wang, 2016). Tissi (2000) proposes a disfluency classification for the context of interpretation comprising pauses, parenthetic remarks, and interruptions. Cecot (2001, p. 70) develops Tissi’s classification further by dividing disfluency into “filled

* Corresponding author at: Department of Linguistics, Zhejiang University, Hangzhou 310058, PR China.
E-mail addresses: shenmx@zucc.edu.cn (M. Shen), linyumeng58@foxmail.com (Y. Lin), lvqianxi@sjtu.edu.cn (Q. Lv), jyleung@zju.edu.cn (J. Liang).

https://doi.org/10.1016/j.lingua.2023.103562
0016-7037/

pauses, glottal clicks, audible breaths, vowel and consonant lengthening, parenthetical sentences, and utterance interruptions" such as repetitions, restructurings, and restarts. Despite the lack of uniform classification, the concept of disfluency offers a window into the cognitive processing involved in interpreting (Tissi, 2000; Piccaluga et al., 2005; Gile, 2008; Plevoets and Defrancq, 2016). Investigating disfluency increases our understanding of interpreters' problems in comprehending given speech properties and reformulating the speech features that place a load on interpreters' cognitive resources.

To capture the nature of disfluency, researchers have been striving to identify its features and causes, either by focusing on trainees in interpreting, or by comparing experienced interpreters with novices or even non-interpreters (Mead, 2002; Piccaluga et al., 2005; Xu, 2010; Wang and Li, 2015). Xu's (2010) study reveals that most pauses made by trainee interpreters are triggered by their rearrangement of information, retrieval of expressions in the target language, and repair of interpretations. Mead (2002) suggests that interpreters' general linguistic knowledge and skills, as well as coping tactics and strategies, can help reduce pauses in interpreting. It is also observed that with the increase of interpreting expertise, pauses often decrease in number, and once interpreters' language proficiency improves, their pauses become shorter (Piccaluga et al., 2005). This last observation echoes the findings of Wang and Li (2015), namely that compared with student interpreters, pauses produced by professional interpreters are shorter in duration and smaller in number. The above researches show the importance of language proficiency and interpreting expertise in the control of pause number and duration.

Interpreters' capacity for working memory has also been found to be a key factor affecting disfluency (Lin et al., 2018; Shen and Liang, 2015). Using scores on reading span tasks to measure working memory, Macnamara and Conway (2016) show that interpreters' capacity for working memory is highly predictive of their interpreting performance. Hence, high capacity for working memory is often regarded as a cognitive advantage and a predictor of successful interpreting learning. Lin et al. (2018) also find that capacity for working memory of interpreting trainees can better predict their interpreting fluency than their language proficiency.

Working memory in the interpreting context refers to interpreters' memory capacity to handle the cognitive load in interpretation. Once cognitive load increases, there is a tendency for the quality of performance to decrease (Wilson and Eggemeier, 2006). Studies on disfluency should therefore examine not only the interpreters' perspectives and competence, but also the difficulty involved in interpreting. Taking this point into consideration, text features and properties of speech delivery have been examined to explore the cognitive load imposed on interpretation and the corresponding impact on disfluency.

Gile (2008) conducts a local analysis of text features at the sentence level and suggests that information density in the sentence can affect cognitive load in simultaneous interpreting. Specifically, sentence endings with high information density require more processing effort. If information in the previous sentence is not processed completely, the resulting cognitive load interferes with incoming information analysis and other concurrent tasks. In addition to the cognitive load caused by text features at the sentence level, Plevoets and Defrancq (2016, p. 202) assess cognitive load by measuring the frequency of *uh(m)*, a type of disfluency as well as an indicator of cognitive effort, and they examine the effect with respect to four variables, namely "delivery rate, lexical density, percentage of numerals, and average sentence length". Their findings suggest that the delivery rate of the original speaker can affect the number of *uh(m)*s, and that the higher the lexical density of the output sentence is, the more *uh(m)*s the interpreter will produce.

Investigations into disfluency have been carried out either from the perspective of interpreters' competence, such as in terms of working memory capacity, interpreting expertise as well as language proficiency, or based on cognitive load caused by different features of information density and properties of speech delivery. However, research on disfluency in relation to linguistic properties has been sparse, despite the promising findings from Plevoets and Defrancq (2016) in examining the effect of lexical density and average sentence length as linguistic features of source and target speeches on disfluency. As sentence comprehension includes the processing of both semantics and syntax (Hagoort, 2000), syntactic features exhibited by sentences are of major relevance to their processing in working memory during sentence comprehension (Friederici and Bornkessel, 2003). To comprehend syntactically complex sentences, additional working memory resources are required and more cognitive effort is needed. For example, converting the syntactic structure of one language such as subject–object–verb into a different structure such as subject–verb–object of another language in simultaneous interpreting generates more cognitive load than if similar syntactic structures are involved (Seeber, 2011). Compared with simple and short sentences, syntactically complex sentences with embedded relative clauses also cause more disfluencies in sight interpretation (Shreve and Lacruz, 2011).

To assess the cognitive load imposed on interpreters, Jiang and Jiang (2020) investigate the impact of the maximum dependency distance (max DD) in the source text on interpreters' disfluency in an experimental setting. Heringer et al. (1980) first propose the concept of dependency distance (DD), which is later introduced by Hudson (1995, p. 16) as a technical term, referring to "the distance between words and their parents, measured in terms of intervening words". Such syntactic relation "between words and their parents" is then labelled as "dependency relation" (Liu, 2008;

Wang and Liu, 2017; Liang et al., 2017) between two words in a sentence, with “one as the governor and the other as the dependent” (Wang and Liu, 2017, p. 135). For a word string $W_1...W_i...W_n$ in a sentence, if W_a and W_b form a dependency relation in syntax, the DD between these two words is the absolute value of the quotient of a minus b. The maximum dependency distance is “the maximum absolute value of DD among all the dependency relations in a sentence” (Jiang and Jiang, 2020, p 6.). To clarify this metric, we cite an example from Jiang and Jiang (2020) in Fig. 1, where the number below each word identifies its linear position. From this example, we can see the max DD is between *lost* and *toy*, i.e. |3-5| = 2. The results of Jiang and Jiang (2020) suggest that sentences with high values of maximum DD may cause the interpreter to yield significantly more disfluencies in sight interpretation and hence more cognitive effort is required to interpret sentences with higher syntactic complexity.

Still, questions remain regarding the impact exerted by syntactic complexity. Research in this domain has generally been conducted in experimental settings, while natural settings, such as on-site conference interpreting, have rarely been explored. Jiang and Jiang (2020, p. 13) similarly acknowledge this research gap and suggest that “corpus-based study may serve as a significant alternative, based on large-scale authentic data in real-life settings.” So far, Plevoets and Defrancq (2018b) have investigated the effect of syntactic complexity on disfluency based on the corpus. Plevoets and Defrancq (2018b, p. 3) trace the frequency of *uh(m)* in real-life French–Dutch simultaneous interpreting and measure syntactic complexity by “the number of subordinate clauses per utterance length”. Their results reveal that the syntactic complexity of the original speech produces a borderline effect, while no significant effect is found from the syntactic complexity of the interpreted speech.

In an effort to examine the effect of syntactic complexity on disfluency more precisely, the current study takes the mean DD of all the dependencies in a sentence as the metric for quantifying syntactic complexity. Gibson (1998, 2000) claimed that keeping track of long incomplete dependencies would impose a great burden on one’s memory, and that the accumulated memory load in turn could cause difficulty in sentence comprehension. Grodner and Gibson (2005, p. 261) also reported that “difficulty associated with integrating a new input item is heavily determined by the amount of lexical material intervening between the input item and the site of its target dependents.” According to Liu (2008), the longer the mean DD among all the dependencies in a sentence (hereafter MDD), the harder it is to comprehend the sentence. In this respect, MDD has been used widely to measure syntactic complexity and to explain the cognitive load involved in sentence comprehension and production, and is our choice of indicator of syntactic complexity for this study. Moreover, as researchers have mainly focused on disfluency in simultaneous or sight interpreting (Piccaluga et al., 2005; Plevoets and Defrancq, 2016; Jiang and Jiang, 2020), this study uses a corpus of on-site consecutive interpreting, another cognitively constrained mode of interpreting, to examine the effect of syntactic complexity on disfluency.

Jones (2008) claims that in simultaneous interpreting, long and/or complicated sentences are broken down into easier and shorter sentences, and that relative and subordinate clauses can be shifted around within a sentence. In consecutive interpreting, interpreters take notes while listening to and analyzing the incoming information. When the speaker pauses, the interpreter reformulates the message using their notes as well as memory. Do interpreters employ different reformulation methods to regulate the cognitive load?

Taking into account the points above, the current study attempts to answer the following main questions:

(1) Does input syntactic complexity affect the disfluency of on-site consecutive interpreting?
(2) Does the syntactic complexity of the output affect interpreters’ disfluency?
(3) Do different reformulation methods interact with syntactic complexity and influence interpreting disfluency?

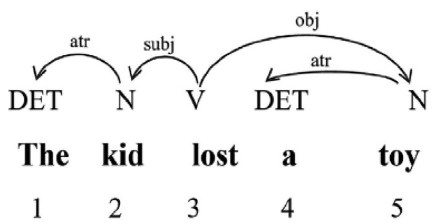

Fig. 1. Sample analysis of dependency structure and dependency distance (Jiang and Jiang, 2020).

2. METHODOLOGY

To investigate the effect of syntactic complexity and different reformulation methods on disfluency in on-site consecutive interpretation, we built a parallel corpus using original speeches and corresponding interpretations from press conferences where Chinese Premiers meet the press.

2.1. Corpus

Five conferences were selected for the building of the corpus. Consecutive interpreting was provided for each conference by a different professional interpreter. In each conference, the Premier makes a few preliminary remarks, and then takes questions from the press and gives his answer. To confine the consecutive interpretation considered in this study to that from Chinese to English, only the Premier's opening remarks, the Premier's answers to questions, and their corresponding interpretations were excerpted. Although the texts of both the original speeches and their English interpretations are available online, we checked all the texts against the conference videos to ensure accuracy when tagging the disfluencies in interpreting. The full stop of each sentence in the texts was also checked based on the videos to ensure the end of an utterance as a sentence, where the speaker or the interpreter dropped his/her voice and then made a short pause (Jones, 2008). Table 1 displays the basic information of the corpus.

2.2. Instruments

2.2.1. Disfluency

This study adopted the evaluation and categorization schema of disfluent phenomena developed by Dayter (2021). We identified and coded six types of disfluency: filled pause (e.g. *uh, mm,* and *ah*), repetition (the interpreter repeats part of the utterance), replacement (the interpreter replaces the already articulated utterance with other items), false start (the interpreter interrupts an unfinished sentence and starts another sentence), mispronunciation (the interpreter mispronounces or truncates words), and non-fluent (non-juncture) silent pause (the interpreter pauses without fillers and the pause does not occur at grammatical junctures).

According to Cenoz (1998, p. 2), filled pauses are "interruptions of speech flow by non-lexical sounds, such as *ah*, *mm*, *er*, and *uh*", and are usually operational indicators of speech disfluency (Tissi, 2000; Watanabe et al., 2008; Dayter, 2021). The occurrence of filled pauses can often be attributed to an increase in informational or cognitive load (Plevoets and Defrancq, 2016, 2018a). False starts, as suggested by Tissi (2000, p. 119), "are not an individual characteristic but rather the result of difficulties in coping with the original text." In comparison, mispronounced words and truncated words result from errors in articulation or slips of the tongue due to the interpreters' stressed mental state. Fluent (juncture) silent pauses, according to Kenny (1996), occur more often at syntactic boundaries. In contrast, non-fluent (non-juncture) silent pauses do not occur at grammatical junctures and are judged to be abnormal. Put differently, non-fluent silent pauses can be identified and differentiated from fluent silent pauses according to their syntactic position. Previous studies (Chambers, 1997; Ahrens, 2004) have also claimed that if pauses occur between sentences or clauses, they belong to syntactic pauses; and when they occur where they are not normally required, they are identified as non-syntactic. Furthermore, Rennert (2010, p. 103) argues that "the position has a strong impact on the minimum length necessary for a pause to be noticed." Wang and Li (2015, p. 7) also examine different syntactic positions of pauses and claim that pauses "between the subject and the predicate verb, between the predicate verb and the object, and inside phrases and between parallel structures" are non-juncture pauses. In the present study, we adopt this criterion to identify non-fluent (non-juncture) silent pauses.

Two of the authors of the current study identified and coded disfluencies in the transcribed interpreting texts using the working definitions of disfluencies described above. Non-fluent (non-juncture) silent pauses were marked as "..." in the sentence and annotated as [SP] at the end of the sentence, while filled pauses were marked as "*uh*" in the sentence and annotated as [FP] at the end of the sentence. Disfluencies of the other types were marked in italics in the sentence and annotated as [RP] (repetition), [RL] (replacement), [FS] (false start), and [MP] (mispronunciation) at the end of the sentence. After becoming familiar with the annotation rules, the two authors coded a conference separately and then

Table 1
Basic information of the corpus.

Interpreting Type	Text Type	Number of Texts	Number of Sentences		Total Word Count (Tokens)	
Consecutive Interpreting	Source Text	5	Chinese	1,090	Chinese	22,073
	Target Text	5	English	1,650	English	30,653

discussed their results. The main discrepancy arose regarding non-fluent silent pauses between parallel structures. For instance, there is a short pause in the sentence "*Third, such cooperation ...there are already important agreements concerning energy cooperation.*" This pause could have been classified as a fluent (juncture) pause as it occurs before a sentence, or as a non-fluent (non-juncture) pause between parallel structures because "*such cooperation*" and "*there are*" both serve as the beginning of a sentence. However, as "*such cooperation*" is a false start and the pause between this false start and the new sentence is quite noticeable, the two coders ultimately agreed to classify this and other similar pauses as non-fluent (non-juncture) pauses between parallel structures. They then re-coded the conference and coded the remaining conferences. The agreement between the coders was 95.5%, and the intercoder reliability Cohen κ was 0.893 ($p < 0.001$), which suggests substantial agreement. Following their coding, they discussed their differences until they reached a consensus. Table 2 shows the details of their disfluency coding in the corpus.

2.2.2. Syntactic complexity and dependency distance (DD)

Computational and cognitive linguists have for decades searched for a metric to measure syntactic complexity (Yngve, 1960; Hawkins, 1994). Hawkins (1994) hypothesized that there is a human preference for linear order and that syntactic complexity is relevant to word order. Hawkins (2004, p. 31) later claimed that "the human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed". Linguists have widely investigated this claim about linear order and syntactic complexity (Gibson, 1998; Liu, 2008) and have explored the relation between the two so as to develop theories and formulas relating to DD to measure syntactic complexity (Liu, 2009; Liu et al., 2017).

Liu (2008, p. 165) proposes a formular to calculate MDD for the measurement of syntactic complexity of a particular sentence where "n is the number of words in a sentence and DD_i is the DD of the i^{th} syntactic link in the sentence":

$$\text{MDD (the sentence)} = \frac{1}{n-1}\sum_{i=1}^{n-1}|DD_i|$$

From the formula, we can see that MDD accounts for the memory burden caused by different DDs (Liu, 2008), which is different from prior studies where the number of subordinate clauses has been used as an indicator to measure syntactic complexity (Setton, 1999; Plevoets and Defrancq, 2018b), although subordinate constructions have served as an effective predictor of processing cost in both first language and second language research (Gordon and Luper, 1989; Norris and Ortega, 2009; Osborne, 2011). Empirical studies have employed MDD in syntactical structure analysis for different languages and have proved that processing load increases with MDD and thus MDD works as an effective indicator of cognitive load (Hsiao and Gibson, 2003; Grodner and Gibson, 2005). For example, MDD is used in psycholinguistic work, where object-relative sentences are deemed to cause more cognitive load than subject-relative sentences (King and Just, 1991; Jay, 2004). By using MDD, it is possible to illustrate that more processing load is caused by object-relative sentences than by subject-relative sentences. To show how MDD is used in this context, we consider the following sentences:

a) The student who criticized the teacher failed the exam. (subject-relative sentence)
b) The student who the teacher criticized failed the exam. (object-relative sentence)

Table 2
Coding of disfluency in the corpus.

Annotation	Meaning	Example[1]
[SP]	Non-fluent silent pause	"We have been successful in avoiding major ups and downs in the economy, preventing... excessive price hikes and keeping prices at a stable level and maintaining stable and fairly rapid economic growth." [SP]
[FP]	Filled pause	"Last year indeed *uh* has been the most *uh* extraordinary for China." [FP][FP]
[RP]	Repetition	"And consultations and negotiations will be *conducted, conducted* on an equal footing." [RP]
[RL]	Replacement	"First of all, let me *set send* my greeting to the 23 million compatriots in Taiwan." [RL]
[FS]	False start	"Third, *such cooperation* ...there are already important agreements concerning energy cooperation." [FS][SP]
[MP]	Mispronunciation	"Fourth, we should enhance infrastructure for the securities market, centring on putting *an por* appropriate system in place." [MP]

[1]The examples are taken from the interpretation transcripts of the questions and answers in Chinese Premiers' Press Conferences which are available online.

If syntactic complexity was quantified by the number of subordinate clauses, the two sentences would have the same syntactic complexity, as both have one subordinate clause. However, by using MDD as the metric, these two sentences entail different syntactic complexity. The illustrations in Fig. 2 provide a comparison in this regard:

Numbers below the words indicate the specific dependency distances between each dependent word and its governor. The MDD of the second sentence is 2.25 (=1 + 5 + 3 + 1 + 1 + 4 + 1 + 2 / 8), while the first sentence has its MDD of 1.875 (=1 + 5 + 1 + 2 + 1 + 2 + 1 + 2 / 8) according to the formula. Therefore, the MDD of the second sentence is higher than that of the first, suggesting that more cognitive load is generated in processing the second sentence in comparison to the first. Given the difference in results provided by the two measures of syntactic complexity in this pair of example sentences, we assume that MDD can be used as a more precise metric.

In the current study, the Chinese sentences and the English sentences were all imported into Stanford Parser, with which their syntactic structures and dependency relationships were generated. The absolute DDs of dependent words from their governors in a sentence were also presented. The parsed results were then programmed into Microsoft Excel for the calculation of MDD.

2.2.3. Different reformulation methods

In consecutive interpreting, interpreters can reverse the order of two sentences from the speaker, merge two sentences into one, or, conversely, divide a long sentence into several shorter sentences (Jones, 2008). For this corpus, we found that the output, when aligning the output with the input, does not always correspond to the input sentence-by-sentence. On the contrary, we identified four output reformulation methods: 1) divide (content in one input sentence is reformulated and rendered as two or more output sentences), 2) correspond (content in one input sentence is reformulated and rendered as one output sentence), 3) merge (content in two or more input sentences is reformulated and rendered as one output sentence), and 4) reconstruct (usually, content in two input sentences is reformulated and rendered as three output sentences or vice versa; the relative and subordinate clauses in one sentence are shifted to another).

After becoming familiar with the different reformulation methods, two authors coded a conference separately by checking the output sentences against the input sentences, and they then discussed their results. Difficulty arose regarding whether certain reformulations should be considered as using the correspond or reconstruct method. For example, the content in two input sentences were reformulated and rendered into the same number of output sentences, but part of the information in one input sentence was shifted into another sentence in the output. Upon discussion, the two coders agreed to classify such reformulation methods as reconstruct and re-coded the conference. They then coded the remaining conferences. The agreement between the coders was 97.2%, and the intercoder reliability Cohen κ was 0.911 ($p < 0.001$), which suggests substantial agreement. Following their coding, they discussed their differences until they reached a consensus.

3. DATA ANALYSIS AND RESULTS

We adopted Poisson regression models to predict the frequencies of disfluencies during interpreting. Poisson regression has been an established way to analyze count data in many fields, including linguistics, social sciences and others (McElreath, 2020; Zuur et al., 2009; Winter and Bürkner, 2021). In the current study, we used the parameter of Poisson distribution 'lambda λ' to specify the mean and variation of disfluency during interpreting. In addition, we included in the regression models the syntactic complexity via MDD of both source speech and interpreted utterances (as "fixed effects") and their variation at four output reformulation types (also as "fixed effects") to analyze the effect of syntactic

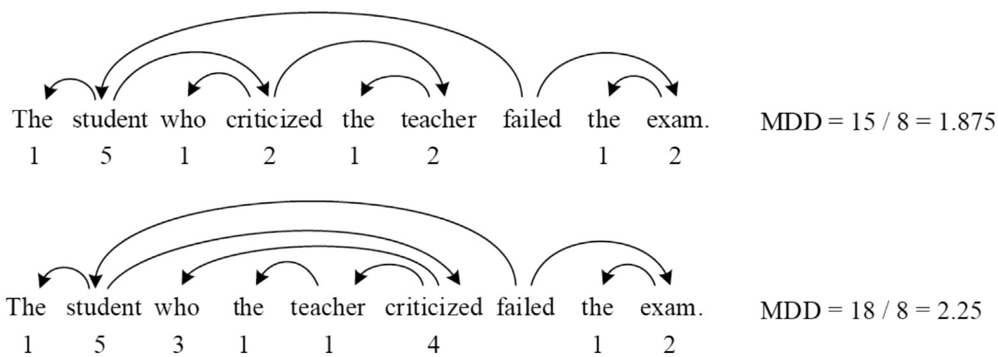

Fig. 2. MDDs of a subject-relative sentence and an object-relative sentence.

complexity on disfluency occurrences in different sentences ("random effects") levels (Nalborczyk et al., 2019; Winter and Bürkner, 2021).

In line with the methodology employed by Jing, Widmer, and Bickel (2021), we employed a Bayesian framework with Markov chain Monte Carlo (MCMC) algorithms to fit Poisson regression models. The purpose was to examine the relationship between disfluency counts and reformulation methods, as well as the syntactic complexity of both source and target utterances, including the interaction between these variables.

Various Poisson regression models were constructed and compared, incorporating different sets of variables as outlined in Table 3. As for the fixed effects, we estimated the coefficients for the MDD of the source text (MDD_CN), MDD of the target text (MDD_EN), and reformulation methods (REC), while accounting for sentence-level adjustments ("Ss_Num"). This model structure not only accounted for sentence-specific variation but also facilitated comparing the effect of MDD and reformulation methods on disfluency occurrences between the population and sentence levels.

To account for the relatively small sample size of the study and prevent potential model overfitting (Jing et al., 2021; McElreath, 2020; Nalborczyk et al., 2019), we allowed for varying slopes across types of reformulations to capture interactions between MDD and reformulation types. The models were fitted using Stan (Stan Development Team, 2020) in the brms (Bürkner, 2017) and cmdstanr packages (Gabry and Češnovar, 2020) within R (R Core Team, 2020). Weakly informative priors were chosen, specifying a normal distribution with mean 0 and standard deviation 0.5 for intercepts and slopes. Each model was run with two MCMC chains for 4,000 iterations, including a warm-up phase of 2,000 iterations. Convergence was assessed and thus the effective sample sizes (ESS) were above 100 and Rhat value below 1.05 for each parameter.

Additionally, we conducted leave-one-out cross-validation (LOO-CV) to assess potential overdispersion and determine whether Poisson regression or negative binomial regression should be utilized (Vehtari et al., 2017). The model's predictive performance on unseen data was evaluated with fixed model variance using the expected log pointwise predictive density (ELPD) and the unbiased sampling estimator of Pareto-smoothed importance sampling (Δelpdpsis) based on posterior log probabilities (Yao et al., 2017; Sivula et al., 2022; McElreath, 2020;). In addition, to analyze the raw elpd estimates, we employed model stacking as a means of comparing model performance. This approach entailed assigning weights to the models to collectively optimize prediction quality (Vehtari et al., 2020).

As can be seen in Table 4, the model of MDD of source text by REC performs slightly better in this particular case than the model of REC by MDD of target text and all other models. But the difference in predictive power is negligible ($\Delta elpd_{psis} = -0.10$) and associated with a comparatively larger standard error ($SE(\Delta elpd_{psis}) = 9.20$). Fig. 3 reports the posterior distributions of the model of MDD of source text by REC.

A full Poisson regression model and a negative binomial regression model were built, where all predictors (MDD of both source and target texts, and reformulation methods) and their mutual interactions were included. The model formula was:

DISFLUENCY_NO~REC*MDD_CN*MDD_EN+(0+REC|MDD_CN)+(0+REC|MDD_EN)+(1+REC|Ss_Num)

Table 3
Models for estimating disfluency frequency of different syntactic complexity (μ) using brms-style notation. DISFLUENCY_NO: the frequency of disfluency occurrence; REC: reformulation methods; MDD_CN: MDD of source text, MDD_EN: MDD of target text, and Ss_num: sentence number, respectively.

Name	Regression
intercept	DISFLUENCY_NO ~ 1 + (1 | Ss_Num)
MDD of source text	DISFLUENCY_NO ~ MDD_CN + (1 + MDD_CN | Ss_Num)
MDD of target text	DISFLUENCY_NO ~ MDD_EN + (1 + MDD_EN | Ss_Num)
Reformulation methods	DISFLUENCY_NO ~ REC4 + (1 + REC4|Ss_Num)
MDD of target text by reformulation methods	DISFLUENCY_NO ~ 1 + MDD_EN + (1 + MDD_EN | REC4) + (1 + MDD_EN | Ss_Num)
MDD of source text by reformulation methods	DISFLUENCY_NO ~ 1 + MDD_CN+(1 + MDD_CN|REC4) + (1 + MDD_CN|Ss_Num)
Reformulation methods with the exposure of MDD	DISFLUENCY_NO ~ REC4 + offset(log(MDD_CN)) + offset(log(MDD_EN)) + (1 + REC4 | Ss_Num)
Reformulation methods & log MDD	DISFLUENCY_NO ~ REC4 + log(MDD_CN) + log(MDD_EN) + (1 + REC4 | Ss_Num)
full model	DISFLUENCY_NO ~ REC4 + MDD_CN + MDD_EN + (0 + MDD_CN | REC4) + (0 + MDD_EN | REC4) + (1 + REC4 | Ss_Num)

Table 4
Model comparisons via approximate leave-one-out cross-validation and model stacking.

Model	$elpd_{psis}$	$\Delta elpd_{psis}$	$SE(\Delta elpd_{psis})$	Weights
MDD of source text by REC	−1216.80	0.00	0.00	0.428
REC by MDD of target text	−1216.90	−0.10	8.10	0.370
MDD of source text	−1218.30	−1.50	2.30	0.100
REC by MDD of source text	−1222.20	−5.40	6.90	0.002
REC	−1227.00	−10.20	7.20	0.000
MDD of target text	−1237.60	−20.80	6.90	0.000
MDD of target text by REC	−1237.80	−21.00	6.50	0.000
intercept	−1253.40	−36.60	6.50	0.000

(REC: reformulation methods)

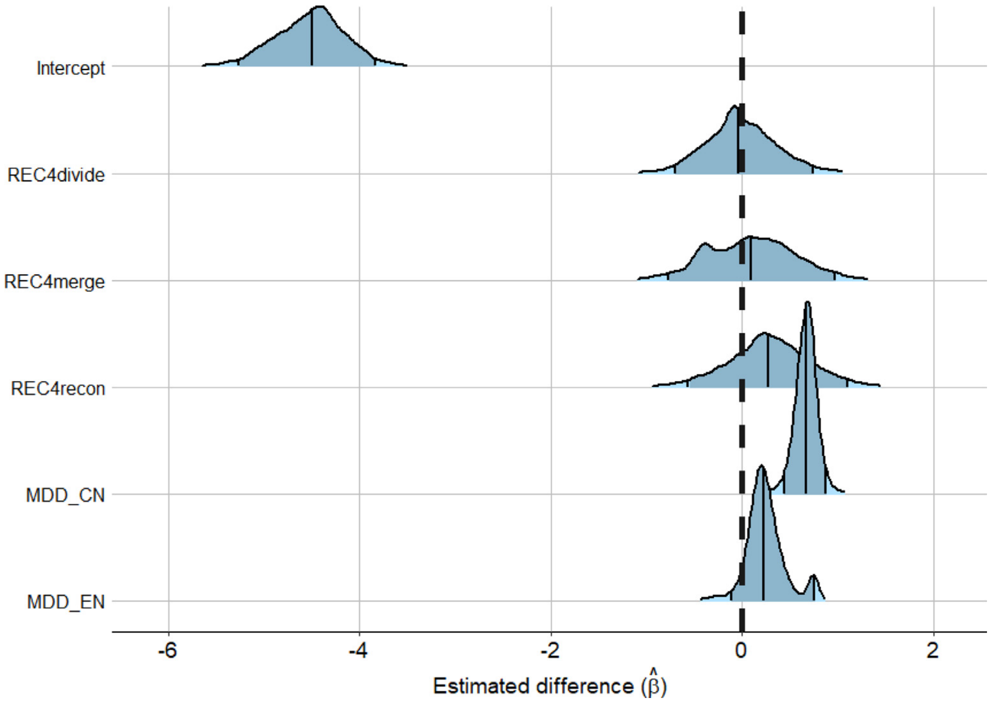

Fig. 3. Posterior distribution of parameters for disfluency occurrences (the model of MDD of source text by REC).

We further performed model comparisons via approximate leave-one-out cross-validation, as well as model stacking on the above models. The Poisson model performed better in predictive power ($\Delta elpd_{psis}$ = −3.10, $SE(\Delta elpd_{psis})$ = 0.8). In addition, the estimate of the shape parameter of the negative binomial regression model was 182.48 (95% of CI = [54.73, 443.83]). Table 5 lists the posterior distributions of the full Poisson regression model. The analyses revealed the effect of MDD of source ($\widehat{\beta}$ = 0.816, 95% CrI = [0.428, 1.20]) and target sentence ($\widehat{\beta}$ = 0.624, 95% CrI = [0.163, 1.09]) on the occurrence of disfluency, as the probabilities of these main effects being larger than zero were > 0.999 and 0.996, respectively (see Figs. 4 and 5). No effect of each reformulation method alone was found (see Table 5 for a full summary). The probability of the interaction between MDD of target text and the reformulation method of divide ($\widehat{\beta}$ = -0.264, 95% CrI = [−0.713, 0.169]) being lower than one was 0.874. As can be seen in Table 5, by using the reformulation method of "divide", the occurrence of disfluency does not increase significantly in the output even when the output sentence has a higher mean dependency distance (A negative correlation between disfluency frequency and the use of reformulation method of "divide" with the increase in MDD is indicated by the negative estimated value of the interaction).

Table 5
Summary of the Bayesian Poisson regression model.

Effect	$\hat{\beta}$	P($\hat{\beta}$)>	P($\hat{\beta}$)<	95% CI effect
Intercept	−5.200	<0.001	>0.999	[−6.51,−3.95]
REC = divide	−0.027	0.479	0.521	[−0.901,0.841]
REC = merge	−0.019	0.479	0.521	[−0.953,0.916]
REC = reconstruct	0.084	0.567	0.433	[−0.845,1.03]
MDD_CN	**0.816**	**>0.999**	**<0.001**	[0.428,1.2]
MDD_EN	**0.624**	**0.996**	**0.004**	[0.163,1.09]
REC = divide:MDD_CN	0.058	0.641	0.359	[−0.253,0.38]
REC = merge:MDD_CN	0.023	0.523	0.477	[−0.758,0.804]
REC = reconstruct:MDD_CN	0.111	0.645	0.355	[−0.489,0.687]
REC = divide:MDD_EN	−**0.264**	**0.126**	**0.874**	[−0.713,0.169]
REC = merge:MDD_EN	0.076	0.596	0.404	[−0.615,0.734]
REC = reconstruct:MDD_EN	0.056	0.564	0.436	[−0.681,0.773]
MDD_CN:MDD_EN	−0.077	0.116	0.884	[−0.204,0.0498]
REC = divide:MDD_CN:MDD_EN	0.013	0.583	0.417	[−0.107,0.134]
REC = merge:MDD_CN:MDD_EN	−0.076	0.316	0.684	[−0.394,0.228]
REC = reconstruct:MDD_CN:MDD_EN	−0.035	0.392	0.608	[−0.294,0.227]

Notes: The second column shows the estimated mean of regression coefficients. The third and fourth columns list the proportion of the posterior estimated that is higher than 0 and lower than 1, respectively. The 95% confidence intervals for the coefficients are given in the last column.

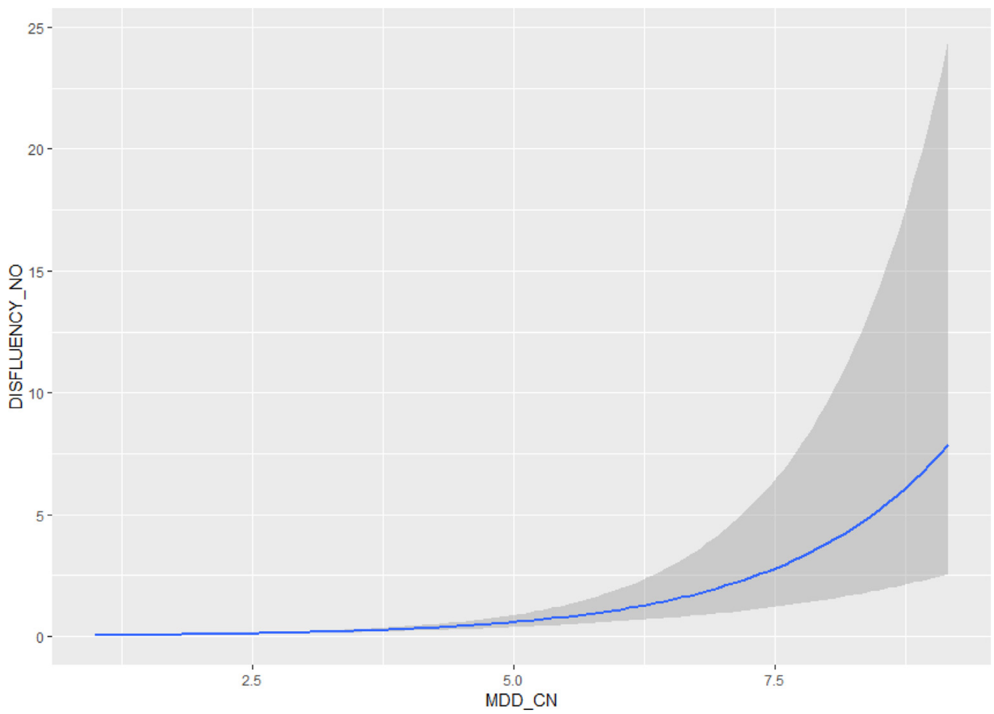

Fig. 4. Estimated values of the number of disfluency occurrence predicted depending on the MDD of the source text.

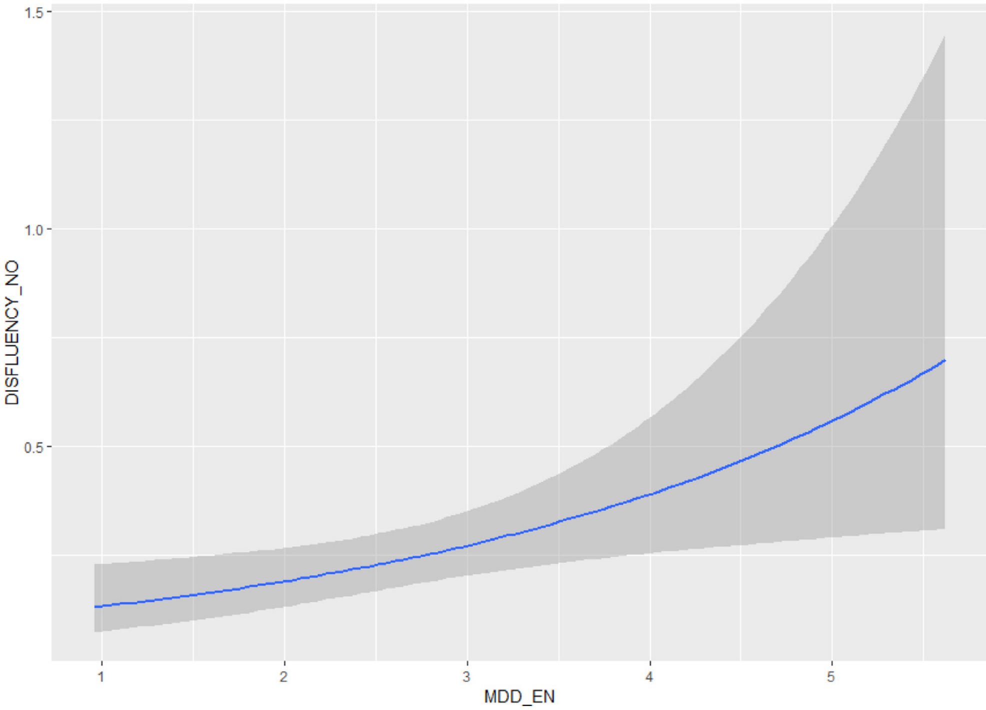

Fig. 5. Estimated values of the number of disfluency occurrence predicted depending on the MDD of the target text.

4. DISCUSSION

4.1. Effect of syntactic complexity on interpreting disfluency

The present study establishes that input syntactic complexity quantified in terms of MDD predicts disfluency in interpreting. As Liu's (2008, p. 159) study reveals that "linguistic complexity is a measure of cognitive difficulty of human language processing", once a sentence is more syntactically complex, comprehension of this sentence yields more cognitive load. Being an indicator of syntactic complexity, MDD reflects the cognitive cost of processing all the dependencies of a sentence, especially the cost on working memory. In consecutive interpretation, the interpreter needs to store dependent words continuously in working memory to fully comprehend and extract the deep syntactic information of a sentence from the original speech. It is only once the corresponding governors are heard that these dependencies can be resolved and the syntactic structure hence perceived and understood. Therefore, the sentence comprehension process is affected by the complexity of the whole sentence. According to Gile's (2015, p. 136) effort model, effort is made for "L (listening and analysis), M (memory), NP (note production), and C (coordination)" in the comprehension phase of consecutive interpreting, while in the reformulation phase, effort is made for "NR (note reading), SR (speech reconstruction from memory), P (production), and C (coordination)". Thus, it is self-evident that cognitive resources shall be assigned properly to each effort to ensure successful interpretation. Difficulties arising in any part of either phase may ignite chain reactions in the whole interpreting process. For instance, experiencing high cognitive pressure during comprehension may result in suboptimal note-taking, which may make reconstruction of the speech from memory and notes during the reformulation phase more difficult. Alternatively, the comprehension process may not have been completed due to higher syntactic complexity of the source speech, and the information in the notes can be incomplete; or the notes may have been taken too rapidly because more time and effort was expended in listening and analysis, causing the notes to become hard to read. These factors may lead to difficulty in the reformulation phase and disfluency.

While features of the source speech have been the main focus of previous studies in examining cognitive load involved in interpretation, this study also examines the target speech. Our results show that the target speech production also entails cognitive load and the syntactic complexity of the output quantified by MDD can also predict the occurrence of interpreting disfluency. This suggests that in consecutive interpreting, reformulating syntactically complex sentences imposes a higher cognitive load than reformulating simple sentences. In monolingual speaking activities, the speaker must plan an utterance when giving spontaneous speeches; certain working memory resources are required to this end (Gibson, 1998; Hawkins, 2004), and structures that contain long syntactic dependencies may be costlier as the speaker has to retain the element initiating that long dependency in short-term memory before articulating the governor or the relevant dependent word (Scontras et al., 2015). Thus, when interpreters reformulate syntactically more complex sentences in their output, they may experience more difficulties in their cognitive processing and output production, which may in turn result in more disfluencies. Our reasoning echoes the finding of Scontras et al. (2015, p. 559) for monolingual speech production, namely that “there is a cost associated with planning and uttering the more syntactically complex structures and this cost manifests in the form of longer durations and disfluencies.” In line with this finding, studies of monolingual activities have often found that syntactically simpler structures are chosen by their participants in picture description tasks (Gennari et al., 2012). These studies all corroborate our efforts here, suggesting that human language use and memory capacity develop synergistically, and that constraints on memory capacity drive the adaptation and change of language use (Gong and Shuai, 2015).

4.2. Cognitive load moderation through output reformulation in interpreting

The original sentence patterns were not strictly followed by the interpreters in our corpus. This fact has motivated us to probe the connection between cognitive load and output reformulation. In press conference interpretation examined in this study, the speaker’s utterance is spontaneous and very often not well-refined. As a result, a governor in a sentence may be quite far away from its dependent, thus requiring the interpreter to use more cognitive resources to fully comprehend the sentence. This situation may explain the rearrangement of information observed in most output sentences, in particular into syntactic structures with smaller MDDs. Such reformulation reduces the burden of comprehension for the listener as well as the difficulty of processing for the interpreter. Fig. 6 displays a sample input sentence and its interpretation using the reformulation method of “divide” and dependency structures in each sentence are labelled.

In the input Chinese sentence of this example, the DDs between the governor “是” (*shi* ‘was’) and its dependent word “难忘” (*nan wang* ‘unable to forget’), between “是” and “知青” (*zhi qing* ‘youth’), and between “难忘” and “岁月” (*sui yue* ‘time’) are all quite long. The sentence structure also includes the head-final noun phrase type (modifier + *de* + noun), which has been identified as a problem, triggering difficulty in second-language sight translation (Su and Li, 2019). In the output sentence, in comparison, the English translations of the dependent words “知青” and “岁月” are more adjacent to those of their governors “是” and “难忘,” respectively. The MDD of the input sentence is 2.833, while the MDDs of the two outputs are 1.545 and 1.923, respectively. Given the lower MDDs of the output sentences, the effect of the target speech MDD on the occurrence of disfluency becomes weaker.

As is shown in the above sample, interpreters reconstruct the content of the complex input in simplified syntactic structures in the output based on notes as well as memory, with mean dependency distance significantly reduced. It

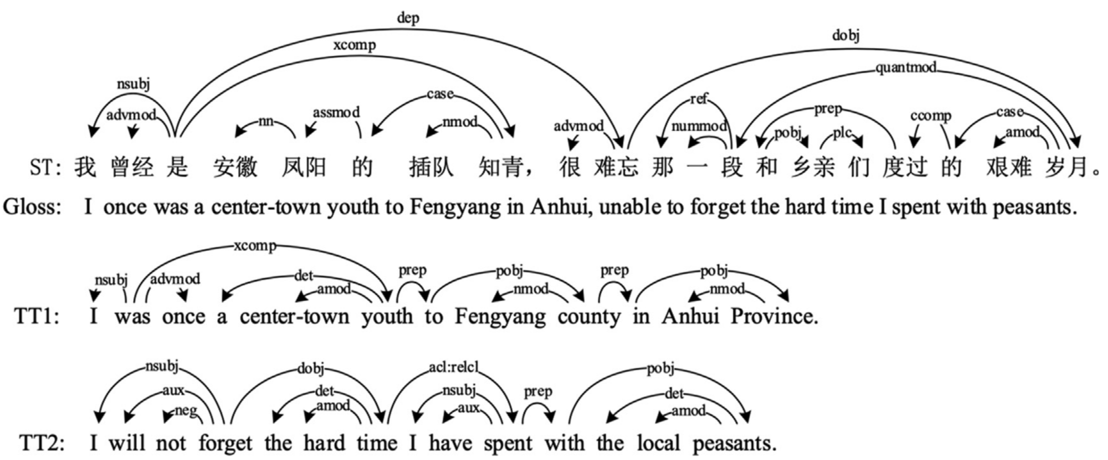

Fig. 6. Dependency structures of a sample input sentence and its output generated with “divide” (nsubj: nominal subject, ccomp: clausal complement, dobj: direct object, advmod: adverbial modifier, plc: plural form, nn: noun compound modifier, pobj: object of a preposition, dep: dependent, aux: auxiliary, nummod: numeric modifier, assmod: associative modifier, case: case, det: determiner, xcomp: open clausal complement, prep: preposition, nmod: noun modifier, amod: adjectival modifier, acl:relcl: clause relative clause, ref: referent, neg: negative).

is found that the syntactic complexity of the output speech quantified by MDD in consecutive interpreting is significantly lower than that of the input speech (Lv, 2020). This shows that interpreters moderate their cognitive load by dynamically reformulating information in different syntactic structures. Thus, information is reformulated and reordered, resulting in differences in the cognitive load that are brought about by source and target speeches (Riccardi, 1998; Plevoets and Defrancq, 2016).

5. CONCLUSION

This study aims to examine the effect of input and output syntactic complexity (indexed by MDD) on disfluency in consecutive interpreting, and to examine how reformulation methods interact with syntactic complexity to moderate cognitive load. The findings suggest that input and output syntactic complexity both serve as predictors of interpreting disfluency. This study has also demonstrated that interpreters reformulate the content of input speech into output sentences with lowered MDDs to moderate their cognitive load.

Despite the above findings, we are conscious of several limitations of this study. Firstly, to ensure non-fluent silent pauses are noticeable interruptions of the acoustic flow of interpretation, coders were instructed to base their evaluation on their perception and the syntactic positions of silent pauses, which could be prone to bias despite our utmost efforts to set the coding criteria to ensure consistency and to conduct discussion between the two coders until they reached consensus and then to re-code. Secondly, as different cognitive processes are involved in consecutive interpretation (including comprehension, note-taking, and memory in the input phase, as well as note-reading, speech reconstruction, and production in the output phase), failure in any process may lead to poor interpreting performance. The link is quite complex between difficult comprehension due to syntactic complexity and sub-optimal note-taking and then disfluencies in the output, but the effect of the quality of note-taking in this chain reaction is not investigated. Future efforts are required to examine the patterns and characteristics of note-taking, an intermediate link between comprehension and reformulation in consecutive interpreting, as these may vary according to the syntactic complexity of the input speech and may reveal other interesting findings.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENTS

This work is supported by the National Social Science Foundation of China, China (Grant No. 17BYY068), Zhejiang Provincial Philosophy and Social Sciences Planning Project, China (Grant No. 21NDJC181YB). Special thanks go to the reviewers for their insightful comments and suggestions, and to Dr. Timothy Osborne for proofreading and polishing the language.

References

Ahrens, B., 2004. Prosodie beim Simultandolmetschen. Peter Lang, Frankfurt/Main.
Bürkner, P.C., 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. J. Stat. Softw. 80 (1), 1–28. https://doi.org/10.18637/jss.v080.i01.
Cecot, M., 2001. Pauses in simultaneous interpretation: a contrastive analysis of professional interpreters' performances. Interpreters' Newslett. 11, 63–85.
Cenoz, J., 1998. Pauses and communication strategies in second language speech. College Stud. 11, 1–11.
Chambers, F., 1997. What do we mean by fluency?. System 25 (4) 535–544. https://doi.org/10.1016/S0346-251X(97)00046-8.
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.
Friederici, A.D., Bornkessel, I., 2003. Missing the syntactic piece. Commentary to Ruchkin et al.. Behav. Brain Sci. 26 (6), 735–736. https://doi.org/10.1017/S0140525X03290166.
Gabry, J., Rok Češnovar. 2020. Cmdstanr: R Interface to 'CmdStan'.

Gennari, S., Mirkovic, J., MacDonald, M., 2012. Animacy and competition in relative clause production: a cross-linguistic investigation. Cogn. Psychol. 65, 141–176. https://doi.org/10.1016/j.cogpsych.2012.03.002.

Gibson, E., 1998. Linguistic complexity: locality of syntactic dependencies. Cognition 68, 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1.

Gibson, E., 2000. The dependency locality theory: a distance-based theory of linguistic complexity. In: Miyashita, Y., Marantz, A., O'Neil, W. (Eds.), Image, Language, Brain. MIT Press, Cambridge, pp. 95–126.

Gile, D., 2008. Local cognitive load in simultaneous interpreting and its implications for empirical research. Forum 2, 59–77. https://doi.org/10.1075/forum.6.2.04gil.

Gile, D., 2015. Effort Models. In: Franz Pöchhacker, F. (Ed.), Routledge Encyclopaedia of Interpreting Studies. Routledge, New York, pp. 135–137.

Gong, T., Shuai, L., 2015. Modeling coevolution between language and memory capacity during language origin. PLoS One 10 (11). https://doi.org/10.1371/journal.pone.0142281.

Gordon, P.A., Luper, H.L., 1989. Speech disfluencies in nonstutterers. Syntactic complexity and production task effects. J. Fluency Disord. 14, 429–445.

Grodner, D., Gibson, E., 2005. Some consequences of the serial nature of linguistic input. Cognit. Sci. 29, 261–290. https://doi.org/10.1207/s15516709cog0000_7.

Guillot, M.N., 1999. Fluency and Its Teaching. Multilingual Matters, Clevedon.

Hagoort, P., 2000. Electrocortical reflections on the binding problem for language. Paper presented at the Fourth European Conference of The Federation of European Psychophysiology Societies.

Hawkins, J.A., 1994. A performance theory of order and constituency. Cambridge University Press, Cambridge.

Hawkins, J.A., 2004. Efficiency and Complexity in Grammars. Oxford University Press, Oxford.

Heringer, H., Strecker, B., Wimmer, R., 1980. Syntax: Fragen-Losungen. Wilhelm Fink Verlag, München, Alternativen.

Hsiao, F., Gibson, E., 2003. Processing relative clause in Chinese. Cognition 90, 3–27. https://doi.org/10.1016/S0010-0277(03)00124-0.

Hudson, R., 1995. Measuring Syntactic Difficulty. Unpublished paper. http://www.pon.ucl.ac/home/dick/difficulty.htm (2008-7-6).

Jay, T.B., 2004. The Psychology of Language. Beijing University Press, Beijing.

Jiang, X.L., Jiang, Y., 2020. Effect of dependency distance of source text on disfluencies in interpreting. Lingua 243, 1–18. https://doi.org/10.1016/j.lingua.2020.102873.

Jing, Y.Q., Widmer, P., Bickel, B., 2021. Word Order Variation is Partially Constrained by Syntactic Complexity. Cognit. Sci. 45, e13056.

Jones, R., 2008. Conference Interpreting Explained. Shanghai Foreign Language Education Press, Shanghai.

Kenny, K.D., 1996. Language Loss and the Crisis of Cognition: Between Socio- and Psycholinguistics. Mouton de Gruyter, New York.

King, J., Just, M.A., 1991. Individual differences in syntactic processing: The role of working memory. J. Mem. Lang. 30, 580–602.

Liang, J.Y., Fang, Y.Y., Lv, Q.X., Liu, H.T., 2017. Dependency distance differences across interpreting types: Implications for cognitive demand. Front. Psychol. 8, 1–10. https://doi.org/10.3389/fpsyg.2017.02132.

Lin, Y.M., Lv, Q.X., Liang, J.Y., 2018. Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting. Front. Psychol. 9, 1–13. https://doi.org/10.3389/fpsyg.2018.01543.

Liu, H.T., 2008. Dependency distance as a metric of language comprehension difficulty. J. Cognit. Sci. 9, 159–191. https://doi.org/10.17791/jcs.2008.9.2.159.

Liu, H.T., 2009. Probability distribution of dependencies based on a Chinese dependency treebank. J. Quant. Linguist. 16, 256–273. https://doi.org/10.1080/09296170902975742.

Liu, H.T., Xu, C.S., Liang, J.Y., 2017. Dependency distance: a new perspective on syntactic patterns in natural languages. Phys. Life Rev. 21, 171–193. https://doi.org/10.1016/j.plrev.2017.03.002.

Lv, Q.X., 2020. An Integrated Approach to Syntactic Processing in Interpreting (Doctoral Dissertation). Zhejiang University, Hangzhou.

Macnamara, B.N., Conway, A.R.A., 2016. Working memory capacity as a predictor of simultaneous language interpreting performance. J. Appl. Res. Mem. Cogn. 5, 434–444. https://doi.org/10.1016/j.jarmac.2015.12.001.

McElreath, R., 2020. Statistical rethinking: A Bayesian course with examples in R and Stan. CRC Press, London.

Mead, P., 2002. Exploring hesitation in consecutive interpreting: An empirical study. In: Garzone, G., Viezzi, M. (Eds.), Interpreting in the 21st Century: Challenges and Opportunities. John Benjamins Publishing Company, Amsterdam, pp. 73–82. https://doi.org/10.1075/btl.43.08mea.

Mead, P., 2005. Methodological issues in the study of interpreters' fluency. Interpreters' Newslett. 13, 39–63.

Nalborczyk, L., Batailler, C., Loevenbruck, H., Vilain, A., Bürkner, P.C., 2019. An introduction to bayesian multilevel models using BRMS: a case study of gender effects on vowel variability in Standard Indonesian. J. Speech Lang. Hear. Res.. https://doi.org/10.1044/2018_JSLHR-S-18-0006.

Norris, J.M., Ortega, L., 2009. Toward an organic approach to investigating CAF in instructed SLA: the case of complexity. Appl. Linguis. 30 (4), 555–578.

Osborne, J., 2011. Fluency, complexity and informativeness in native and non-native speech. Int. J. Corpus Linguist. 16 (2), 276–298.

Piccaluga, M., Nespoulous, J.L., Harmegnies, B., 2005. Disfluencies as a window on cognitive processing: An analysis of silent pauses in simultaneous interpreting. Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop. 10–12 September 2005, Aix-en-Provence, France, pp. 151–155.

Plevoets, K., Defrancq, B., 2018b. Lexis or parsing? A corpus-based study of syntactic complexity and its effect on disfluencies in interpreting. https://biblio.ugent.be/publication/8562920

Plevoets, K., Defrancq, B., 2016. The effect of informational load on disfluencies in interpreting. Transl. Interpret. Stud. 2, 202–224. https://doi.org/10.1075/tis.11.2.04ple.

Plevoets, K., Defrancq, B., 2018a. The cognitive load of interpreters in the European Parliament. Interpreting 20 (1), 1–28. https://doi.org/10.1075/intp.00001.ple.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rennert, S., 2010. The impact of fluency on the subjective assessment of interpreting quality. Interpreters' Newslett. 15, 101–115.

Riccardi, A., 1998. Interpreting strategies and creativity. In: Beylard-Ozeroff, A., Kralova, J., Moser-Mercer, B. (Eds.), Translator's Strategies and Creativity. John Benjamins, Philadelphia.

Scontras, G., Badecker, W., Shank, L., Lim, E., Fedorenko, E., 2015. Syntactic complexity effects in sentence production. Cognit. Sci. 39 (3), 559–583.

Seeber, K.G., 2011. Cognitive load in simultaneous interpreting. Interpreting 13 (2), 176–204.

Setton, R., 1999. Simultaneous interpretation: A cognitive-pragmatic analysis. John Benjamins, Amsterdam.

Shen, M.X., Liang, J.Y., 2015. A comparative study on working memory between professional interpreters and student interpreters in simultaneous interpreting through analysis of their high-risk omissions. J. Foreign Lang. 5, 78–83.

Shreve, G.M., Lacruz, I., 2011. Sight translation and speech disfluency. In: Alvstad, C., Hild, A., Tiselius, E. (Eds.), Methods and Strategies of Process Research. John Benjamins Publishing Company, Amsterdam, pp. 93–120.

Sivula, T., Magnusson, M., Vehtari, A., 2022. Unbiased estimator for the variance of the leave-one-out cross-validation estimator for a Bayesian normal model with fixed variance. Commun. Statist.-Theory Methods. https://doi.org/10.1080/03610926.2021.2021240.

Stan Development Team, 2020. RStan: the R interface to Stan. R package version 2.21.2. http://mc-stan.org/

Su, W.C., Li, D.F., 2019. Identifying translation problems in English-Chinese sight translation. Translation and Interpreting. Studies. 14 (1), 110–134. https://doi.org/10.1075/tis.00033.su.

Tissi, B., 2000. Silent pauses and disfluencies in simultaneous interpreting: a descriptive analysis. Interpreters Newslett. 10, 103–112 http://hdl.handle.net/10077/2455.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., Bürkner, P.C., 2020. Rank Normalization, Folding, and Localization: An Improved R! for Assessing Convergence of MCMC. Supplementary Material. Bayesian Anal.. https://doi.org/10.1214/20-BA1221SUPP.

Wang, J., 2016. Effects of chunk cognition training on fluency of simultaneous interpreting: an empirical study. Foreign Lang. Teach. Res. 48, 765–775.

Wang, B.H., Li, T., 2015. An empirical study of pauses in Chinese-English simultaneous interpreting. Perspectives 1, 124–142. https://doi.org/10.1080/0907676X.2014.948885.

Wang, Y.Q., Liu, H.T., 2017. The effects of genre on dependency distance and dependency direction. Lang. Sci. 59, 135–147. https://doi.org/10.1016/j.langsci.2016.09.006.

Watanabe, M., Hirose, K., Den, Y., Minematsu, N., 2008. Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. Speech Comm. 50, 81–94. https://doi.org/10.1016/j.specom.2007.06.002.

Wilson, G.F., Eggemeier, F.T., 2006. Mental workload measurement. In: Karwowski, W. (Ed.), International Encyclopedia of Ergonomics and Human Factors. CRC Press, FL, pp. 814–817.

Winter, B., Bürkner, P.C., 2021. Poisson regression for linguists: a tutorial introduction to modelling count data with BRMS. Lang. Linguist. Compass. https://doi.org/10.1111/lnc3.12439 e12439.

Xu, H.M., 2010. Pauses in conference consecutive interpreting from English into Chinese: An empirical study. Foreign Lang. Res. 119 (1), 64–71.

Yao, Y.L., Vehtari, A., Simpson, D., Gelman, A., 2017. Using Stacking to Average Bayesian Predictive Distributions. Bayesian Anal. 1–78. https://doi.org/10.1214/17-BA1091.

Yngve, V., 1960. A model and a hypothesis for language structure. Proc. Am. Philos. Soc. 104, 444–466.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, Berlin.