frontiers | Frontiers in Language Sciences

TYPE Original Research
PUBLISHED 26 January 2026
DOI 10.3389/flang.2025.1717090

Check for updates

OPEN ACCESS

EDITED BY
Seana Coulson,
University of California, San Diego,
United States

REVIEWED BY
Yuhua Liu,
South China Agricultural University, China
Jia Liu,
Hebei Normal University, China
Lingxi Fan,
Beihang University, China

*CORRESPONDENCE
Yue Lang
✉ langyue88@163.com

RECEIVED 01 October 2025
REVISED 24 December 2025
ACCEPTED 31 December 2025
PUBLISHED 26 January 2026

CITATION
Lang Y, Weissheimer J, Finger I and Kroll J (2026) Speech connectedness and linguistic complexity as predictors of Chinese—English bilinguals' interpreting performance. *Front. Lang. Sci.* 4:1717090. doi: 10.3389/flang.2025.1717090

Speech connectedness and linguistic complexity as predictors of Chinese—English bilinguals' interpreting performance

Yue Lang[1]*, Janaina Weissheimer[2], Ingrid Finger[3] and Judith Kroll[4]

[1]Ocean University of China, Qingdao, China, [2]Universidade Federal do Rio Grande do Norte, Natal, Brazil, [3]Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [4]University of California Irvine, Irvine, CA, United States

Interpreting provides a unique window into bilingual processing, as it requires the rapid transfer of meaning across languages under high temporal and cognitive constraints. This study investigates speech connectedness and linguistic complexity as predictors of interpreting performance in Chinese–English non-interpreter bilinguals. Consecutive interpreting data from 917 participants in the Parallel Corpus of Chinese EFL Learners (PACCEL) were analyzed using two computational tools: SpeechGraphs, to measure speech connectedness, and Coh-Metrix, to assess linguistic complexity. Results showed that higher interpreting scores were associated with greater long-range connectedness and fewer short-range recurrences, suggesting that more fluent and coherent speech reflects more efficient cognitive processing. Coh-Metrix indices further revealed that better-performing participants produced texts containing less frequent words, higher lexical specificity, and were more complex. Together, these findings indicate that measures of speech connectedness and linguistic complexity might serve as reliable predictors of interpreting competence in Chinese–English non-interpreter bilinguals. By integrating large-scale learner data with computational analyses, the study highlights how linguistic output can index underlying cognitive–linguistic mechanisms in bilingual interpreting.

KEYWORDS

consecutive interpreting, corpus, linguistic complexity, non-interpreter bilinguals, speech connectedness

1 Introduction

Interpreting is a cognitively demanding bilingual activity (Zhao et al., 2024) that requires the integration of linguistic and cognitive resources under real-time constraints (Gile, 2009). Although extensive research has examined specific cognitive abilities involved in interpreting (e.g., working memory, attentional control) and a range of linguistic skills relevant to comprehension and production, these two domains have largely been investigated independently (García, 2014). As a result, we still lack a clear understanding of how cognitive and linguistic mechanisms jointly shape interpreting output. This gap is particularly salient for non-interpreter bilinguals, a population that constitutes the majority of second language learners yet remains understudied. Their interpreting output represents early-stage processing before strategy acquisition and professional training, offering an important baseline for understanding skill development (Zhang et al., 2025). At the same time, recent advances in computational linguistics provide new opportunities to examine interpreting performance through quantitative indicators of cognitive processing (e.g., speech connectedness) and linguistic complexity (e.g., Coh-Metrix indices). However, the joint contribution of these two dimensions has not been systematically examined, especially using large-scale, ecologically valid corpus data from non-interpreter bilinguals.

Against this background, the present study investigates how cognitive and linguistic attributes collectively predict interpreting performance in Chinese–English non-interpreter bilinguals. Using interpreting output from the Parallel Corpus of Chinese EFL Learners, we integrate speech connectedness measures with linguistic complexity metrics to provide a comprehensive account of the cognitive-linguistic interface in early-stage interpreting.

Specifically, this study addresses the following research questions:

> Question 1. How are cognitive (speech connectedness) and linguistic (complexity) features associated with interpreting performance in Chinese–English non-interpreter bilinguals?
> Question 2. To what extent do these cognitive and linguistic features jointly predict interpreting performance in Chinese–English non-interpreter bilinguals?

2 Literature review

2.1 The relationship between cognitive and linguistic abilities in interpreting

Successful interpreting depends on a range of linguistic abilities, such as distinguishing speech from background noise (Elmer et al., 2014), comprehending sentences (Bajo et al., 2000; Yudes et al., 2013), detecting semantic errors (Fabbro et al., 1991), storing lexical items efficiently (Christoffels et al., 2006), and effectively transferring meaning from source to target (Chernov, 2004; Padilla et al., 2005). At the same time, these linguistic operations are supported by underlying cognitive resources including working memory (Liu et al., 2004; Aben et al., 2012), attentional control and shifting (Christoffels and de Groot, 2005; Nour et al., 2020).

Empirical findings further demonstrate the interdependence between linguistic and cognitive abilities in interpreting. Some studies show that interpreters' advantages emerge most clearly in linguistic abilities that are directly involved in interpreting tasks, rather than in broader executive-function abilities (Santilli et al., 2019; Babcock et al., 2017), while other studies highlight their strengths in cognitive abilities implicated in interpreting tasks (Bajo et al., 2000; Yudes et al., 2011). Importantly, linguistic skills such as translation efficiency and cognitive skills such as working memory act as separable but complementary contributors to interpreting performance (Christoffels et al., 2003; García, 2014). Collectively, these studies suggest that interpreting performance is shaped by the joint influence of linguistic competence and cognitive processing.

Understanding these mechanisms in non-interpreter bilinguals is theoretically significant because their performance reflects early-stage skill development, free from the confounds of professional training (García et al., 2014). Examining this population provides a baseline for interpreting competence and responds to calls for greater attention to variability in bilingual experience (De Bruin, 2019; Kroll et al., 2021). Moreover, focusing on non-interpreter bilinguals minimizes potential self-selection biases associated with individuals who actively pursue interpreter training, thereby offering clearer insight into how interpreting-related abilities emerge prior to formal specialization.

To advance this line of inquiry, it is essential to use datasets that are sufficiently robust and ecologically valid. However, most prior experimental studies have relied on small, laboratory-based samples, limiting the generalizability of their findings (Macnamara et al., 2011). To address this gap, the present study draws on a large-scale corpus of Chinese–English non-interpreter bilinguals' oral production elicited during an interpreting task. This corpus-based approach allows for a more comprehensive examination of how cognitive features (indexed through speech connectedness) and linguistic features (indexed through Coh-Metrix measures) jointly contribute to interpreting performance.

2.2 Speech connectedness as a proxy of cognitive processing

The development of computational methods has opened new avenues for examining oral and written discourse production. Within the Speech Graphs paradigm, speech is represented as a graph in which words are nodes and transitions are edges, enabling the quantification of structural connectedness (Mota et al., 2012, 2014). Metrics reflect short-range recurrences (e.g., word repetitions) and long-range recurrences (connections across distant parts of discourse), which together capture the degree of speech connectedness. These indices have been shown to reflect underlying cognitive processes in both typical and atypical populations (Mota et al., 2023). In addition, graph structures preserve the online flow of speakers' word-by-word decisions, offering a direct window into real-time discourse organization.

A growing body of research has been devoted to developing innovative approaches to investigate aspects of the bilingual experience that shape cognitive and linguistic processing. In this context, graph analysis has emerged as a promising tool for mapping the relationship between language and cognition. For example, Botezatu et al. (2021) showed that speech connectedness correlates with verbal fluency in learners of Spanish and Chinese as second languages, suggesting that higher levels of connectedness reflect more advanced proficiency. However, despite these encouraging findings, applications of graph-based approaches to interpreting remain rare. The present study therefore advances this line of research by applying graph analysis directly to interpreting performance in bilinguals, a domain that places particularly high demands on both cognitive processing and linguistic competence.

2.3 Coh-Metrix as a tool to measure linguistic performance

Coh-Metrix is a computational system originally designed to assess cohesion and coherence in written text (Graesser et al., 2004). Its most recent version automatically generates a total of 108 indices characterizing linguistic features at various levels of language, discourse, meaning and conceptual analysis (Graesser and McNamara, 2011). The specific description of the indices can be retrieved from Coh-Metrix version 3.0 indices (memphis.edu). Generally, these indicators can be categorized into three levels, namely surface code, textbase code, and situation model, representing the exact form of the text, the propositional network of the text and the state of affairs referred to by the text (Zwaan, 1994). These indices have been widely used to examine readability (Crossley et al., 2008), syntactic difficulty (Lu, 2017), writing quality (Latifi and Gierl, 2021), and language proficiency (Crossley et al., 2012), among others.

Although originally designed for written texts, Coh-Metrix has also been used to analyze oral production, capturing features such as lexical diversity, word frequency, and cohesion (Crossley et al., 2011). Importantly, it has demonstrated explanatory power for the analysis of speaking proficiency (Crossley and McNamara, 2011, 2013) and, more recently, for interpreting quality. Ouyang et al. (2021), analyzing data from All China Interpreting Competition, showed that Coh-Metrix indices such as lexical diversity, verb hypernyms, and word frequency accounted for 60% of the variance in human ratings. These findings suggest that linguistic measures alone can capture a substantial portion of interpreting performance variability.

Building on this work, the present study applies SpeechGraphs and Coh-Metrix to interpreting output to systematically assess how linguistic complexity interacts with speech connectedness. Together, these tools allow us to explore the cognitive-linguistic interface in bilingual interpreting.

2.4 Linking speech connectedness and linguistic complexity to interpreting ability

Although linguistic features (Coh-Metrix) and speech connectedness (SpeechGraphs) of oral output have been examined separately, interpreting performance relies on the combined effect of cognitive processing and linguistic competence. Under a heavy time pressure, bilinguals are unable to polish their renderings (Shlesinger and Malkiel, 2005). Therefore, the interpreting outputs as the final results of their decision making are more likely to reflect bilinguals' default processing pattern, making it an informative source for assessing underlying cognitive-linguistic mechanisms.

From a cognitive perspective, higher level of speech connectedness reflects more efficient discourse planning and information retrieval under time pressure. These skills are essential for maintaining coherent output during interpreting. Conversely, less connected speech may signal difficulty in managing cognitive load, leading to incoherence or fragmentation.

From a linguistic perspective, Coh-Metrix captures linguistic sophistication through indicators such as lexical diversity, syntactic complexity, cohesion markers, and semantic richness. Interpreters with stronger linguistic skills tend to produce more coherent, varied, and context-appropriate output. Thus, Coh-Metrix provides a granular profile of linguistic performance that underlies interpreting quality.

Together, SpeechGraphs and Coh-Metrix index complementary dimensions of interpreting performance. Their combined use allows the identification of patterns of linguistic-cognitive synergy that characterize more proficient interpreting.

3 Methods

The procedures and apparatus we have employed in the present study are illustrated in Figure 1 and described in the following sections.

3.1 Corpus

The present study used the Parallel Corpus of Chinese EFL Learners (PACCEL) as the primary data source. PACCEL is an open-access corpus containing translation and interpreting discourse produced from 2003 to 2007 by Chinese English learners during the Test for English Majors-Band 8 (TEM-8), a nationwide English proficiency test for senior undergraduates majoring in English (Wen and Wang, 2008). Although these students may have taken introductory translation or interpreting courses as part of their undergraduate curriculum, such courses were designed primarily for general language learning rather than for developing professional interpreting skills. They did not receive systematic or specialized interpreter training. Therefore, the speakers in PACCEL can be appropriately classified as Chinese–English non-interpreter bilinguals.

The corpus includes recordings of both English–Chinese and Chinese–English interpreting tasks as well as transcriptions of these recordings. All recordings were evaluated by human raters, who provided comprehensive performance scores based on accuracy, completeness of information transmission, pronunciation, word choice, and grammatical correctness (see Wen and Wang, 2008). Each transcription is tagged with metadata, including test year, task type (spoken or written), gender, and score. Errors in grammar, pronunciation, disfluencies, fillers, and repetitions are preserved and annotated, though the annotation scheme does not differentiate between error types. Accordingly, all annotated errors were analyzed uniformly in this study.

Given that the computational tools selected (SpeechGraphs and Coh-Metrix) could only process English texts, only Chinese–English interpreting data were included in the analysis. The initial dataset consisted of transcriptions from 930 bilingual participants; after excluding 13 cases due to missing data, 917 valid transcriptions were retained for analysis (see Table 1).

3.2 SpeechGraphs

SpeechGraphs (https://neuro.ufrn.br/softwares/speechgraphs) is an open-access software tool designed for analyzing text characteristics through graph-theoretical measures. It generates various indices of speech connectedness, such as the largest connected component (LCC), the largest strongly connected component (LSC), and short-recurrence attributes such as parallel

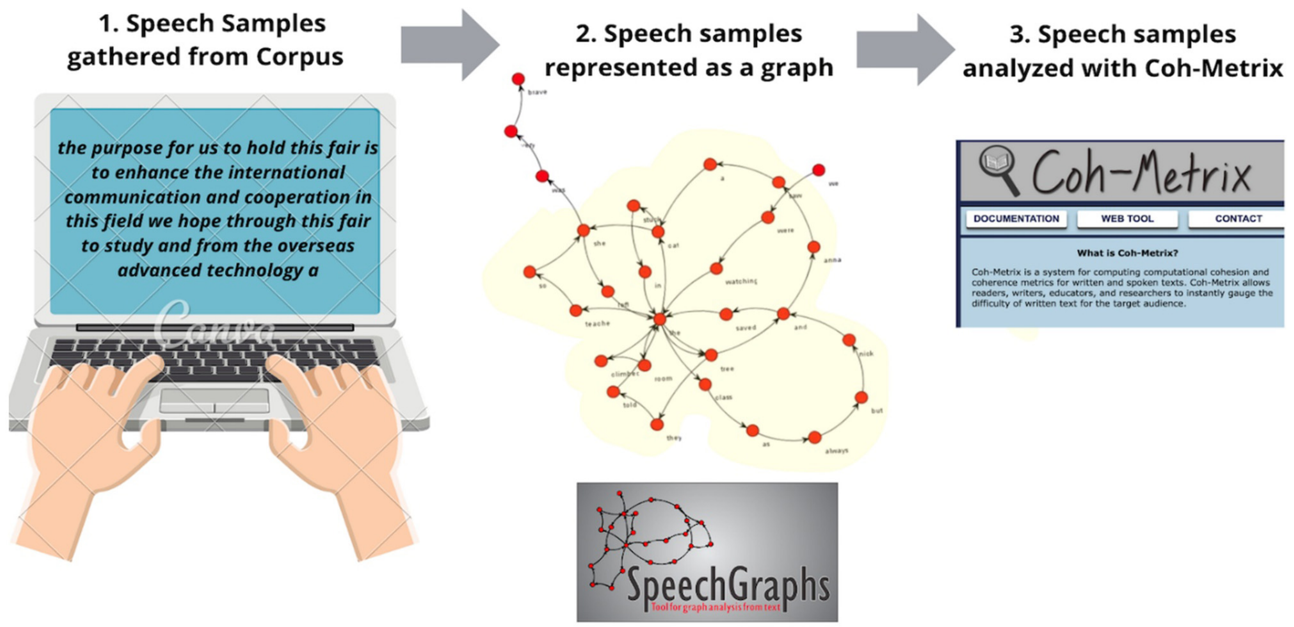

FIGURE 1
The procedures and apparatus of our study. (1) Speech samples of 917 participants were collected from the Parallel Corpus of Chinese EFL Learners (PACCEL); (2) Speech samples were analyzed with SpeechGraphs, including LCC, LSC, PE and RE as attributes; (3) The same speech samples were submitted to Coh-Metrix, considering surface code, textbase, and situation model as the main levels.

TABLE 1 Basic information of the corpus.

Year	Number of Texts	Number of Words
2003	180	26,494
2004	188	28,881
2005	187	32,721
2006	174	31,386
2007	188	30,740

edges (PE) and repeated edges (RE). This tool has proven effective for the analysis of orally produced texts (Mota et al., 2014, 2016).

In the present study, all annotated elements were removed from the oral samples prior to analysis. Only raw transcriptions in the target language (English) were retained to ensure that the measures reflected genuine features of speech production. Because SpeechGraphs does not support punctuation marks, all punctuation was also removed before analysis.

3.3 Coh-Metrix

Coh-Metrix (http://cohmetrix.memphis.edu/cohmetrixhome/) was employed to evaluate how linguistic features contribute to human judgments of interpreting performance. Transcribed texts were submitted to the Coh-Metrix platform, and the resulting indices were downloaded for further statistical analysis. Since spoken language does not naturally contain paragraphs, indices related to paragraph-level features (e.g., number of paragraphs, paragraph length) were excluded (Ouyang et al., 2021), leaving 100 indices for analysis.

Following Ouyang et al. (2021), the retained indices were grouped into three levels: surface code, textbase, and situation model. These correspond respectively to lexical and syntactic features, cohesion and coherence features, and conceptual/thematic information. The surface code reflects the most superficial level of representation (word- and syntax-based features); the textbase captures semantic content through networks of propositions extracted from the text's inherent meaning, independent of prior knowledge; and the situation model integrates textual information with the reader's or listener's prior knowledge (Fletcher and Chrysler, 1990). Coh-Metrix operationalizes these levels into specific linguistic indices and scales them to allow quantitative analysis.

4 Data analysis and results

4.1 Correlation analysis

We first conducted correlation analyses among the indices generated by Coh-Metrix. A total of 62 indices were found to be significantly correlated with interpreting performance (see Table 2), including 29 surface-level indices, 13 textbase indices, and 20 situation model indices. We also examined correlations between the speech connectedness indices (LCC, LSC, PE, RE) and interpreting performance, as well as between the scaled number of errors and interpreting performance. All indices demonstrated significant relationships with interpreting performance (see Table 2). From these, we selected the 20 most strongly correlated indices for further analysis. These indices covered multiple linguistic constructs, including readability (Flesch–Kincaid Grade Level [RDFKGL], Flesch Reading Ease

TABLE 2 Results of the correlation analysis between the interpreting score and linguistic indices, speech connectedness, and the number of errors.

Index	Correlation	*p*	Index	Correlation	*P*
Surface Code					
DESSC	−0.171	1.80E−07	WRDADJ	−0.139	2.34E−05
DESWC	0.195	2.70E−09	WRDPRP1s	−0.222	1.02E−11
DESSL	0.284	1.90E−18	WRDPRP1p	0.121	0.000
DESSLd	0.170	2.09E−07	WRDPRP2	−0.130	8.14E−05
DESWLsy	0.283	2.62E−18	WRDPRP3s	−0.109	0.001
DESWLsyd	0.307	1.85E−21	WRDFRQc	−0.374	8.30E−32
DESWLlt	0.340	2.76E−26	WRDFRQa	−0.250	1.45E−14
DESWLltd	0.302	7.83E−21	WRDFAMc	−0.352	3.42E−28
DRAP	0.112	0.001	WRDMEAc	−0.109	0.001
DRPP	0.099	0.003	WRDHYPn	0.232	1.16E−12
DRNEG	−0.09	0.006	WRDHYPv	0.200	1.01E−09
DRINF	0.167	3.84E−07	WRDHYPnv	0.236	5.06E−13
DRVP	0.072	0.030	LDTTRc	0.196	2.19E−09
LDMTLD	0.265	3.27E−16	LDTTRa	0.140	2.03E−05
LDVOCD	0.315	1.24E−22			
Textbase Code					
CRFCWO1	−0.182	2.72E-08	CNCAll	0.200	9.22E−10
CRFCWO1d	−0.099	0.003	CNCLogic	0.127	0.000
CRFCWOa	−0.161	9.60E-07	CNCADC	0.133	5.41E−05
CRFCWOad	−0.102	0.002	CNCTempx	0.129	8.70E−05
CRFANP1	0.222	9.59E-12	CNCAdd	0.209	1.76E−10
CRFANPa	0.190	7.17E-09	CNCPos	0.208	2.15E−10
CNCNeg	0.113	0.001			
Situation Model					
LSASS1	−0.158	1.53E−06	PCSYNp	−0.178	5.89E−08
LSAGN	−0.179	4.67E−08	PCCNCz	0.193	3.87E−09
LSAGNd	−0.152	3.54E−06	PCCNCp	0.217	2.99E−11
RDFRE	−0.452	2.56E−47	PCREFz	−0.154	2.82E−06
RDFKGL	0.421	9.40E−41	PCDCz	0.090	0.007
RDL2	−0.316	9.33E−23	PCDCp	0.067	0.044
SMINTEr	0.066	0.046241	PCVERBz	−0.192	5.39E−09
SMCAUSlsa	−0.234	7.84E−13	PCCONNz	−0.197	5.65E−09
SMTEMP	−0.143	1.47E−05	PCCONNp	−0.189	7.74E−09
PCTEMPp	−0.131	7.00E−05	PCTEMPz	−0.137	3.05E−05
Speech Connectedness					
LSC	0.294	2.2e−16	PE	−0.270	2.2e−16
LCC	0.345	2.2e−16	RE	−0.242	1.035e−13
Errors					
Errors	−0.068	0.037			

[RDFRE], Coh-Metrix L2 Readability [RDL2]), word length (mean number of letters [DESWLlt], standard deviation of letters per word [DESWLltd], mean number of syllables [DESWLsyd], standard deviation of syllables per word [DESWLsy]), lexical diversity (VOCD [LDVOCD] and MTLD [LDMTLD]), sentence length (mean number of words per sentence [DESSL]), lexical specificity (hypernymy for nouns and verbs [WRDHYPnv] and for nouns only [WRDHYPn]), word frequency (CELEX frequency for content words [WRDFRQc] and all words [WRDFRQa]), word familiarity (content-word familiarity [WRDFAMc]), concreteness (content-word concreteness [PCCNCp]), and cohesion (latent semantic analysis causal cohesion [SMCAUSlsa], anaphor overlap [CRFANP1], and additive connectives [CNCAdd]). The full definitions of these indices are provided in Table 3.

TABLE 3 Detailed description of features of Coh-Metrix Indices.

Coh-Metrix dimension/Index	Description
RDFKGL	Flesch-kincaid grade level
DESWLlt	Word length, number of letters, mean
LDVOCD	Lexical diversity, VOCD, all words
DESWLsyd	Word length, number of syllables, standard deviation
DESWLltd	Word length, number of letters, standard deviation
DESSL	Sentence length, number of words, mean
DESWLsy	Word length, number of syllables, mean
LDMTLD	Lexical diversity, MTLD, all words
WRDHYPnv	Hypernymy for nouns and verbs, mean
WRDHYPn	Hypernymy for nouns, mean
RDFRE	Flesch reading ease
WRDFRQc	CELEX word frequency for content words, Mean
WRDFAMc	Familiarity for content words, mean
RDL2	Coh-Metrix L2 Readability
WRDFRQa	CELEX Log frequency for all words, mean
SMCAUSlsa	LSA verb overlap
WRDPRP1s	First person singular pronoun incidence
CRFANP1	Anaphor overlap, adjacent sentences
PCCNCp	Text Easability PC Word concreteness,Percentile
CNCAdd	Additive connectives incidence

4.2 Multiple regression model

Building on the correlation results, we applied a multiple regression model to evaluate the predictive power of both cognitive and linguistic attributes. In addition to the four speech connectedness indices (LCC, LSC, PE, RE), the number of errors was also included. Prior to model construction, all data were standardized using scaling and logarithmic transformation. We initially conducted a stepwise regression analysis, which retained 16 linguistic indices and three speech connectedness attributes (LCC, LSC, RE). Multicollinearity diagnostics were then performed, and six indices with high variance inflation factor (VIF) values were excluded. The final model included 13 predictors: 10 linguistic indices (LDVOCD, DESWLsy, WRDHYPnv, WRDFAMc, WRDFRQa, WRDPRP1s, CRFANP1, CNCAdd, PCCNCp, RDFRE), number of errors, and two speech connectedness indices (LSC, RE). The model explained a substantial portion of the variance in interpreting performance, with $R^2 = 0.45$ (see Table 4). This implies that the model accounts for approximately 45% of the variations in interpreting task outcomes.

4.3 Partial least square structural equation modeling

After establishing direct associations among individual indicators, we employed partial least squares structural equation modeling (PLS-SEM) to investigate how latent constructs—comprising linguistic and speech connectedness measures—jointly contribute to interpreting performance. Unlike covariance-based SEM (CB-SEM), which focuses on model fit, PLS-SEM maximizes explained variance of dependent variables using an ordinary least squares estimation method (Fornell and Bookstein, 1982). PLS-SEM is particularly suitable for exploratory research where theoretical frameworks are less developed, for smaller sample sizes, or for data that deviate from normality (Hair et al., 2019). Furthermore, the constructs in our study are formative: each indicator captures unique aspects of the construct and cannot be interchanged (Hair and Alamer, 2022).

The analysis was performed using the plspm package in R (version 4.2.1). Based on Section 3.3, the formative constructs included surface code (six indicators: DESWLsy, WRDFRQa, WRDFAMc, LDVOCD, WRDHYPnv, WRDPRP1s), textbase (two indicators: CNCAdd, CRFANP1), and situation model (two indicators: RDFRE, PCCNCp). Performance and Errors were treated as single-item constructs, with construct scores identical to their observed values. Speech connectedness was operationalized using LSC (long-range connectedness) and RE (short-range repetitions).

Prior to the analysis, indicators negatively correlated with performance were rescaled by multiplying values by −1 to align orientation. These were renamed with an "*n*" prefix (e.g., RE → nRE) for interpretability. Multicollinearity was assessed via VIF values, all of which were < 3, suggesting no collinearity issues. A bootstrap procedure with 5,000 iterations was conducted to ensure robustness. Relative contributions were evaluated by examining the *p*-values of outer weights. Although three indicators (DESWLsy, nWRDFRQa, nRE) did not show significant weights, their loadings were significant (≥ 0.5) and thus retained. Notably, nWRDFRQa's loading was 0.498, slightly below the conventional threshold, but it was also retained given its theoretical importance in capturing word frequency.

The resulting structural model is presented in Figure 2. All path coefficients were statistically significant (see Table 5). Detailed

TABLE 4 Results for multiple regression analysis.

Predictors	Estimates	Std Error	t-value	P
(Intercept)	4.20	0.005	734.3331	$< 2e{-}16$ ***
Surface level				
LDVOCD	0.056	0.007	7.972	4.70e−15 ***
DESWLsy	−0.1010	0.011	−8.936	$< 2e{-}16$ ***
WRDHYPnv	0.0257	0.007	3.652	0.000 ***
WRDFAMc	−0.041	0.007	−5.824	7.96e−09 ***
WRDFRQa	0.007	0.007	1.006	0.314
WRDPRP1s	−0.017	0.006	−1.866	0.062
Textbase				
CNCAdd	0.014	0.006	2.2597	0.024*
CRFANP1	0.0171	0.007	2.619	0.009*
Situation model				
PCCNCp	0.021	0.006	3.462	0.000***
RDFRE	−0.147	0.010	−14.694	$< 2e{-}16$ ***
Speech Connectedness				
LSC	0.059	0.009	6.408	2.36e−10 ***
RE	0.010	0.010	1.106	0.269
Errors				
serror	−0.018	0.006	−3.148	0.002 **

Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

measurement-model indicator weights and loadings are reported in Tables 6 and 7.

Specifically, Errors (β = 0.07, 95% CI [0.02, 0.12]) showed a modest effect, while Surface Code (β = 0.24, 95% CI [.19, 0.32]), Textbase (β = 0.17, 95% CI [.11, 0.22]), and Speech Connectedness (β = 0.12, 95% CI [0.05, 0.20]) showed moderate effects. The Situation Model exerted the strongest influence (β = 0.35, 95% CI [0.29, 0.41]). The model's predictive power, as assessed by R^2 = 0.389, indicates a moderate explanatory capacity in the context of second language research (Hair and Alamer, 2022).

5 Discussion

5.1 Question 1. how are cognitive (speech connectedness) and linguistic (complexity) features associated with interpreting performance in Chinese–English non-interpreter bilinguals?

To address our first research question, we analyzed cognitive and linguistic attributes separately in relation to interpreting performance using correlation and regression analyses (Figure 3).

5.1.1 Speech Graph analysis

Participants who produced more long-range connectedness (LSC) achieved higher interpreting scores, whereas those who produced more short-range repetitions (RE) performed more poorly. This aligns with previous work demonstrating that greater connectedness and fewer repetitions reflect higher oral second language production proficiency (Botezatu et al., 2021). The negative association between error counts and interpreting performance further supports the long-established practice of treating errors as a key indicator of interpreting output quality (Gile, 1995).

5.1.2 Coh-Metrix indices analysis

In the analysis of linguistic features, the results show that, among the 100 Coh-Metrix indices, 62 showed significant correlations with interpreting performance (29 surface code indices, 13 textbase indices, 20 situation model indices). Through an exploratory refinement procedure, ten indices (LDVOCD, DESWLsy, WRDHYPnv, WRDFAMc, WRDFRQa, WRDPRP1s, CRFANP1, CNCAdd, PCCNCp, RDFRE) were retained as meaningful predictors (for details, see Figure 3).

(1) Surface Level

At the surface level, indicators of word length in terms of syllables (DESWLsy), lexical diversity as measured by vocd algorithm (LDVOCD), specificity of word use as measured by the hypernym of noun and verbs (WRDHYPnv) were found to positively correlate with interpreting performance, while indicators related to the incidence of first-person pronouns (WRDPRP1s), word familiarity (WRDFAMc) and average word frequency for all words (WRDFRQa) were negatively associated with interpreting performance. These patterns converge in highlighting the importance of a broad and sophisticated lexical repertoire for successful interpreting.

First, we found that Chinese–English bilinguals with better performance tend to produce longer words, which are expected to express more complex meanings (Bergen et al., 2012) and concepts (Lewis and Frank, 2016). In interpreting tasks, where the goal is to convey the informational content of the source under time pressure, the use of longer lexical forms allows speakers to transmit denser semantic content. This aligns with Han's (2016) claim that information completeness is central to interpreting quality. However, in the multiple regression model, word length in syllables showed a negative contribution to interpreting performance once other linguistic predictors were controlled. This indicates that word length *per se* does not independently enhance interpreting performance.

Second, lexical diversity and lexical specificity emerged as positive predictors of interpreting performance. Bilinguals who achieved higher scores demonstrated access to a wider vocabulary range and were able to select more specific, semantically informative lexical items. This finding is consistent with Ouyang et al. (2021), who showed that lexical richness is a strong predictor of interpreting quality, and resonates with research on explicitation in interpreting (Gumul, 2006). According to this line of work, interpreters frequently replace a source expression with a more explicit or specific target-language form (Perego, 2003). For learners, the ability to use explicit and specific language in

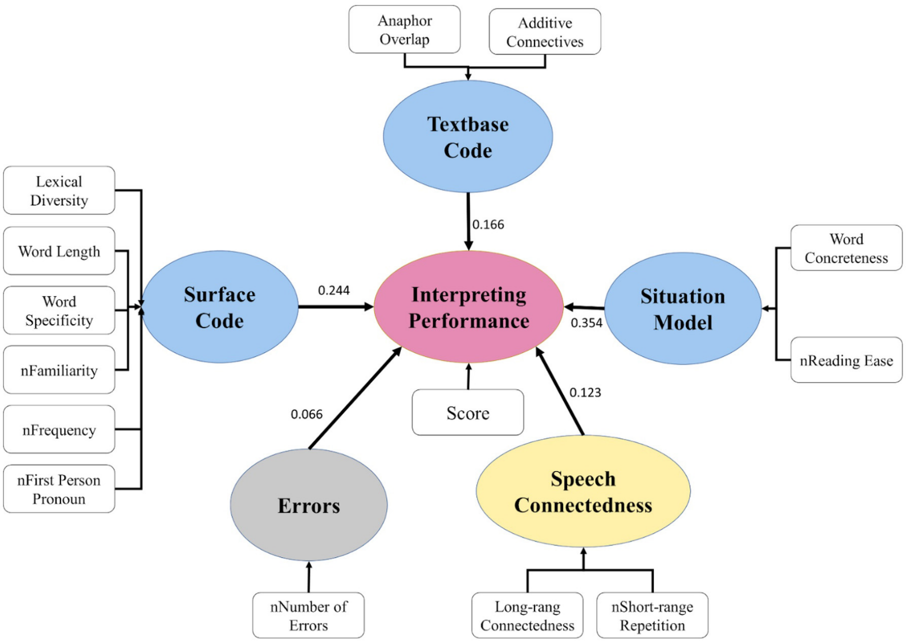

FIGURE 2
The structural model.

TABLE 5 Results of bootstrapped paths.

Path	Original	Mean.Boot	Std.Error	perc.025	perc.975
Surface -> Performance	0.2441	0.2506	0.0331	0.1856	0.316
Text -> Performance	0.1661	0.1664	0.0274	0.1121	0.219
Situation -> Performance	0.3536	0.3511	0.0286	0.2947	0.406
Connectedness -> Performance	0.1230	0.1246	0.0360	0.0526	0.196
Error -> Performance	0.0657	0.0653	0.0260	0.0161	0.118

consecutive interpreting has also been interpreted as evidence of deeper comprehension and increasing audience orientation through training (Li and Dong, 2022). Because specific lexical items tend to be less frequent (Kim et al., 2018), the present finding that bilinguals with better interpreting performance tend to use less familiar and less frequent words reinforces the interpretation that these speakers possess richer and more accessible lexical resources (Crossley et al., 2009). The ability to retrieve low-frequency words is cognitively demanding (Meara, 2005), and therefore serves as a sign of rich linguistic resources in their memory, which is a good indicator of interpreting ability.

Finally, the negative association between interpreting performance and first-person pronoun use (WRDPRP1s) suggests that greater reliance on self-referential expressions may reflect reduced attentional focus on the source message. As noted by Ouyang et al. (2021), frequent use of first-person pronouns may signal intrusions of the speaker's own perspective or hesitations, which could disrupt the flow and accuracy of interpreting.

Taken together, the surface-level results reveal a coherent pattern: high-quality interpreting is associated with the ability to mobilize a broad, precise, and semantically rich lexicon, while reduced reliance on frequent, familiar, or self-focused lexical items appears to reflect more limited linguistic resources and less effective message reformulation.

(2) Textbase Level

At the textbase level, the additive connectives (CNCAdd) and the anaphor overlap of adjacent sentences (CRFANP1) were positive indicators of a bilingual's interpreting ability.

The positive effect of anaphor overlap suggests that bilinguals with better interpreting performance are better able to preserve referential ties across clauses and sentences. By maintaining continuity of reference, interpreters help audiences track follow the narrative or argument more easily, enhancing their comprehension. This aligns with prior work showing that referential cohesion is a key determinant of interpretive clarity and discourse coherence (Graesser and McNamara, 2011; Gieshoff and Albl-Mikasa, 2024).

TABLE 6 Results of bootstrapped weights.

Variable	Original	Mean.Boot	Std.Error	perc.025	perc.975
Surface-DESWLsy	0.0448	0.0457	6.96e−02	−0.0913	0.183
Surface-nWRDFRQa	−0.0339	−0.0350	8.28e−02	−0.1955	0.129
Surface-nWRDFAMc	0.5098	0.5041	6.53e−02	0.3751	0.628
Surface-LDVOCD	0.5599	0.5549	6.45e−02	0.4232	0.676
Surface-WRDHYPnv	0.2210	0.2196	7.32e−02	0.0756	0.363
Surface-nWRDPRP1s	0.4042	0.3999	6.22e−02	0.2765	0.519
Text-CNCAdd	0.6022	0.5976	1.05e−01	0.3744	0.786
Text-CRFANP1	0.6680	0.6629	9.81e−02	0.4608	0.840
Situation-nRDFRE	0.9006	0.9001	3.76e−02	0.8215	0.965
Situation-PCCNCp	0.4272	0.4239	5.65e−02	0.3079	0.530
Connectedness-LSC	0.8543	0.8490	1.28e−01	0.5820	1.083
Connectedness-nRE	0.1869	0.1867	1.55e−01	−0.1161	0.491
Error-nserror	1.0000	1.0000	1.22e−16	1.0000	1.000
Performance-logscore	1.0000	1.0000	1.05e−16	1.0000	1.000

TABLE 7 Results of bootstrapped loadings.

Variable	Original	Mean.Boot	Std.Error	perc.025	perc.975
Surface-DESWLsy	0.563	0.558	4.76e−02	0.462	0.649
Surface-nWRDFRQa	0.498	0.493	6.36e−02	0.367	0.612
Surface-nWRDFAMc	0.702	0.694	4.76e−02	0.596	0.781
Surface-LDVOCD	0.628	0.622	5.95e−02	0.498	0.732
Surface-WRDHYPnv	0.469	0.464	6.41e−02	0.334	0.585
Surface-nWRDPRP1s	0.442	0.437	6.72e−02	0.303	0.565
Text-CNCAdd	0.761	0.755	8.21e−02	0.581	0.894
Text-CRFANP1	0.811	0.805	7.42e−02	0.645	0.932
Situation-nRDFRE	0.904	0.904	2.64e−02	0.848	0.951
Situation-PCCNCp	0.435	0.429	7.91e−02	0.268	0.573
Connectedness-LSC	0.992	0.986	1.60e−02	0.942	1.000
Connectedness-nRE	0.816	0.812	6.14e−02	0.681	0.919
Error-nserror	1.000	1.000	1.11e−16	1.000	1.000
Performance-logscore	1.000	1.000	1.11e−16	1.000	1.000

The role of additive connectives is more nuanced. In written discourse, additive connectives are generally considered less informative about the unfolding text and may even hinder comprehension (Kleijn et al., 2019). However, interpreting imposes constraints that differ fundamentally from written processing. Because speakers must reformulate incoming information under severe time pressure, additive connectives offer a cognitively economical way to maintain the flow of speech without engaging in substantial syntactic restructuring. Their use enables interpreters to link ideas smoothly, minimize pauses and disfluencies, and sustain an uninterrupted information flow, which are features that are essential for audience comprehension in oral interpreting. Thus, although additive connectives are not highly informative, they serve an important functional role interpreting production. Their frequent presence in bilinguals' output likely reflects a strategic adaptation to the cognitive demands of interpreting: they prioritize maintaining fluency and local cohesion, ensuring that the message unfolds in a continuous and accessible manner for the target-language audience.

Together, the two textbase-level indicators suggest that effective interpreting depends not only on lexical sophistication but also on the ability to construct locally cohesive discourse. These cohesion-building skills become especially important in interpreting, where rapid processing leaves little opportunity for revision or restructuring.

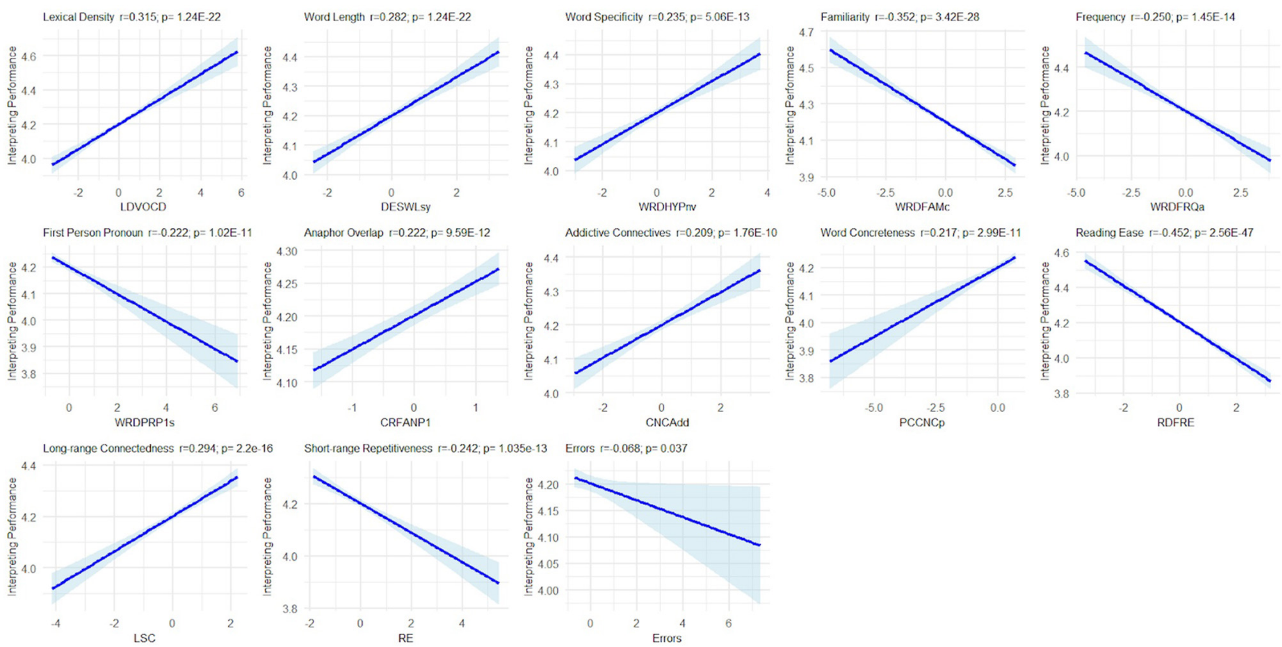

FIGURE 3
Relationships of linguistic attributes and speech connectedness with interpreting performance.

(3) Situation Model Level

At the situation model level, our findings align with Ouyang et al. (2021) in showing that bilinguals' interpreting performance was negatively associated with readability (RDFRE), as indexed by Flesch Reading Ease scores. In other words, participants who produced linguistically simpler and therefore more readable output tended to receive lower interpreting scores. Although seemingly counterintuitive, this pattern likely reflects characteristics of non-interpreter bilinguals' data. As the test aims to assess English proficiency, bilinguals who produced output with greater lexical and syntactic complexity (and thus lower readability) tended to achieve higher interpreting scores, as higher linguistic complexity is widely recognized as an indicator of better second language proficiency (Crossley and McNamara, 2011).

In addition, word concreteness was found to be a positive indicator of interpreting performance. This result diverges from Crossley et al. (2009), who reported that speakers with higher lexical proficiency tended to use more abstract words. However, it converges with findings from Ouyang et al. (2021) and may reflect features specific to interpreting. Because interpreting often involves real-time meaning clarification and audience-oriented reformulation, interpreters may prefer more concrete lexical choices to enhance explicitness and ensure message transparency. This aligns with explicitation tendencies observed in interpreting research (Li and Dong, 2022).

Together, these findings suggest that situation model features are crucial for interpreting performance among non-interpreter bilinguals. Lower readability scores indicate the use of more advanced linguistic structures, while higher concreteness reflects efforts to convey information explicitly and clearly. The greater linguistic sophistication and increased explicitness capture the dual demands of interpreting at the conceptual level, where interpreters must both understand the communicative intent and render it in a way that is accessible to audiences under tight processing constraints.

5.1.3 Summary

The findings for RQ1 show that cognitive and linguistic features contribute to interpreting performance in distinct yet complementary ways. On the cognitive side, greater long-range connectedness and fewer short-range repetitions were associated with higher interpreting performance, indicating that bilinguals who can retrieve and organize information efficiently under time pressure produce more coherent output. Error rates also behaved as expected, with more errors predicting poorer performance.

On the linguistic side, features across all three Coh-Metrix levels meaningfully differentiated performance. Surface code indicators such as higher lexical diversity and greater lexical specificity characterized bilinguals with better interpreting performance, whereas frequent use of familiar, highly frequent words and first-person pronouns predicted weaker performance. At the textbase level, cohesive devices, particularly additive connectives and anaphor overlap, were positively associated with interpreting performance, underscoring the importance of maintaining local coherence in real time. Situation-model attributes also contributed: lower readability scores and higher concreteness were both linked to better performance, consistent with the dual demands of linguistic sophistication and explicit meaning construction in interpreting.

Taken together, these results demonstrate that interpreting performance reflects a synergy between real-time cognitive processing and linguistic resource deployment. Speech connectedness indexes the efficiency with which bilinguals manage information flow, while Coh-Metrix indices capture the sophistication and cohesion of the linguistic output. The convergence of these cognitive and linguistic pathways provides empirical support for the multifactorial nature of interpreting competence in non-interpreter bilinguals.

5.2 Question 2. To what extent do these cognitive and linguistic features jointly predict interpreting performance in Chinese–English non-interpreter bilinguals?

To address the second research question, we employed partial least squares structural equation modeling to examine the combined predictive power of cognitive and linguistic features.

The PLS-SEM results showed that all five constructs significantly predicted interpreting performance, but their relative contributions differed in strength. Among the linguistic predictors, the situation model exerted the strongest standardized effect ($\beta = 0.35$), followed by surface code ($\beta = 0.24$) and textbase features ($\beta = 0.17$). Speech connectedness also contributed to performance, although its influence was comparatively modest ($\beta = 0.12$), and error counts showed only a small effect ($\beta = 0.07$).

Although speech connectedness exerted a smaller effect than linguistic features, its role remains meaningful. In our study, speech connectedness was measured by long-range recurrences and short-range repetitions in the produced output. In this sense, a comparison of two illustrative Chinese–English bilinguals from our sample (see Figure 4) shows that the participant with a high interpreting score (on the right) was able to produce more long-range connectedness and fewer short-range repetitions (Example 2). In contrast, the participant with a low interpreting score (on the left) seems to struggle to produce longer utterances and repeats words frequently (Example 1). For the bilingual depicted on the left, the shorter sentences, lack of connectives and overlap of content words seem to be indicators of their limited vocabulary size, which might contribute to their lower interpreting score.

These observations underscore that speech connectedness seems to capture an important dimension of real-time cognitive processing, reflecting "the rapid, smooth, accurate, lucid and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing" (Lennon, 2000, p. 26). However, in interpreting, a bilingual's output is not solely the translation of their own thought or intention; it is highly constrained by the source text and the linguistic repertoire of the speaker. Thus, connectedness interacts with the linguistic resources available to the speaker, making its predictive contribution meaningful but inherently constrained.

Although our correlation and regression analyses suggested that surface-level linguistic attributes were more strongly related to interpreting performance, their predictive power in the structural model was lower than that of the situation model. This result is partially consistent with Ouyang et al. (2021)'s findings, where surface-level features explained a larger proportion of variance in performance. The discrepancy likely reflects differences in participant populations. Ouyang et al.'s data come from novice interpreters competing in the later stages of a professional contest, whereas our data come from non-interpreter bilinguals. For the participants in our sample, the overall comprehensibility of their interpreting output (the situation model) appears to take priority over aspects pertaining to the surface of the text. However, in professional interpreting contexts, especially in the final and semi-final rounds, basic comprehensibility is already taken for granted, and greater emphasis is placed on surface level quality, including the use of more accurate, varied, and advanced words or expressions.

Although the present study focuses on non-interpreter bilinguals, the findings have implications for understanding how interpreting-relevant skills evolve across different levels of expertise. The patterns observed in our non-interpreter bilinguals, such as reliance on short-range repetition, limited lexical diversity, and simpler cohesive structures, likely reflect early-stage cognitive overload and less automatized linguistic access. These findings thus serve as a baseline profile for understanding how interpreting outcomes evolve as bilinguals progress from no experience to novice and expert levels which are expected to produce more cohesive discourse with fewer repetitions and clearer global structure, reflecting the integration of cognitive efficiency and linguistic resources. Future research could directly compare these profiles to trace developmental trajectories in cognitive-linguistic integration during interpreting.

6 Conclusion

This study used SpeechGraphs and Coh-Metrix computational tools to explore cognitive and linguistic processes that jointly shape consecutive interpreting performance in Chinese–English non-interpreter bilinguals. The analyses revealed that greater long-range connectedness was positively associated with interpreting performance, whereas higher levels of short-range repetition and errors were negatively associated with it. These findings highlight the importance of fluent, globally integrated speech for successful real-time interpreting.

At the linguistic level, features across all three Coh-Metrix layers contributed to explaining performance differences. Surface code indicators such as higher lexical diversity and greater lexical specificity predicted better performance, while frequent use of highly familiar words and first-person pronouns predicted lower performance. Textbase features, including additive connectives and anaphor overlap, were also positively associated with interpreting performance, underscoring the role of local cohesion in supporting audience comprehension. At the situation model level, lower readability and higher lexical concreteness were associated with better performance, illustrating the dual importance of linguistic sophistication and explicit meaning construction in interpreting output.

Results from the partial least squares structural equation modeling further showed that linguistic attributes were stronger predictors of interpreting performance than speech connectedness.

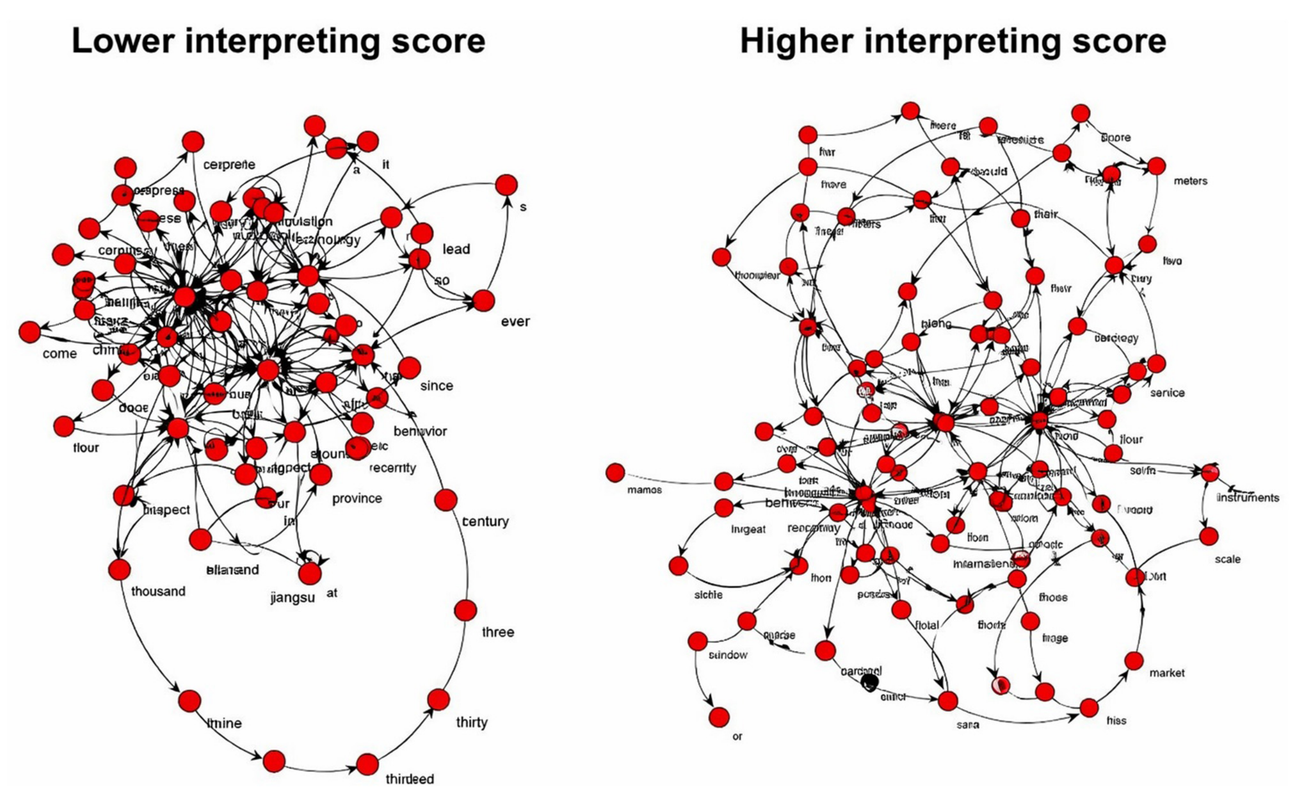

FIGURE 4
Examples of long-range connectedness and short repetitions and their relationship with interpreting scores.

Among the linguistic constructs, the situation model exerted the strongest effect, followed by surface code and textbase features. Speech connectedness also contributed positively, although its effect was comparatively smaller, and error counts showed only a weak influence. Together, these findings demonstrate that interpreting performance in non-interpreter bilinguals reflects the combined influence of cognitive efficiency and linguistic resourcefulness, with linguistic features playing a particularly prominent role.

Overall, the study provides empirical evidence for the distinct yet complementary contributions of cognitive and linguistic processes to interpreting performance. By focusing on non-interpreter bilinguals, it establishes a baseline cognitive-linguistic profile that can be used for comparison with novice and expert interpreters in future research. More broadly, because interpreting represents an intensive form of bilingual language use, the findings also contribute to theoretical models of bilingual processing, suggesting that speech connectedness and linguistic complexity jointly support performance under high processing demands.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YL: Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. JW: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. IF: Conceptualization, Investigation, Methodology, Supervision, Writing – review & editing. JK: Conceptualization, Investigation, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by the Social Sciences Planning Research Project of Shandong Province (Grant No. 25CYYJ13).

Conflict of interest

The author (s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author (s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aben, B., Stapert, S., and Blokland, A. (2012). About the distinction between working memory and short-term memory. *Front. Psychol.* 3:301. doi: 10.3389/fpsyg.2012.00301

Babcock, L., Capizzi, M., Arbula, S., and Vallesi, A. (2017). Short-term memory improvement after simultaneous interpretation training. *J. Cogn. Enhanc.* 1, 254–267. doi: 10.1007/s41465-017-0011-x

Bajo, M. T., Padilla, F., and Padilla, P. (2000). "Comprehension processes in Simultaneous Interpreting," in *Translation in Context: Selected Papers from The EST Congress*, eds. A. Chesterman, San Salvador, N. G., and Gambier, Y. Granada 1998 (Amsterdam: John Benjamins), 127–142.

Bergen, L., Goodman, N. D., and Levy, R. (2012). "That's what she (could have) said: how alternative utterances affect language use," in *Proceedings of the 34th Annual Conference of the Cognitive Science Society,* eds. M. Naomi, P. David, and P. C. Richard (Austin, TX: Cognitive Science Society), 120–125.

Botezatu, R., Weissheimer, J., Ribeiro, M., Guo, T., Finger, I., and Mota, N. B. (2021). Graph structure analysis of speech production among second language learners of Spanish and Chinese. *Front Psychol.* 9: 940269. doi: 10.3389/fpsyg.2022.940269

Chernov, G. (2004). *Inference and Anticipation in Simultaneous Interpreting: A Probability-Prediction Model.* Amsterdam:John Benjamins., 251.

Christoffels, I., and de Groot, A. (2005). "Simultaneous interpreting: a cognitive perspective," in *Handbook of bilingualism: Psycholinguistic Approaches*. eds, J. Kroll and A. de Groot, (New York, NY: Oxford University Press), 454–479.

Christoffels, I., de Groot, A., and Kroll, J. (2006). Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *J. Mem. Lang.* 54, 324–345. doi: 10.1016/j.jml.2005.12.004

Christoffels, I., de Groot, A., and Waldorp, L. (2003). Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. *Biling. Lang. Cogn.* 6, 201–211. doi: 10.1017/S1366728903001135

Crossley, S., and McNamara, D. (2013). Applications of text analysis tools for spoken response grading. *Lang. Learn. Technol.* 17, 171–192. doi: 10.64152/10125/44329

Crossley, S., Salsbury, T., and McNamara, D. (2009). Measuring L2 lexical growth using hypernymic relationships. *Lang. Learn.* 59, 307–334. doi: 10.1111/j.1467-9922.2009.00508.x

Crossley, S. A., Greenfield, J., and McNamara, D. (2008). Assessing text readability using cognitively based indices. *TESOL Q.* 42, 475–493. doi: 10.1002/j.1545-7249.2008.tb00142.x

Crossley, S. A., and McNamara, D. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *Int. J. Contin. Eng. Educ. Life-long Learn.* 21,170–191. doi: 10.1504/IJCEELL.2011.040197

Crossley, S. A., Salsbury, T., and McNamara, D. S. (2012). Predicting the proficiency level of language learners using lexical indices. *Lang. Test.* 29, 243–263. doi: 10.1177/0265532211419331

Crossley, S. A., Salsbury, T., McNamara, D. S., and Jarvis, S. (2011). What is lexical proficiency? Some answers from computational models of speech data. *TESOL Q.* 45, 182–193. doi: 10.5054/tq.2010.244019

De Bruin, A. (2019). Not all bilinguals are the same: a call for more detailed assessments and descriptions of bilingual experiences. *Behav. Sci.* 9, 33. doi: 10.3390/bs9030033

Elmer, S., Klein, C., Kühnis, J., Liem, F., Meyer, M., and Jäncke, L. (2014). Music and language expertise influence the categorization of speech and musical sounds: behavioral and electrophysiological measurements. *J. Cogn. Neurosci.* 26, 2356–2369. doi: 10.1162/jocn_a_00632

Fabbro, F., Gran, B., and Gran, L. (1991). Hemispheric specialization for semantic and syntactic components of language in simultaneous interpreters. *Brain Lang.* 41, 1–42. doi: 10.1016/0093-934X(91)90108-D

Fletcher, C. R., and Chrysler, S. T. (1990). Surface forms, textbases, and situation models: recognition memory for three types of textual information. *Discourse Process.* 13, 175–190. doi: 10.1080/01638539009544752

Fornell, C. G., and Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *J. Mark. Res.* 19, 440–452. doi: 10.1177/002224378201900406

García, A. (2014). The interpreter advantage hypothesis: preliminary data patterns and empirically motivated questions. *TIS* 9, 219–238. doi: 10.1075/tis.9.2.04gar

García, A., Ibáñez, A., Huepe, D., Houck, A. L., Michon, M., Lezama, C., et al. (2014). Word reading and translation in bilinguals: the impact of formal and informal translation expertise. *Front. Psychol.* 5, 1–14. doi: 10.3389/fpsyg.2014.01302

Gieshoff, A. C., and Albl-Mikasa, M. (2024). Interpreting accuracy revisited: a refined approach to interpreting performance analysis. *Perspectives* 32, 210–228. doi: 10.1080/0907676X.2022.2088296

Gile, D, 2nd edn. (2009). *Basic Concepts and Models for Interpreter and Translator Training: Revised.* Amsterdam: John Benjamins.

Gile, D. (1995). Fidelity assessment in consecutive interpretation: an experiment. *Target Int. J. Transl. Stud.* 7, 151–164. doi: 10.1075/target.7.1.12gil

Graesser, A. C., and McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Top. Cogn. Sci.* 3, 371–398. doi: 10.1111/j.1756-8765.2010.01081.x

Graesser, A. C., McNamara, D. S., Louwerse, M. M., and Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behav. Res. Methods Instrum. Comput.* 36, 193–202. doi: 10.3758/BF03195564

Gumul, E. (2006). Explicitation in simultaneous interpreting: a strategy or a by-product of language mediation? *Across Lang. Cult.* 7, 171–190. doi: 10.1556/Acr.7.2006.2.2

Hair, J., and Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research: guidelines using an applied example. *RMAL* 1:100027. doi: 10.1016/j.rmal.2022.100027

Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* 31, 2–24. doi: 10.1108/EBR-11-2018-0203

Han, C. (2016). Investigating score dependability in English/Chinese interpreter certification performance testing: a generalizability theory approach. *Lang. Assess. Q.* 13, 186–201. doi: 10.1080/15434303.2016.1211132

Kim, M., Crossley, S. A., and Kyle, K. (2018). Lexical sophistication as a multidimensional phenomenon: relations to second language lexical proficiency, development, and writing quality. *Mod. Lang. J.* 102, 120–141. doi: 10.1111/modl.12447

Kleijn, S., Pander Maat, H. L. W., and Sanders, T. J. M. (2019). Comprehension effects of connectives across texts, readers, and coherence relations. *Discourse Processes.* 56, 447–464. doi: 10.1080/0163853X.2019.1605257

Kroll, J. F., Tabori, A. T., and Navarro-Torres, C. A. (2021). Capturing the variation in language experience to understand language processing and learning. *Lang. Interact. Acquis.* 12, 82–109. doi: 10.1075/lia.20018.kro

Latifi, S., and Gierl, M. (2021). Automated scoring of junior and senior high essays using Coh-Metrix features: implications for large-scale language testing. *Lan. Test.* 38, 62–85. doi: 10.1177/0265532220929918

Lennon, P. (2000). "The lexical element in spoken second language fluency,". in *Perspectives on Fluency,* ed, Riggenbach H.. (Ann Arbor, MI: University of Michigan Press), 25–42.

Lewis, M. L., and Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition* 153, 182–195. doi: 10.1016/j.cognition.2016.04.003

Li, Y., and Dong, Y. (2022). Use of explicitation by interpreting students and its contribution to consecutive interpreting performance: a developmental perspective. *Perspectives.* 30, 103–119. doi: 10.1080/0907676X.2021.1892158

Liu, M., Schallert, D. L., and Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting* 6, 19–42. doi: 10.1075/intp.6.1.04liu

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Lan. Test.* 34, 493–511. doi: 10.1177/0265532217710675

Macnamara, B. N., Moore, A. B., Kegl, J. A., and Conway, A. R. A. (2011). Domain-general cognitive abilities and simultaneous interpreting skill. *Interpreting* 13, 121–142. doi: 10.1075/intp.13.1.08mac

Meara, P. (2005). Lexical frequency profiles: a Monte Carlo analysis. *Appl Linguist* 26, 32–47. doi: 10.1093/applin/amh037

Mota, N. B., Furtado, R., Maia, P. P. C., Copelli, M., and Ribeiro, S. (2014). Graph analysis of dream reports is especially informative about psychosis. *Sci. Rep.* 4:3691. doi: 10.1038/srep03691

Mota, N. B., Vasconcelos, N. A. P., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., et al. (2012). Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS ONE* 7:e34928. doi: 10.1371/journal.pone.0034928

Mota, N. B., Weissheimer, J., Finger, I., Ribeiro, M., Malcorra, B., and Hübner, L. (2023). Speech as a graph: developmental perspectives on the organization of spoken language. *BP:CNNI* 8, 985–993. doi: 10.1016/j.bpsc.2023.04.004

Mota, N. B., Weissheimer, J., Madruga, B., Adamy, N., Bunge, S. A., Copelli, M., et al. (2016). A naturalistic assessment of the organization of children's memories predicts cognitive functioning and reading ability. *MBE* 10, 184–195. doi: 10.1111/mbe.12122

Nour, S., Struys, E., Woumans, E., Hollebeke, I., and Stengers, H. (2020). An interpreter advantage in executive functions?: a systematic review. *Int. J. Res. Pract. Interpreting* 22, 163–186. doi: 10.1075/intp.00045.nou

Ouyang, L., Lv, Q., and Liang, J. (2021). "Coh-Metrix Model-Based Automatic Assessment of Interpreting Quality,". in *Testing and Assessment of Snterpreting*, eds, J. Chen and C. Han (Singapore: Springer), 179–200

Padilla, F., Bajo, M. T., and Macizo, P. (2005). Articulatory suppression in language interpretation: working memory capacity, dual tasking and word knowledge. *Biling. Lang. Cogn.* 8, 207–219. doi: 10.1017/S1366728905002269

Perego, E. (2003). Evidence of explicitation in subtitling: toward a categorisation. *Across Lang. Cult.* 4, 63–88. doi: 10.1556/Acr.4.2003.1.4

Santilli, M., Vilas, M. G., Mikulan, E., Martorell Caro, M., Muñoz, E., Sedeño, L., et al. (2019). Bilingual memory, to the extreme: lexical processing in simultaneous interpreters. *Biling. Lang. Cogn.* 22, 331–348. doi: 10.1017/S13667289180 00378

Shlesinger, M. and B., Malkiel. (2005). Comparing modalities: cognates as a case in point. *Across Lang. Cult* 6, 173–193 doi: 10.1556/Acr.6.2005.2.2

Wen, Q., and Wang, J. (2008). *Parallel Corpus of Chinese EFL Learners*. Beijing: Foreign Language Teaching and Research Press.

Yudes, C., Macizo, P., and Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Front. Psychol.* 2:309. doi: 10.3389/fpsyg.2011.00309

Yudes, C., Macizo, P., Morales, L., and Bajo, M. T. (2013). Comprehension and error monitoring in simultaneous interpreters. *Appl. Psycholinguist.* 34, 1039–1057. doi: 10.1017/S0142716412000112

Zhang, X., Zheng, B., Wang, R., and He, H. (2025). Is interpreter advantage a gift or an effect of training? *Cognitive changes and interpreting acquisition at the early stage of training. Biling. Lang. Cogn.* 1–11. doi: 10.1017/S136672892500032X

Zhao, H., Chen, X., and Dong, Y. (2024). Young interpreting trainees' better adaptation to the flanker conflicting environment: an ERP study. *J. Neurolinguistics* 70:101181. doi: 10.1016/j.jneuroling.2023.101181

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 920–933. doi: 10.1037//0278-7393.20.4.920