



Correlations Among Disfluency Markers in Simultaneous Interpreting: Speech Production in Bilingual Mode

- Elaheh Ghanavati***  MA in Translation Studies, Department of English Language and Literature, Azarbaijan Shahid Madani University, Tabriz, Iran
- Reza Yalsharzeh**  Assistant Professor, Department of English Language and Literature, Azarbaijan Shahid Madani University, Tabriz, Iran
- Davoud Amini**  Associate Professor, Department of English Language and Literature, Azarbaijan Shahid Madani University, Tabriz, Iran
- Mehran Soleymani**  Associate Professor, Department of Educational Sciences and Psychology, Azarbaijan Shahid Madani University, Tabriz, Iran

Abstract

Speech disfluency, including pauses and repetitions, is a key yet underexplored factor, and correlations among disfluency markers remain poorly understood. This study examined intercorrelations among nine speech disfluency markers in English-to-Persian simultaneous interpreting, using recordings from 71 purposively sampled bilingual adults. Pearson correlations among disfluency markers showed significant positive associations (e.g., filled pauses and multisyllabic word repetitions) and a negative relationship between blocks and filled pauses. These findings highlight interconnected disfluency patterns, providing insights into speech production mechanisms in the bilingual mode, when both languages are fully active.

Keywords: Speech disfluency, disfluency markers, simultaneous interpreting, bilingual modes, speech production

*Corresponding author: elaheghhanavati01@gmail.com

Cite this article: Ghanavati, E., Yalsharzeh, R., Amini, D., & Soleymani, M. (2026). Correlations among disfluency markers in simultaneous interpreting: Speech production in bilingual mode. *Translation and Interpreting Research*, 3(10), 1-16. <https://doi.org/10.22054/tir.2026.93611.1068>



© 2026 – Allameh Tabataba'i University Press

Print ISSN: 3092-6599

Online ISSN: 3060-7019

Publisher: Allameh Tabataba'i University Press

Published: 30.06.2026

Accepted: 19.05.2026

Received: 02.04.2026

1. Introduction

According to Grosjean (1997, 2001, 2008, 2013), in exceptional instances, both languages can be fully active in the bilingual mode. Two examples stand out: first, when a bilingual listens to two individuals, each using a different language; and second, when someone is interpreting from one language to another. Unlike the typical bilingual mode, where one language is more active than the other, interpreting requires both the source and target languages to be equally active, as both are needed for comprehension and production. Although the two languages are equally active, the processing mechanisms differ: the input of both the source and target languages is active, whereas the target language output mechanism is active, with the source language output mechanism being inactive or even exceptionally inhibited (Figure 1). Investigating the links between different disfluency markers in interpreting, therefore, provides valuable insights into its constructs and enhances our understanding of disfluency from various viewpoints. This exploration also contributes to a better grasp of the information-processing mechanisms involved in simultaneous interpreting and the bilingual mode.

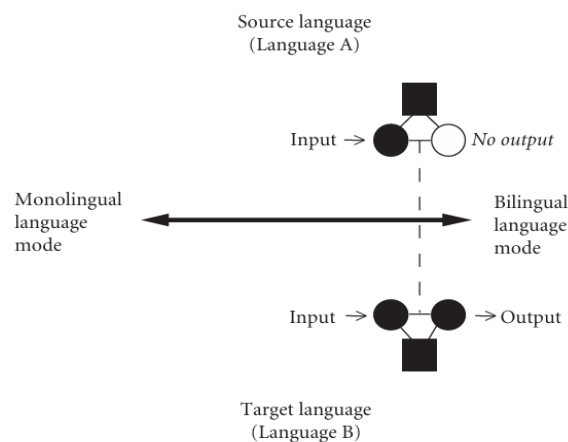


Figure 1. The Mechanisms of Bilingual Mode during Simultaneous Interpreting
Note. Reprinted from Grosjean (2001, p. 62)

Studying disfluencies is crucial as they offer meaningful insights into the functioning of the fluent speech production mechanism (Shattuck-Hufnagel, 1979). Speech can be interrupted when a speaker experiences articulatory disfluencies, such as stuttering (Zhao, 2022). Non-stuttering speakers may also show articulatory disfluencies, often struggling to produce a syllable or word. These articulatory disfluencies arise from challenges in speech programming, not from intentional repetition (Zhao, 2022). Stuttering is characterized by an unusual number of functional disfluencies (Bayerl et al., 2023). Stuttering, therefore, is a severe form of disfluency (Zhao, 2022).

In simultaneous interpreting, one must listen to, understand, and process incoming speech while producing their output at the same time, all while continuing to process further output (Tóth, 2011). In the course of speech production during simultaneous interpreting, individuals possess a monitoring function (Petite, 2005) that allows them to revise their output, resulting in self-modifications that can cause disfluencies in their speech. However, self-modification is not the only factor contributing to these disfluencies in interpreters' output (Tóth, 2011). Yet, the phenomenon of speech disfluency in the outcome of simultaneous interpreting has not been extensively explored (Tóth, 2011).

Over the past years, research has provided insights into the general characteristics of normal disfluencies in various languages. However, per our existing data, there is currently no normative reference for disfluent speech in Persian during simultaneous interpreting involving non-stuttering bilinguals. Furthermore, to the best of our knowledge, no study has conducted a comprehensive

analysis of the correlations between disfluency markers during simultaneous interpreting, as a special case of the bilingual mode (Grosjean, 2001). It is also important to note that prior research has consistently shown that bilingual individuals (who do not stutter) exhibit significantly higher rates of speech disfluencies than their monolingual peers, frequently surpassing traditional monolingual clinical thresholds (Saad Merouwe et al., 2024; Aslan & Eggers, 2025).

The literature suggests that while certain disfluency markers may be related, potentially due to shared causes like cognitive load or production difficulties, direct, systematic studies on correlations between disfluency types are scarce. Previous research has predominantly focused on speech disfluency markers in isolation and their roles in specific contexts (e.g., language acquisition or pathology), with little attention given to their interrelationships. In this study, we aim to examine the potential co-occurrence of nine disfluency markers, namely monosyllabic word repetitions, part-word repetitions, prolongations, blocks, broken words, filled pauses, revisions and abandoned sentences, multi-syllable word repetitions, and phrase repetitions in bilinguals' speech, positing that such markers may indicate shared underlying mechanisms tied to the cognitive and linguistic demands associated with simultaneous interpreting as an exceptional case of the bilingual mode (Grosjean, 2001). Additionally, the research seeks to determine whether disfluency markers are interconnected rather than independent, potentially due to the pressures and constraints inherent in real-time language processing during simultaneous interpreting. The study attempts to address the following research question: Are there any correlations among bilinguals' speech disfluency markers (monosyllabic word repetitions, part-word repetitions, prolongations, blocks, filled pauses, broken words, revisions and abandoned sentences, multisyllabic word repetitions, and phrase repetitions) when they are in the bilingual mode, performing simultaneous interpreting?

2. Literature Review

According to Shriberg (2001), disfluencies are significant in the processes of speech production and comprehension, as they offer insights into the underlying mechanisms influencing human speech and language production. While spoken languages appear to follow a linear order, with one word following another in the acoustic channel, they are actually governed by various non-linear processes, as speakers continuously work on their output to reshape the course of their delivery. Spoken languages thus often involve cycles of what are commonly known as "fluent" versus "disfluent" (Kosmala, 2024). Disfluencies are prevalent in human speech, with research suggesting that they occur between six and ten times for every hundred words (Kosmala, 2024; Shriberg, 1994).

In the 1950s, Goldman-Eisler (1958) discovered that the occurrence of pauses in speech was not accidental. Language production availability issues are commonly observed in daily life, where low availability in speech production leads primarily to pauses and hesitations that can disrupt the flow of speech, making it unpleasant for listeners. It may also result in a failure to convey the intended message accurately if speakers give up on finding the right words in favor of more readily available ones that don't quite capture their meaning (Gile, 2009).

The stuttering classification commonly utilized in clinical and research settings, put forward by the Illinois Stuttering Research Program (Ambrose & Yairi, 1999), includes two categories of disfluencies (Figure 2): stuttering-like disfluencies (SLD) and other disfluencies (OD). The stuttering-like disfluencies (SLD), such as part-word repetitions, are characteristic of stuttering speech. They are distinguished from other disfluencies due to their higher level of uncontrollability: prolongations of syllables and blocks (occurring in more than 10% of words) are displayed with irregular durations. Additionally, these disfluencies feature visible and audible changes in speech, including significant fluctuations in speed (either very slow or excessively fast), voice intensity (either too weak or too strong), and pitch (ranging from very high to very low). Other disfluencies (OD) are common in the speech of individuals who do

not stutter, but they can also appear in stuttering speech, albeit less frequently than stuttering-like disfluencies (SLD). This category encompasses multisyllabic and phrase repetitions, interjections, revisions, and incomplete phrases (For additional details, see Natarelli, 2018; also refer to Jansson-Verkasalo et al., 2020).

Articulatory disfluencies often arise due to difficulties in speech planning (Zhao, 2022). Furthermore, speech naturally contains disfluencies, and the frequency of these disfluencies in an individual's speech is affected by different factors, such as their cognitive load and psychological traits. These disfluencies may manifest as interjections, which are frequently employed to provide speakers with additional time to think while speaking. They can also arise when an individual seeks to modify or edit the initial sentence by restarting or rephrasing it (Bayerl et al., 2023).

Some studies have delved into speech disfluencies in simultaneous interpreting. Pöschhacker (1995) investigated instances of speech repairs in the speaker's speech and the corresponding output of simultaneous interpreters. The study revealed more speech repairs in the target language output. Also, while speakers' output primarily featured simple errors and false starts, interpreters' output predominantly encompassed lexical and structural blends as the most frequent forms of disfluencies. Additionally, Pöschhacker (1995) analyzed the occurrence of slips and shifts in simultaneous interpreters' output, and the findings indicated that interpreters show a higher frequency of slips and shifts in their output than speakers. Moreover, Cecot (2001) examined the occurrence of non-fluency in simultaneous interpreting, particularly in relation to changes in the speaker's speech rate, and found that the target text exhibited an average increase in filled pauses, revisions, and lengthening of consonants and vowels.

Furthermore, numerous studies have explored speech disfluency and examined various influencing or influenced variables. Research on cognition and speech planning includes foundational works by Bock (1986), Bock and Levelt (1994), Levelt (1989), Dell (1986), Fromkin (1971, 1973, 1980), Goldman-Eisler (1958), MacKay (1970, 1972, 1973), Garrett (1975), Maclay and Osgood (1959), Shattuck-Hufnagel and Klatt (1979), Nooteboom (1969), and Shattuck-Hufnagel (1979, 1982), as well as more recent contributions by Song and Li (2020) and Lin et al. (2018). Additionally, studies addressing affect and disfluency encompass Bayerl et al. (2023), Zhao (2022), and Zhao et al. (2023). Moreover, the role of the interpreter's sex in their speech disfluency has been investigated by Collard and Defrancq (2019). However, while existing research has identified various disfluency markers in speech, there is a notable gap in understanding how these markers correlate. Consequently, there is no established threshold or reference benchmark in the literature to rely upon. Moreover, another gap in the literature on interpreting processes lies in the under-exploration of Grosjean's (2001) concept of the bilingual mode, despite its relevance for understanding how simultaneous interpreting (full co-activation of languages) takes place.

3. Methodology

The study comprised a total of 71 Iranian adult bilinguals proficient in both Persian (L1) and English (L2), including 8 (PhD) university professors (11.2%), 16 master's students (22.6%), and 47 bachelor's students (66.2%), selected through purposive sampling to ensure adequate language comprehension, facilitating a more targeted and meaningful exploration of speech disfluency. The participant pool consisted of professors and students from Azarbaijan Shahid Madani University of Iran, trained in subfields of English translation, English language teaching, and English literature (all completed the interpreting and translation courses). The participants were aged between 20 and 50 years. This age range encompasses a spectrum of experiences across different life stages, which may enhance the diversity and richness of the data collected.

The assessments were conducted individually in a language laboratory, allowing for focused data collection. This one-on-one approach ensured participants received the necessary attention and support, minimizing confounding variables. The procedure was initially piloted with two English language professors (the second and third authors), allowing for the identification and rectification of any procedural flaws before the actual implementation of the study. Furthermore, the data that did not meet the requisite standards of language proficiency were systematically excluded from the analysis to ensure the integrity and validity of the findings.

Initially, a speaking test was administered to all participants to assess their proficiency levels in English and fluent speaking ability. To ensure objectivity in the evaluation, the speaking test was taken, analyzed, and scored using the AI capabilities of the *Speechace* website. The three questions posed were uniform for all participants to maintain consistency in the assessment process. The selected metric for this study was IELTS, and participants were required to achieve a minimum band score of 7 on at least one question to qualify for participation in the subsequent test.

Subsequently, participants were presented with an audio for interpretation to induce the bilingual mode and full co-activation of languages. The audio selected for this task was a TED Talk by Celeste Headlee (2015), entitled *10 Ways to Have a Better Conversation*. Participants were instructed to listen to the first four minutes of the talk to familiarize themselves with the topic, the speaker's accent, the pace, and the density of the speech. Simultaneous interpreting commenced from the fifth minute of the audio, with participants interpreting from English (L2) to Persian (L1). The total duration for the simultaneous interpreting task, which involved active interpreting rather than just listening, was five minutes and thirteen seconds. Participants were not permitted to pause the audio or replay segments during the interpretation process. The task was conducted under the careful observation of the examiner (first author), ensuring adherence to the task's guidelines.

All recorded audio interpretations and their corresponding waveforms provided by participants were examined through repeated listening by the experimenter (first author), utilizing the Sonic Visualiser software, version 5.0.1. We established clear criteria for analysis and utilized a structured framework (Figure 2). Also, to enhance the validity and reliability of our findings, we randomly selected five audio recordings from our dataset, and an English language professor (third author), with 23 years of academic experience and prior expertise in speech analysis, as evidenced by a published article (Rahimpour & Amini, 2013), was asked to independently evaluate these five recordings based on the specified criteria (Figure 2). The professor's analyses were subsequently compared with those of the primary examiner (first author) to assess inter-rater reliability using the Pearson correlation coefficient (Table 2), computed via IBM SPSS Statistics. This approach was designed to ensure that our labeling system is explicit, transparent, and well-defined. Additionally, the inter-rater agreement score serves as a reference for the reliability of the speech disfluency assessment itself.

For the analysis of audio recordings, auditory techniques were implemented to ensure the precision of acoustic phonetics evaluations. Audio files were systematically manipulated, including slowing down, speeding up, and making specific speed adjustments as required, to optimize auditory clarity and precision in the identification of disfluencies. These modifications were complemented by repeated and focused listening sessions, where particular attention was directed towards detecting any sign of disfluency in speech, even at the vowel level. The analysis primarily focused on identifying signs of disfluencies, which were labeled using a classification of nine disfluency types, based on the criteria proposed by Ambrose and Yairi (1999) and Stes (2004) (Figure 2). For the purposes of this study, a speech disfluency was classified as a block if it lasted for five seconds. This classification implies that a block lasting, for example, ten seconds, would be considered two separate blocks. This distinction differentiates speech blocks from mere queuing phenomena and silent pauses, indicating that they can be considered extreme forms of disfluency. The technique, known as queuing, involves briefly stopping the interpretation for two or three seconds in order to simply listen. This technique is applied

specifically when a segment of speech is highly dense or rapid; when there has been a distraction or momentary loss of attention; when low volume or poor audio quality makes it challenging to follow the speaker; or in any other situation where the interpreter starts to lose track of the thread of meaning (Camayd-Freixas, 2011). Furthermore, in the analysis of the “revisions”, our approach extended beyond the sentence level to examine linguistic elements at the word level, allowing us to uncover subtle changes that might not be noticeable in broader syntactic assessments. Additionally, we focused specifically on analyzing the number of occurrences of disfluent markers in the spoken discourse, rather than examining their duration or other related features. This approach allowed us to quantify how frequently various disfluencies manifest within the speech samples. This quantitative analysis provided a clear and objective framework for understanding the patterns of speech disfluencies and facilitated the comparison between different speakers.

In furtherance of this, the translated text (Persian version) of the original audio was obtained from the *Downsub* website. This text served as a reference for comparing the interpretation audio provided by participants. The objective of this comparison was to identify instances of “abandoned sentences” in participants’ interpretations. By analyzing discrepancies between the participants’ audio and the reference text, we aimed to enable an objective assessment and inspect the completeness of the interpretations, to ascertain whether low disfluency rates indicated authentic fluency or, conversely, stemmed from strategic avoidance, including the use of safe but potentially inaccurate formulations or deliberate silence. To accomplish this task, we segregated each sentence in the source speech and coded for final interpreting analysis. The overall count of sentences was 82 (the cumulative count of abandoned sentences was then merged with the total number of revisions to calculate the conclusive score for the “revision and abandoned sentences” measure). Therefore, it should be noted that the term “revision” encompasses the total count of both revised and abandoned sentences; however, for the purpose of conciseness, it will be designated simply as “revision.”

	Disfluency type (abbreviation)	Description	Examples in English
Stuttering like disfluencies (SLD)	Part word repetition (PWR)	Repetition of a sound or syllable within or in the beginning of the word	thi thi this
	Monosyllabic word repetition (MonWR)	Repetition of a monosyllabic word	and and
	Prolongation (P)	Prolongation of a sound resulting in an atypical duration of that sound	mmmy coookie
	Block (B)	Stopping airflow and sound during or before production of a vowel or consonant	...girl
	Broken words (BW)	Stopping of airflow in the middle of a vowel within a word	o...pen
Other disfluencies (OD)	Filled pause (FP)	Filler words or non-linguistic sounds used within an utterance	Hey look uhmm at that
	Revision and abandoned sentences (R)	Correction of word choice, grammatical or phonological errors or add or delete lexical information	Mom ate fixed dinner. I want Hey look at that
	Multisyllable word repetition (MulWR)	Repetition of a multi-syllabic word	because because
	Phrase repetition (PR)	Repetition of a phrase within an utterance	I want I want to go

Figure 2. Speech Disfluencies Classification

Note. Ambrose & Yairi (1999) and Stes (2004), reprinted from Jansson-Verkasalo et al. (2020, p. 10). Also refer to Natarelli (2018).

Subsequently, all the data were subjected to analysis using IBM SPSS Statistics. To quantify the relationships between all disfluency markers, Pearson correlation coefficients were calculated for each pair based on their frequency of occurrence in the samples. These samples were derived from audio recordings of participants’ performances, which were transcribed and coded for the presence of each disfluency type. The Pearson correlation method was selected as it measures the linear association between variables, enabling the identification of potential patterns or clusters of disfluency markers that may co-occur. Furthermore, the analysis provides insights into the nature of these relationships;

positive correlations might suggest shared underlying mechanisms, such as cognitive load, while negative correlations could indicate distinct processing challenges.

4. Results

4.1. Descriptive Statistics

The descriptive statistics of the nine disfluency markers are provided in Table 1 below.

Table 1. Descriptive Statistics of Speech Disfluency Markers

	Participants	Minimum	Maximum	Mean	Std. Deviation
PWR	71	.00	40.00	7.2676	5.85286
MonWR	71	.00	14.00	2.6197	2.49781
P	71	2.00	53.00	17.9014	8.97004
B	71	.00	42.00	17.3521	10.06848
BW	71	.00	4.00	.3803	.83437
FP	71	.00	102.00	32.9014	21.02390
R	71	16.00	66.00	40.9718	11.50773
MuIWR	71	.00	9.00	3.3662	2.42746
PR	71	.00	10.00	2.4085	2.06036

Note. Part-word repetitions (PWR), monosyllabic word repetitions (MonWR), prolongations (P), blocks (B), broken words (BW), filled pauses (FP), revisions and abandoned sentences (R), multisyllabic word repetitions (MuIWR), and phrase repetitions (PR).

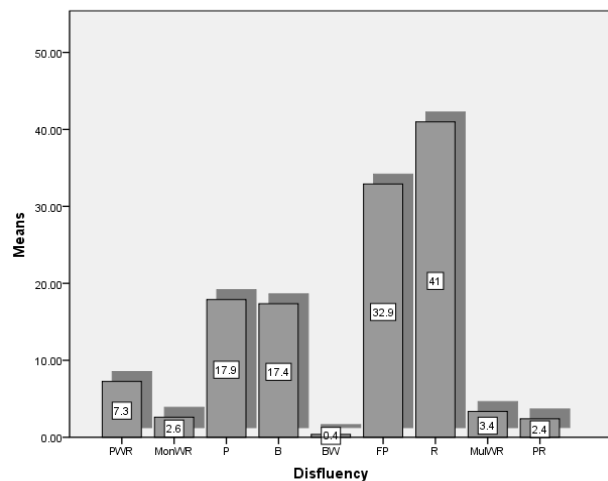


Figure 3. Overall Histogram of Disfluency Markers

The frequency of speech disfluencies varied widely across the nine types measured. The ranking of the disfluencies by mean frequency (highest to lowest) is as follows:

1. Revisions
2. Filled Pauses
3. Prolongations
4. Blocks
5. Part-Word Repetitions
6. Multisyllabic Word Repetitions

7. Monosyllabic Word Repetitions
8. Phrase Repetitions
9. Broken Words

The analysis reveals significant variation in the frequency of speech disfluencies. Revisions and filled pauses were the most prevalent. These disfluencies are typically associated with normal speech patterns. Revisions reflect self-correction or reformulation, and filled pauses (e.g., “um,” “uh”) serve as hesitation markers during speech planning. Their high frequency suggests that participants frequently engaged in these behaviors, possibly due to the cognitive demands of the speaking task (simultaneous interpreting) and the full co-activation of languages in the bilingual mode. Prolongations and blocks were moderately frequent. These disfluencies are often linked to stuttering, involving elongated sounds (prolongations) or interruptions in airflow (blocks). Their presence at moderate levels might indicate a mix of typical and atypical speech patterns in the sample. Part-word repetitions, followed by multisyllabic word repetitions, monosyllabic word repetitions, and phrase repetitions, were less common, suggesting that these repetition-based disfluencies are not dominant features of the sample’s speech. Broken words were exceptionally rare. The standard deviations provide further insight: filled pauses and revisions showed high variability, suggesting individual differences in speech production and cognitive processes, while broken words had low variability, reinforcing their rarity. The wide spread in standard deviations suggests that these measures are not consistent across all speakers. This reflects the multifaceted nature of speech, which is shaped by a mix of personal traits.

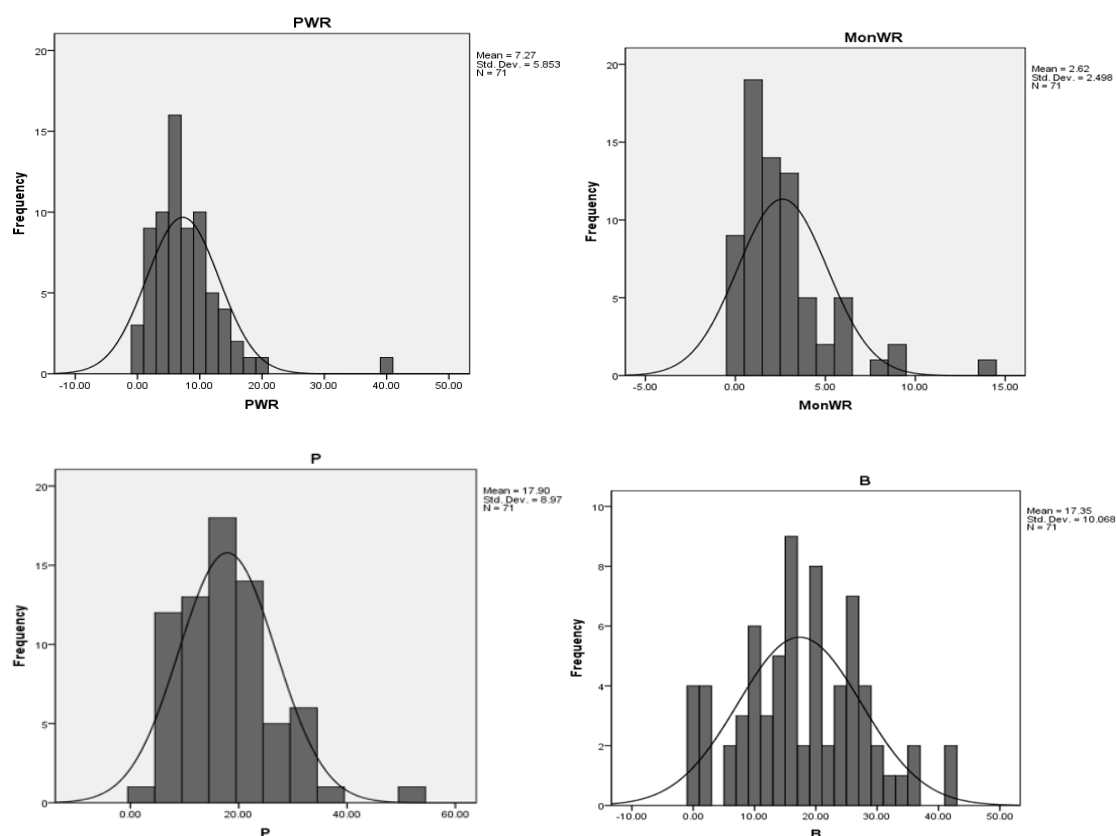


Figure 4. Histograms of Disfluency Markers – Part 1

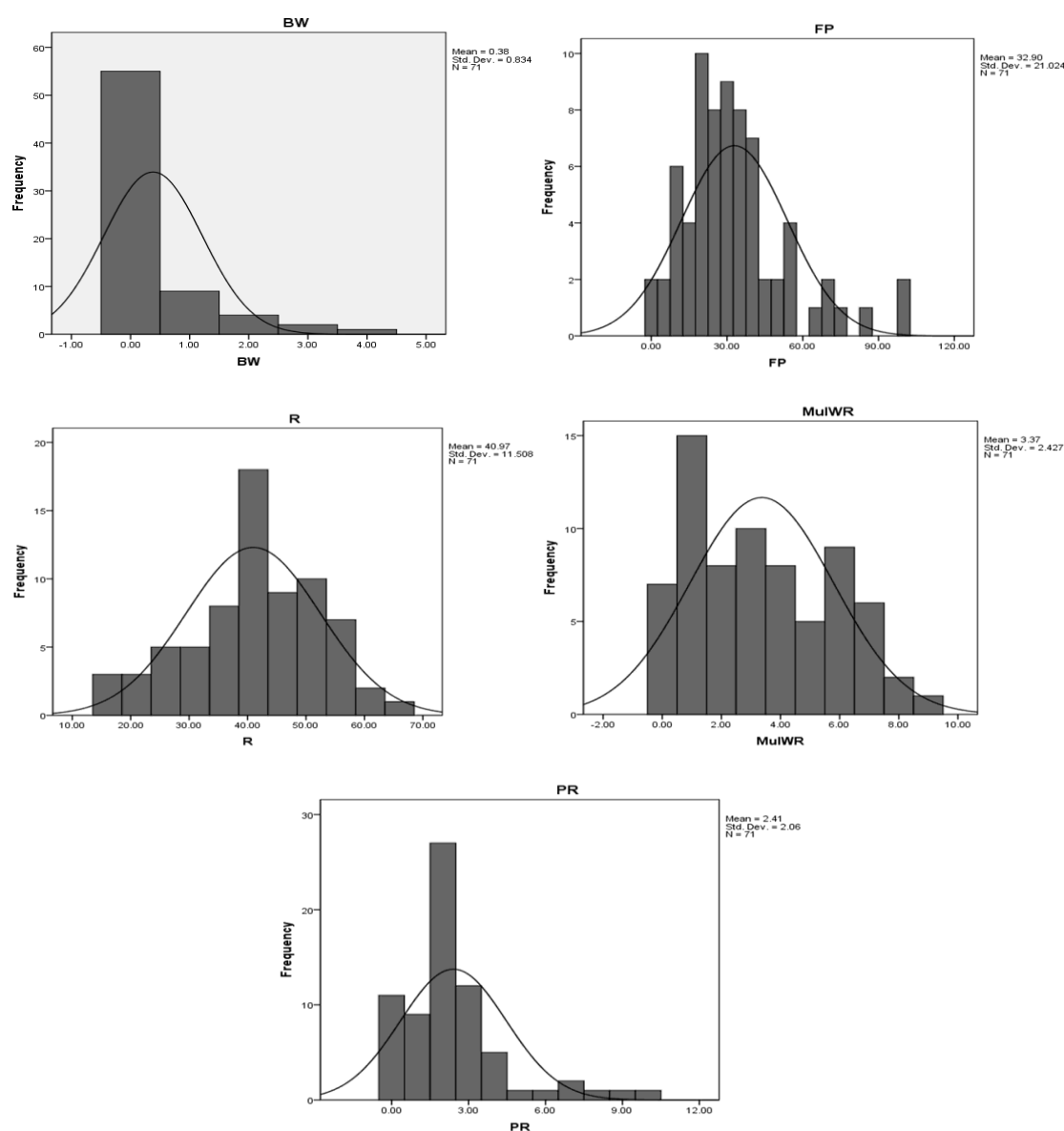


Figure 5. Histograms of Disfluency Markers – Part 2

4.2. Inter-rater Reliability

In this study, two raters (the first and third authors) independently evaluated 5 audio recordings to score the occurrence of 9 disfluency markers (Figure 2). Afterwards, the Pearson correlation coefficient was computed between their scores. The results, shown in the table below, reveal a very strong positive correlation, demonstrating a high level of agreement between the two raters (Table 2). The significant level of agreement further confirms that this strong correlation is not due to chance, providing solid evidence of the reliability of the rating process. The sample size of 45 cases (5×9) rated by both raters supports the robustness of this finding. The high inter-rater reliability observed here ensures that the data derived from the ratings are consistent, enhancing the validity of the study's conclusions and its suitability for further analysis.

Table 2. Inter-rater Reliability

		VAR00002	VAR00001
VAR00002	Pearson Correlation	1	.929**
	Sig. (2-tailed)		.000
	N	45	45
VAR00001	Pearson Correlation	.929**	1
	Sig. (2-tailed)	.000	
	N	45	45
**. Correlation is significant at the 0.01 level (2-tailed).			

4.3. Correlation Analyses

Pearson correlation was calculated between part-word repetition (PWR), monosyllabic word repetitions (MonWR), and prolongations (P). The results are presented in Table 2 below.

Table 3. Pearson Correlation Coefficients between PWR, MonWR, and P

		PWR	MonWR	P
PWR	Pearson Correlation	1	.605**	.082
	Sig. (2-tailed)		.000	.497
	N	71	71	71
MonWR	Pearson Correlation	.605**	1	-.002
	Sig. (2-tailed)	.000		.985
	N	71	71	71
P	Pearson Correlation	.082	-.002	1
	Sig. (2-tailed)	.497	.985	
	N	71	71	71
*. Correlation is significant at the 0.05 level (2-tailed).				
**. Correlation is significant at the 0.01 level (2-tailed).				

Note. Part-word repetitions (PWR), monosyllabic word repetitions (MonWR), and prolongations (P)

Part-word repetition and monosyllabic word repetition were strongly correlated: The occurrence of a disfluency marker can prompt the emergence of an additional disfluency marker, contributing to a deterioration in speech production. This strong positive correlation highlights a significant relationship at the word level. Part-word repetition (e.g., “b-ba-ball”) and monosyllabic word repetition (e.g., “I I want”) both involve disruptions in word production, possibly stemming from word retrieval difficulties or motor speech issues. This finding echoes early research by Maclay and Osgood (1959), who noted that hesitations and repetitions frequently co-occur in spontaneous speech, suggesting a shared mechanism that, when triggered, may worsen speech outcomes.

Table 4. Pearson Correlation Coefficients between FP, R, MulWR, and PR

		FP	R	MulWR	PR
FP	Pearson Correlation	1	.001	.252*	.321**
	Sig. (2-tailed)		.994	.034	.006
	N	71	71	71	71
R	Pearson Correlation	.001	1	.055	.272*
	Sig. (2-tailed)	.994		.648	.022
	N	71	71	71	71
MulWR	Pearson Correlation	.252*	.055	1	.524**
	Sig. (2-tailed)	.034	.648		.000
	N	71	71	71	71
PR	Pearson Correlation	.321**	.272*	.524**	1

	Sig. (2-tailed)	.006	.022	.000	
	N	71	71	71	71

Note. Filled pauses (FP), revisions and abandoned sentences (R), multi-syllable word repetitions (MulWR), and phrase repetitions (PR).

Filled pauses exhibited positive correlations with multisyllabic word repetition: The emergence of one disfluency marker can result in the activation of another, thereby amplifying the challenges associated with speech production. The positive correlation suggests that these disfluencies may co-occur as speakers grapple with speech production challenges. Filled pauses (e.g., “um,” “uh”) are commonly associated with hesitation or planning under cognitive load (Clark & Fox Tree, 2002), while multisyllabic word repetitions may reflect difficulties in articulating complex lexical items. Their co-occurrence could indicate that cognitive strain prompts multiple disfluency strategies to emerge simultaneously, amplifying production challenges.

A notably strong correlation was observed between multisyllabic word repetition and phrase repetition: Likewise, the presence of one disfluency marker can trigger the appearance of another disfluency marker, thereby exacerbating speech deficits. This correlation indicates that challenges with complex linguistic units may manifest across both word and phrase levels. Repeating multisyllabic words and entire phrases could reflect difficulties in managing linguistic complexity, where the emergence of one disfluency triggers another, further impairing fluency. This pattern underscores how interconnected disfluencies can amplify speech deficits under production pressure.

Filled pauses exhibited positive correlations with phrase repetition: Again, the existence of a disfluency marker can catalyze the presence of other disfluency markers, potentially worsening speech deficits. The positive correlation points to a related dynamic. Phrase repetition, where a speaker repeats a segment of speech, may function as a stalling mechanism akin to filled pauses, allowing time for planning or reformulation. This clustering of disfluencies aligns with observations that they often reflect heightened cognitive demand (Shriberg, 1994), suggesting that the presence of one marker can catalyze another, potentially exacerbating speech disruptions.

Revision was positively linked to phrase repetition: Once more, the emergence of one disfluency marker can precipitate the occurrence of additional disfluency markers, thereby intensifying speech deficits. This positive correlation suggests that self-correction and phrase repetition are linked, possibly as part of a monitoring process. Speakers may repeat phrases while revising their speech, consistent with Levelt’s (1983) speech production model, which includes mechanisms for error detection and repair. This co-occurrence could intensify speech deficits as one disfluency precipitates another during real-time correction efforts.

Table 5. Pearson Correlation Coefficients between B, BW, and FP

		B	BW	FP
B	Pearson Correlation	1	.064	-.332**
	Sig. (2-tailed)		.597	.005
	N	71	71	71
BW	Pearson Correlation	.064	1	-.017
	Sig. (2-tailed)	.597		.886
	N	71	71	71
FP	Pearson Correlation	-.332**	-.017	1
	Sig. (2-tailed)	.005	.886	
	N	71	71	71

Note. Blocks (B), broken words (BW), and filled pauses (FP).

Blocks showed a negative correlation with filled pauses. In contrast, this negative correlation suggests a divergent relationship. Blocks, marked by sudden speech cessation, differ from filled pauses, which fill silence with vocalizations. This inverse association may indicate that these disfluencies represent alternative responses to production difficulties—blocks linked to motor execution issues, as seen in stuttering (Bloodstein & Bernstein Ratner, 2008), and filled pauses tied to planning. Speakers prone to one may be less likely to exhibit the other, reflecting distinct underlying processes rather than a compounding effect.

Some further observations are listed below:

- Filled pause and revision: No correlation is observed, with an *r*-value near zero.
- Revision and multisyllabic word repetition: A negligible positive correlation is not significant.
- PWR and prolongation: A weak positive correlation is not significant.
- Monosyllabic word repetition and prolongation: No correlation exists.
- Block and broken word: A negligible positive correlation is not significant.
- Broken word and filled pause: No correlation exists.

5. Discussion

The analysis of disfluency markers in the dataset reveals a noteworthy pattern: three specific markers (monosyllabic word repetition, filled pause, and phrase repetition) emerge as particularly dominant, each appearing prominently across three distinct correlations. This prevalence underscores their significance in the broader landscape of speech disfluencies, suggesting that they may serve as core indicators of underlying cognitive or articulatory processes. By prioritizing these dominant markers, researchers and practitioners can develop more targeted training interventions aimed at mitigating disfluencies. Furthermore, this dominance facilitates the identification of clustered disfluency markers and may help to propose a new typology of disfluency markers for bilinguals, categorizing them based on frequency and correlation strength. Such a typology could classify markers into hierarchical levels (e.g., high-level to low-level phonological disruptions).

This study acknowledges several limitations that future research could address. A noteworthy limitation of this study is the use of a five-minute task duration and the TED Talk audio style, which may constrain the generalizability of the findings. The five-minute duration may not fully capture the dynamics of longer tasks or sustained engagement, potentially limiting the applicability of the results.

Another notable limitation of this study is the relatively small sample size ($N = 71$). This limited number may affect the statistical power of the analysis and the generalizability of the findings to broader populations. While this sample was chosen to ensure focused and in-depth data collection, it may not fully capture the diversity of individuals present in the wider community. The requirement for in-person participation and direct measurement of physical outcomes necessitated a manageable sample size to ensure the feasibility and integrity of the research protocol. However, we should also highlight that, compared to similar studies which had smaller sample sizes of 22 (Lin et al., 2018), 28 (Song & Li, 2020), 53 (Zhao, 2022; Zhao et al., 2023), and 54 (Jansson-Verkasalo et al., 2020), our study with 71 participants exceeded these thresholds.

Additionally, due to the constraints of this study, only one individual (the examiner, first author) conducted repeated listening and analysis of all audio recordings to check for signs of disfluencies. Although we employed inter-rater reliability techniques to enhance the robustness of our findings, analyzing the audio recordings by multiple independent researchers could provide an even more comprehensive understanding. It is noteworthy that, in contrast to similar studies such as Zhao (2022),

which utilized two additional raters to evaluate two audio recordings, our study engaged one additional rater to assess five audio recordings. While our approach involved one fewer rater, the analysis of a larger number of recordings enhances the representativeness of the dataset and strengthens the robustness of our validation process. Moreover, based on our analysis, there is no report of the specific scores or mean values obtained for inter-rater reliability in similar works such as Lin et al. (2018), Song and Li (2020), Zhao (2022), and Zhao et al. (2023), which limited our ability to draw direct comparisons and set a threshold for inter-rater reliability.

6. Conclusion

This study aimed to investigate the correlations among nine specific speech disfluency markers (monosyllabic word repetitions, part-word repetitions, prolongations, blocks, broken words, filled pauses, revisions and abandoned sentences, multisyllabic word repetitions, and phrase repetitions) in the speech of bilinguals during simultaneous interpreting, performed to induce bilingual mode and full co-activation of languages. The analysis revealed significant positive correlations between several markers, including filled pauses with both multisyllabic word repetitions and phrase repetitions, as well as a strong correlation between part-word and monosyllabic word repetitions. Additionally, multisyllable word and phrase repetitions were found to be strongly linked, while blocks showed a negative correlation with filled pauses. These findings fill a critical gap by demonstrating that certain disfluency markers tend to co-occur, potentially reflecting shared cognitive or motor origins, while others, such as blocks and filled pauses, may represent distinct processing challenges. Moreover, our analysis identified monosyllabic word repetition, filled pause, and phrase repetition as dominant disfluency markers, each prominent in three different correlations, highlighting their central role. Prioritizing these markers enables targeted training interventions and helps to recognize new clustered disfluency patterns for bilinguals, paving the way for a proposed typology that categorizes markers by frequency and, more importantly, interrelations, offering a framework for future research and potentially also addressing the limitations of applying monolingual norms to bilingual populations (Aslan & Eggers, 2025). These insights have implications for theories of bilingual speech production and practical applications, such as improving language assessment in interpreters by targeting clusters of disfluencies rather than treating them in isolation. Future studies could consider incorporating neurological perspectives into the study of disfluencies and, by utilizing brain imaging techniques, they can explore how different patterns of brain activity (such as Broca and Wernicke's areas) correspond to a specific disfluency marker.

Ethical Statement

A code of ethics was obtained from the Research Ethics Committee of Azarbaijan Shahid Madani University, and informed written consent was signed and obtained from each participant.

References

- Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research: JSLHR*, 42(4), 895–909. <https://doi.org/10.1044/jslhr.4204.895>
- Aslan, G., & Eggers, K. (2025). Multilingualism, speech disfluencies, and stuttering: A scoping review. *Journal of Speech, Language, and Hearing Research: JSLHR*, 68(8), 3869–3886. https://doi.org/10.1044/2025_JSLHR-24-00479

- Bayerl, S. P., Gerczuk, M., Batliner, A., Bergler, C., Amiriparian, S., Schuller, B., Nöth, E., & Riedhammer, K. (2023). Classification of stuttering – The ComParE challenge and beyond. *Computer Speech & Language*, 81, Article 101519. <https://doi.org/10.1016/j.csl.2023.101519>
- Bloodstein, O., & Bernstein Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Delmar. <http://dx.doi.org/10.1016/j.jfludis.2009.10.002>
- Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(4), 575–586. <https://doi.org/10.1037/0278-7393.12.4.575>
- Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.
- Camayd-Freixas, E. (2011). Cognitive theory of simultaneous interpreting and training. In *Proceedings of the 52nd Conference of the American Translators Association* (pp. 1–10). American Translators Association. https://www.academia.edu/303839/Cognitive_Theory_of_Simultaneous_Interpretation
- Cecot, M. (2001). Pauses in simultaneous interpretation: A contrastive analysis of professional interpreters' performances. *EUT - Edizioni Università di Trieste*. <http://hdl.handle.net/10077/2448>
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111. [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3)
- Collard, C., & Defrancq, B. (2019). Disfluencies in simultaneous interpreting, a corpus-based study with special reference to sex. In L. Vandevoorde, J. Daems, & B. Defrancq (Eds.), *New empirical perspectives on translation and interpreting* (pp. 36–56). Routledge.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321. <https://doi.org/10.1037/0033-295X.93.3.283>
- Downsub. (n.d.). *Downsub: Download subtitles from YouTube*. <https://downsub.com>
- Fromkin, V. A. (1971). The nonanomalous nature of anomalous utterances. *Language*, 47(1), 27–52. <https://doi.org/10.2307/412187>
- Fromkin, V. A. (Ed.). (1973). *Speech errors as linguistic evidence*. Mouton Publishers. <https://doi.org/10.1515/9783110888423>
- Fromkin, V. A. (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen and hand*. Academic Press. <https://doi.org/10.1017/S0022226700013785>
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133–177). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60270-4](https://doi.org/10.1016/S0079-7421(08)60270-4)
- Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (Rev. ed.). John Benjamins. <https://doi.org/10.1075/btl.8>
- Goldman-Eisler, F. (1958). Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology*, 10(2), 96–106. <https://doi.org/10.1080/17470215808416261>
- Grosjean, F. (1997). The bilingual individual. *Interpreting*, 2(1/2), 163–187. <https://doi.org/10.1075/intp.2.1-2.07gro>

- Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Blackwell.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press. <https://doi.org/10.1017/S0022226709990089>
- Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). John Wiley & Sons.
- Headlee, C. (2015). *10 ways to have a better conversation* [Video]. TED Conferences. https://www.ted.com/talks/celeste_headlee_10_ways_to_have_a_better_conversation
- Jansson-Verkasalo, E., Silvén, M., Lehtiö, I., & Eggers, K. (2020). Speech disfluencies in typically developing Finnish-speaking children – Preliminary results. *Clinical Linguistics & Phonetics*, 35(8), 707–726. <https://doi.org/10.1080/02699206.2020.1818287>
- Kosmala, L. (2024). *Beyond disfluency: The interplay of speech, gesture, and interaction*. John Benjamins. <https://doi.org/10.1075/ais.11>
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104. [https://doi.org/10.1016/0010-0277\(83\)90026-4](https://doi.org/10.1016/0010-0277(83)90026-4)
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press. <https://doi.org/10.7551/mitpress/6393.001.0001>
- Lin, Y., Lv, Q., & Liang, J. (2018). Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting. *Frontiers in Psychology*, 9, Article 1543. <https://doi.org/10.3389/fpsyg.2018.01543>
- MacLay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15(1), 19–44. <https://doi.org/10.1080/00437956.1959.11659682>
- MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, 8(3), 323–350. [https://doi.org/10.1016/0028-3932\(70\)90078-3](https://doi.org/10.1016/0028-3932(70)90078-3)
- MacKay, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3(2), 210–227. [https://doi.org/10.1016/0010-0285\(72\)90004-7](https://doi.org/10.1016/0010-0285(72)90004-7)
- MacKay, D. G. (1973). Complexity in output systems: Evidence from behavioral hybrids. *The American Journal of Psychology*, 86(4), 785–806. <https://doi.org/10.2307/1422086>
- Natarelli, G. (2018). *Anxiety and stuttering: Filling the research gaps in early adolescence* (Doctoral thesis, Università degli Studi di Padova). <https://www.research.unipd.it/handle/11577/3426700>
- Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.), *Nomen (Society): Leyden studies in linguistics and phonetics* (pp. 114–132). Mouton.
- Petite, C. (2005). Evidence of repair mechanisms in simultaneous interpreting: A corpus-based analysis. *Interpreting*, 7(1), 27–49. <https://doi.org/10.1075/intp.7.1.03pet>
- Pöschhacker, F. (1995). Slips and shifts in simultaneous interpreting. In J. Tömmola (Ed.), *Topics in interpreting research* (pp. 73–90). The University of Turku, Centre for Translation and Interpreting.
- Rahimpour, M., & Amini, D. (2013). The effect of frequent exposure to L2 vocabulary on speech production fluency. In S. Baleghizadeh (Ed.), *The handbook of current research on teaching*

- English language skills* (pp. 97-109). Shahid Beheshti University Publications. <https://espace.library.uq.edu.au/view/UQ:328349>
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Lawrence Erlbaum Associates.
- Shattuck-Hufnagel, S. (1982). Three kinds of speech error evidence for the role of grammatical elements in processing. In L. K. Obler & L. Menn (Eds.), *Exceptional language and linguistics* (pp. 133–142). Academic Press. <https://doi.org/10.1016/B978-0-12-523680-5.50014-5>
- Shattuck-Hufnagel, S., & Klatt, D. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior*, 18(1), 41–55. [https://doi.org/10.1016/S0022-5371\(79\)90554-1](https://doi.org/10.1016/S0022-5371(79)90554-1)
- Saad Merouwe, S., Bertram, R., & Eggers, K. (2024). Speech disfluencies in bilingual Lebanese children who do and do not stutter. *American Journal of Speech-Language Pathology*, 33(5), 2291–2310. https://doi.org/10.1044/2024_AJSLP-23-00311
- Shriberg, E. (1994). *Preliminaries to a theory of speech disfluencies* (Unpublished doctoral dissertation). University of California, Berkeley.
- Shriberg, E. (2001). To ‘errrr’ is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31(1), 153–169. <https://doi.org/10.1017/S0025100301001128>
- Song, S., & Li, D. (2020). The predicting power of cognitive fluency for the development of utterance fluency in simultaneous interpreting. *Frontiers in Psychology*, 11, Article 1864. <https://doi.org/10.3389/fpsyg.2020.01864>
- Speechace. (n.d.). *Speaking test*. <https://www.speechace.com/speaking-test/>
- Stes, R. (2004). *Vloeiendheidsstoornissen: Rationale eclecticische sociaal-cognitieve gedragstherapeutische benadering – Deel III: Normaal (on)vloeiende spreekcomponenten* [Fluency disorders: Rationale eclectic social-cognitive behavioral therapeutic approach – Part III: Normal (dis)fluent speech components]. Lecture at Postgraduate Stuttering Therapy, Antwerp.
- Tóth, A. (2011). Speech disfluencies in simultaneous interpreting: A mirror on cognitive processes. *SKASE Journal of Translation and Interpretation*, 5(2), 23–31. http://www.skase.sk/Volumes/JTI06/pdf_doc/03.pdf
- Zhao, N. (2022). Speech disfluencies in consecutive interpreting by student interpreters: The role of language proficiency, working memory, and anxiety. *Frontiers in Psychology*, 13, Article 881778. <https://doi.org/10.3389/fpsyg.2022.881778>
- Zhao, N., Cai, Z. G., & Dong, Y. (2023). Speech errors in consecutive interpreting: Effects of language proficiency, working memory, and anxiety. *PLOS ONE*, 18(10), Article e0292718. <https://doi.org/10.1371/journal.pone.0292718>