LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 29, No. 2, October 2026, pp. 476-501

LLT Journal: A Journal on Language and Language Teaching
http://ejournal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

ASSOCIATIONS BETWEEN COGNITIVE-LINGUISTIC SKILLS AND OCULOMOTOR PATTERNS IN SENTENCE READING

Beatriz Chaves-Yuste[1*] and Cristina de-la-Peña[2]
[1]Universidad Politécnica de Madrid, Spain
[2]Universidad Internacional de La Rioja, Spain
beatriz.chaves@upm.es[1], cristina.delapena@unir.net[2],
*correspondence: beatriz.chaves@upm.es
https://doi.org/10.24071/llt.v29i2.203
received 18 December 2025; accepted 15 September 2026

Abstract
This study examined how executive functions and cognitive-linguistic skills relate to online lexical processing during sentence reading as a function of word frequency, word length, and concreteness in 171 first- and second-grade children. Executive functions (interference, verbal working memory, and fluency) and linguistic abilities (vocabulary, Rapid Automatized Naming, and phonological awareness) were assessed and linked to two oculomotor indices of word processing, namely First Fixation Duration (FFD) and Total Fixation Duration (TFD), computed for six lexical areas of interest, using descriptive analyses and regression models. Children processed high-frequency words more efficiently than low-frequency words in both FFD and TFD. Early efficiency was associated with fluency and phonological awareness (PA), whereas later efficiency was linked to interference control. Although long words descriptively elicited longer fixations, neither age nor cognitive and linguistic skills significantly predicted differences in FFD or TFD, suggesting that word length was not a reliable predictor of lexical processing efficiency in this task. No consistent advantage was found for concrete over abstract words. Abstract words imposed greater demands during the initial stages of processing (FFD), but these differences diminished in TFD and were not explained by cognitive or linguistic skills. Taken together, these findings provide preliminary evidence that lexical processing efficiency in emergent readers may depend more strongly on word familiarity and on cognitive and linguistic skills than on word length or concreteness, highlighting the interaction between stimulus properties and individual competencies in early reading. However, the findings should be interpreted with caution, as they are task-specific and based on limited item sampling, with only one item representing each lexical category.

Keywords: cognitive-linguistic skills, eye tracking, executive functions, lexical properties, sentence reading.

Introduction
Eye tracking has been extensively employed to investigate the cognitive mechanisms involved in reading. The duration of gaze fixations and regressions provides fine-grained information about processing demands during reading. Eye

This work is licensed under CC BY-SA.
Creative Commons Attribution-ShareAlike 4.0 International License

tracking is now a widely used, temporally sensitive method for investigating reading, providing fine-grained measures of ongoing language processing and a detailed view of readers' interactions with text (Palmović et al., 2023; Papadopoulos et al., 2026; Schotter & Dillon, 2025). Gaze-based metrics such as first fixation, gaze duration, and total fixation time are routinely used to index moment-to-moment processing demands during reading in both experimental and more naturalistic paradigms (Papadopoulos et al., 2026; Schotter & Dillon, 2025). In particular, fixation-based measures provide temporally sensitive indices of word recognition and meaning access, while regressions can signal disruptions in integration processes.

Measuring sentence reading with eye tracking has become a more ecologically valid approach to studying reading because it offers moment-by-moment evidence about ongoing cognitive and linguistic processing that is difficult to obtain from offline assessments such as comprehension tests or self-reports (Schotter & Dillon, 2025). These measures also make it possible to detect variation in processing efficiency and strategy use that may not be evident in offline outcomes (de-la-Peña, 2024; Papadopoulos et al., 2026). More proficient readers generally show more efficient eye-movement patterns over time, whereas less proficient readers tend to show longer fixations, more fixations, and more regressions, reflecting greater processing difficulty (Hautala et al., 2024; Kim et al., 2022).

Word properties also affect eye movements during reading, with readers showing longer or shorter fixation times depending on the lexical and semantic characteristics of the words being processed (Amini et al., 2022; Magnabosco & Hauk, 2024; Schotter & Dillon, 2025; Yan & Pan, 2023). Ambiguous, long, and low-frequency words are expected to receive longer fixations since their processing typically requires greater cognitive effort. Besides lexical properties, reader-level skills and developmental stage also shape eye-movement behavior. Lexical properties are central to accounts of word recognition and remain highly relevant in applied contexts. In readability assessment, lexical features are routinely used to model text difficulty (Pan et al., 2025), and they also inform text adaptation and simplification research (Smirnova et al., 2025).

Executive and linguistic skills, such as phonological awareness (PA), rapid automatized naming (RAN), or morphological awareness, have been shown to correlate with reading (de-la-Peña, 2024; Yan & Pan, 2023). Cognitive-linguistic skills, particularly vocabulary, RAN, and PA, have been linked to oculomotor indices that reflect word-level processing efficiency (Baron et al., 2026; de-la-Peña, 2024). In addition to linguistic skills, executive functions also play a crucial role in shaping eye movement patterns during reading. Readers with strong executive functioning skills are better able to manage attentional resources, maintain information in verbal working memory (VWM), and switch flexibly between different cognitive tasks, all of which are essential for efficient reading. For instance, VWM is essential for integrating information across sentences and paragraphs. Eye tracking research has shown that individual differences in VWM are associated with variation in online reading processes, including eye-movement measures related to ongoing word recognition and rereading (Kim et al., 2021).

However, although previous studies have addressed several of these components in combination, fewer studies have examined how executive resources

and cognitive-linguistic skills jointly relate to online processing when lexical difficulty is manipulated within the same sentence-reading task (e.g., by frequency, length, and concreteness), particularly during early literacy development. More specifically, prior work has often examined lexical variables, reading-related skills, and eye-movement efficiency in partial or separate combinations; accordingly, integrative evidence remains comparatively limited regarding how executive and cognitive-linguistic predictors operate together when children process different types of words within the same sentence-reading context (Baron et al., 2026; de-la-Peña, 2024; Yan & Pan, 2023).

Although related integrative work exists, fewer studies have examined these predictors together within the same analytical framework and under the same sentence-reading task conditions in emergent readers. This question is especially relevant in early readers, because word recognition, lexical access, attentional control, and memory-based integration are all still developing and may contribute differently depending on the lexical demands of the word being read. By examining interference, VWM, fluency, vocabulary, RAN, and PA together across contrasts of frequency, length, and concreteness, the present study aims to clarify whether different lexical properties are associated with different constellations of predictors at early versus more sustained stages of processing. In this way, the study builds on previous work by bringing together executive and cognitive-linguistic factors within a single sentence-reading task and by testing their relation to complementary oculomotor indices of online lexical processing in beginning readers.

First-grade children are still at an early stage of reading development, whereas second-grade children have usually acquired more stable foundational reading skills. Eye-movement research has shown that reading efficiency improves over the early school years, with fixation durations and fixation counts decreasing as children develop (Hautala et al., 2024; Kim et al., 2022). Accordingly, second-grade readers are expected to show more efficient reading patterns than first-grade readers, as predicted by the dual-route cascaded model of reading (Hawelka et al., 2010) and the E-Z Reader model (Reichle et al., 2013) as well as by the Lexical Quality Hypothesis (LQH) (Perfetti & Hart, 2002), all of which posit increasingly automatized lexical access and more efficient coordination of cognitive and linguistic processes as literacy develops. However, recent work in child eye-tracking also emphasizes the importance of distinguishing developmental change from task- and item-specific variance when interpreting reading patterns. This distinction is particularly relevant in developmental research, where both individual differences in executive and linguistic skills and the specific properties of lexical items may shape observed reading patterns. In this sense, the present study does not aim to directly test or extend these models, but rather to refine their empirical scope by examining their applicability under conditions of emergent reading and constrained stimulus sampling.

Given the item-specific nature of the design, the study focuses on between-participant variability, while acknowledging that the extent to which observed effects generalize across lexical items remains an open question. This perspective aligns with ongoing discussions in the mixed-effects modeling literature regarding the importance of disentangling item- and participant-level variance in reading research. Against this background, the study addresses the following research question: To what extent are executive and linguistic skills associated with

individual differences in lexical processing efficiency, as indexed by first fixation duration (FFD) and total fixation duration (TFD), across word categories varying in frequency, length, and concreteness in first- and second-grade readers?

We hypothesize that these skills will differentially predict lexical processing efficiency across word categories, reflecting distinct cognitive demands associated with specific lexical properties. Three hypotheses were formulated:

1. H1. Lexical processing is expected to be more efficient for high-frequency than for low-frequency words, as reflected in shorter fixation durations (FFD and TFD). We also expect frequency-related variation in eye movements to be associated with executive and linguistic skills.
2. H2. Lexical processing is expected to be more efficient for short than for long words, as reflected in shorter fixation durations (FFD and TFD). We also expect length-related variation in eye movements to be associated with executive and linguistic skills.
3. H3. Lexical processing is expected to be more efficient for concrete than for abstract words, as reflected in shorter fixation durations (FFD and TFD). We also expect concreteness-related variation in eye movements to be associated with executive and linguistic skills.

Given the observational design and the limited number of lexical items per condition, these hypotheses are tested as directional and exploratory expectations rather than as strong confirmatory predictions.

Method

This study adopted a non-experimental (observational) design, as no independent variables were manipulated and participants were not randomly assigned to conditions. The variables of interest were measured as they naturally occurred, without experimental manipulation, and the data were collected during a single assessment period. The study included descriptive analyses of executive and cognitive-linguistic skills and eye-movement measures, inferential comparisons between first- and second-grade students, correlational analyses of associations between skills and eye-movement indices, and hierarchical regression models examining whether executive and cognitive-linguistic skills predicted eye-movement measures during sentence reading. All data collection methods are quantitative and cross-sectional.

Sample

The sample was selected based on accessibility to educational institutions. Parental consent was obtained, allowing them to withdraw their children from participation at any time. Inclusion criteria required students to be enrolled in first or second grade, have parental consent, have Spanish as their first language, and have no previous formal participation in similar testing and no diagnosis of a learning disability or sensory/neurological condition.

Figure 1 presents the flow diagram of participant selection. Out of 192 students initially recruited from two charter schools with comparable socioeconomic and cultural backgrounds, 21 were excluded for incomplete test responses (9 students), lack of parental consent for eye-tracking measures (3 students), presence of a neurodevelopmental disorder (3 students), or unreliable eye-tracking data (6 students). The final sample included 171 students: 97 first-

graders (47% female, 53% male) with a mean age of 78 months (SD = 5.3), and 74 second-graders (49% female, 51% male) with a mean age of 89 months (SD = 5.6).

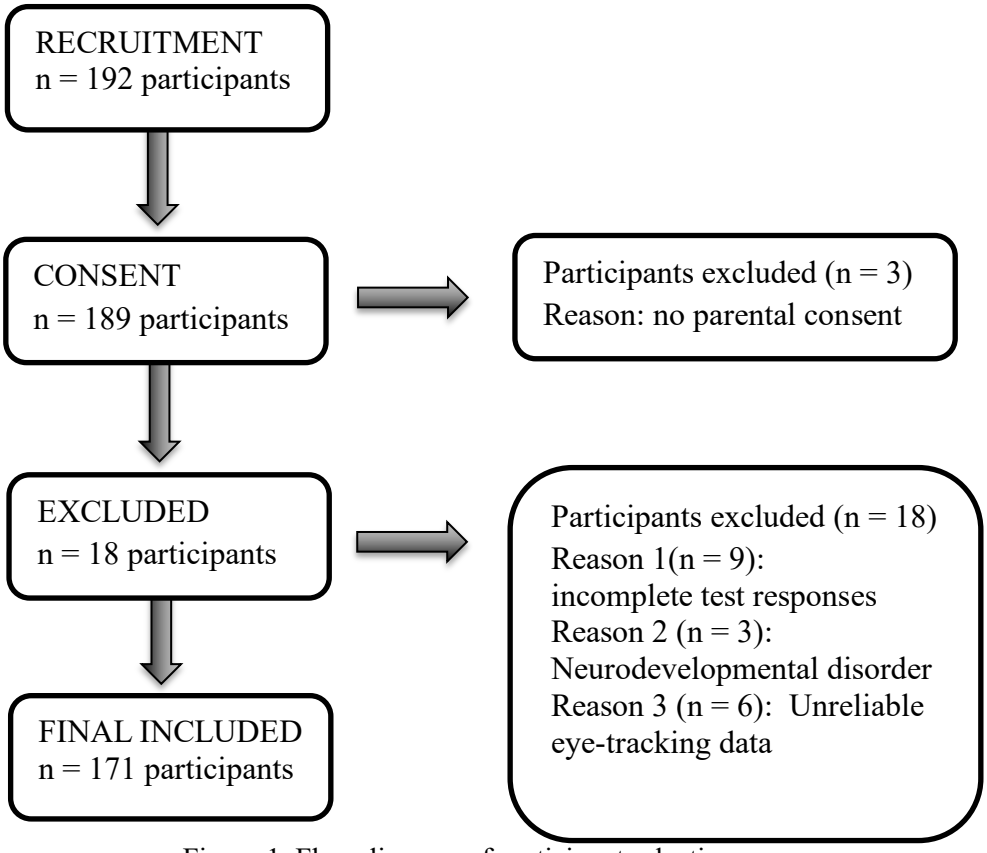


Figure 1. Flow diagram of participant selection

The study received approval from the Ethics Committee of the International University of La Rioja (approval no. PI019/2023). Data were collected between February and May 2024 in the participating schools during the regular school period. The schools were contacted, and verbal consent was obtained from the principals to administer the assessments. All assessments were administered individually in a quiet room at the school. The tasks were presented in the same order for all participants: interference control, verbal working memory, fluency, vocabulary, rapid automatized naming, phonological awareness, and the eye-tracking task. The full assessment lasted approximately 45 minutes, followed by an additional 15-minute eye-tracking session. Parents were provided, by the classroom teachers, with an active consent form detailing the purpose of the study, the assessments their children would complete, the implications of participation, and the contact person for any questions or clarifications.

Instruments

The study assessed executive functions, linguistic skills, and eye-movement behavior during sentence reading. For each instrument, the score used in the analyses is specified below. All assessments used were standardized and validated for children aged 6 and 7 years, and each measure yielded a single score per participant. Because these were direct performance-based measures, internal

consistency indices were not calculated. The validity of these measures is supported by their prior use in child populations and, in the present study, by their ability to discriminate between age groups and by the associations observed among theoretically related variables (convergent validity). To examine these associations, Spearman correlations were conducted on the full sample while controlling for age: vocabulary ↔ phonological awareness ($r = 0.372$, $p < .001$), vocabulary ↔ fluency ($r = 0.359$, $p < .001$), phonological awareness ↔ fluency ($r = 0.193$, $p = .012$), and RAN ↔ fluency ($r = -0.255$, $p < .001$). The results showed significant associations consistent with developmental theory, confirming that the instruments capture the expected constructs and providing evidence of convergent validity in this sample. This approach maximizes statistical power and allows convergent validity to be assessed across the developmental range under study.

Inhibitory control was assessed with the Interference Test from the Neuropsychological Assessment of Executive Functions in Children (Portellano et al., 2009). In this task, children were required to name the ink colour of printed words when the word meaning and ink colour were incongruent. The final interference score, ranging from 1 to 10, was used in the analyses. This test provides a valid, age-appropriate measure of inhibitory control, or interference control, in Spanish children.

Verbal working memory was measured with the Letters and Numbers subtest from the WISC-V Scale (Wechsler, 2015). Children repeated orally presented mixed sequences of letters and numbers in numerical and alphabetical order. The total number of correct responses was used as the final score (range 0–30 points). This standardized measure is widely used in child populations and provides a reliable index of executive functioning. Its Spanish adaptation ensures its validity for the target population. Semantic verbal fluency was assessed with the semantic fluency task from the NEPSY-II Child Neuropsychological Battery (Korkman et al., 2014). Children named as many animals as possible in one minute and as many foods and drinks as possible in a second one-minute trial. The total number of correct responses was used as the fluency score, (range 0–43 points). This standardized task provides a valid measure of semantic fluency for the age range examined in the study.

Receptive vocabulary was assessed with the PPVT-III Peabody (Dunn et al., 2010). Children selected the picture that matched a spoken word from four response options. The total number of correct responses was used as the vocabulary score (range 0–162 points). This instrument is widely used in linguistic research and demonstrates high validity and reliability. Standardized administration procedures were followed.

Naming speed was measured with the Rapid Automatized Naming subtest from the DST-J Test (Fawcett & Nicolson, 2016). Children named 40 familiar objects as quickly as possible. The total naming time in seconds was used in the analyses (range 0–82 points), with lower scores indicating faster performance. This measure provides a reliable index of lexical access speed and has been validated in Spanish child populations.

Phonological awareness was assessed with the Segmentation subtest from the DST-J Test (Fawcett & Nicolson, 2016). Children segmented words into syllables, initial sounds, final sounds, or letters. The total number of correct responses was used as the final score (range 0–12 points). This test is widely used in

developmental and educational contexts and provides valid measures of phonological processing.

Eye movement data were obtained from a sentence reading task based on materials adapted from the standardized text-comprehension subtest of the PROLEC-R (Cuetos et al., 2014). The task included one practice sentence and two experimental sentences, which were presented in Arial 11-point font, black on a white background, with 2 cm vertical spacing between lines. Children read the sentences aloud and then answered two unannounced oral comprehension questions, one literal and one inferential. Comprehension was scored from 0 to 2 points. First Fixation Duration (FFD) and Total Fixation Duration (TFD) were recorded using a Tobii Spark device (60 Hz). Data quality was ensured through individual calibration, a minimum accuracy threshold of 0.5°, and the exclusion of unreliable recordings, thereby guaranteeing the reliability of the oculomotor measures. These measures are widely accepted as valid indicators of real-time reading processes.

For the eye-tracking analyses, target words were selected to represent three lexical contrasts: frequency (*selva* vs. *okapi*), length (*vive* vs. *mamífero*), and concreteness (*animal* vs. *frías*). To minimise positional and structural effects, the first and last word of each line, function words, and words shorter than four characters were excluded from target-word selection. The final set of target words was selected in consultation with four classroom teachers, who agreed on their familiarity and developmental appropriateness for the age group while preserving the intended contrasts.

Procedure

Permissions were obtained from schools, and parental consent was secured at an information meeting. Testing occurred from February to May 2024, respecting academic schedules. The testing order was identical for all participants (interference, VWM, fluency, vocabulary, RAN, PA, and eye-tracking) and conducted individually in a standardized room environment. The study lasted about 45 minutes, with an additional 15 minutes for the eye-tracking task. Instructions were delivered orally, and the experimenter monitored students from a one-meter distance during the eye-tracking test. Students were positioned 62 cm from the screen for calibration accuracy. The researcher verbally instructed the students to read aloud the sentences that would appear on the screen, demonstrating the process by reading the first sentence as an example. Subsequently, the students began reading the two experimental sentences, with measurements being taken from the beginning of the second sentence to the end of the third sentence.

The eye-tracking component used a 60 Hz Tobii Spark, connected to a computer screen with 1920 x 1080 pixels and a minimum spatial accuracy of 0.5° of visual angle under optimal conditions. In addition, a five-point calibration procedure was used and repeated whenever necessary to ensure recording quality. To extract fixation durations, a velocity threshold of 30°/s was applied for the identification of fixations and saccades, along with a minimum duration threshold of 60 ms to avoid the inclusion of spurious microfixations. The duration of each fixation was calculated as the sum of consecutive samples classified within the same fixation event. Accordingly, fixation durations were discretized in multiples of the sampling interval (one sample every 16.67 ms).

Data analysis

Descriptive statistics, including means and standard deviations, were calculated for each assessment. The Kolmogorov-Smirnov test was used to assess the univariate normality of each variable and indicated a non-normal distribution for all variables ($p < .05$). Therefore, non-parametric tests (Mann–Whitney U tests) were applied, as these do not require homogeneity of variances, linearity, or multicollinearity. The independence of observations was ensured by the methodological design: each child was assessed once and independently, such that one participant's results did not affect those of another. In addition, the validity of the median comparisons was supported by verifying that both groups were comparable in size and that the eye-tracking measurements were consistent and complete, allowing the observed differences to be interpreted as meaningful differences between the age groups. Comparisons between grades were conducted using the Mann–Whitney U test, thereby ensuring the statistical validity of the results and grounding the analyses in solid empirical evidence. Effect sizes were interpreted through the biserial rank method (López-Martín & Ardura-Martínez, 2023).

For eye-tracking data, Tobii Pro Lab software (Tobii AB, 2024) was used to calculate mean fixation times, first fixation duration (FFD) —defined as the duration of the first fixation within an area of interest (AOI) during a given interval— and total fixation duration (TFD), which corresponds to the total duration of all the fixations within an AOI during an interval). Six AOIs were predefined based on word properties, specifically high-frequency, low-frequency, short, long, abstract, and concrete words. In the second (experimental) sentence, five AOIs were identified (*selva, okapi, vive, mamífero*, and *animal*), while in the third (experimental) sentence, a single AOI (*frías*) was defined. Thus, word-level AOIs were mapped onto the targeted contrasts (frequency, length, and concreteness) for FFD and TFD extraction. This design implies that eye-movement measures reflect responses to a limited set of predefined lexical items and should therefore be interpreted as item-constrained effects rather than generalizable patterns of lexical processing. Accordingly, the results are task- and item-specific.

The eye-tracking dataset underwent a rigorous cleaning process, which involved removing six students from the sample because of calibration failure. Multiple linear regressions using ordinary least squares (OLS) were conducted on each participant's mean fixation values by word category (FFD and TFD) and the six behavioral measures. Each child contributed a single aggregated value per condition, allowing analyses at the participant level. Under these conditions, the assumption of independence was considered satisfied, as no repeated observations were included within each condition. The variables were transformed into difference scores (e.g., FFD_FreqMinusInfreq) and standardized as z-scores. OLS assumptions were assessed using residual plots (ZRESID vs. ZPRED) and P–P plots. Homoscedasticity and linearity were adequate, the residuals were approximately normally distributed, and VIF values were below 2 for all predictors, indicating the absence of multicollinearity.

There were no missing data; therefore, neither imputation nor bootstrapping was applied. Given the sample size ($N = 171$) and the absence of relevant assumption violations, the regression models were considered robust and appropriate for the data analyzed. Hierarchical regression by blocks was used, with

age entered in Block 1 and cognitive and linguistic skills entered in Block 2, following a theoretical rationale that distinguishes developmental factors from specific skills and allowing the additional explained variance of these skills beyond age (ΔR^2) to be examined; predictors were entered using the Enter method. It is acknowledged that, in studies involving repeated measurements by word or area of interest, mixed-effects models would be the preferred approach, including random intercepts for participants and items, and random slopes where appropriate. However, given the limited number of items per condition, such models were not feasible in the present study. Accordingly, aggregation at the participant level was adopted, and inferential conclusions are restricted to between-participant variability rather than item-level generalization. This approach therefore allows the hypotheses regarding word frequency, length, and concreteness in relation to cognitive and linguistic skills to be evaluated at the participant level. All analyses were performed using SPSS and R with a significance level of $p < .05$.

Findings and Discussion

Findings

The findings are presented in terms of between-grade differences in skills and eye-movement indices, associations, and multiple regression models.

Descriptive and inferential analyses

This section presents preliminary analyses conducted prior to hypothesis testing in order to examine students' performance in each age group across all study variables. Table 1 indicates that second-grade students obtained higher mean scores than first-grade students across the assessed skills, with statistically significant differences for VWM and vocabulary, and a more modest—though still meaningful—difference for phonological awareness (PA). The effect sizes indicate marked differences between the two age groups in the ability to maintain and manipulate information ($r_{(}rb_{)} = .99$) and in vocabulary performance ($r_{(}rb_{)} = .91$), whereas relatively high levels of phonological awareness (PA) are already reached at age 6 ($r_{(}rb_{)} = .40$). Therefore, although second-grade children showed stronger performance on several measures, the magnitude of between-group differences varied across skills. These findings suggest that verbal working memory (VWM) and vocabulary show stronger between-group differences than PA between ages 6 and 7. However, these effect sizes should be interpreted in light of the constrained design and limited stimulus sampling, which may influence the magnitude of observed effects.

Table 1. Results of reading skills for first- and second-grade students

	1st		2nd		U	p	Effect size
	M	SD	M	SD			
Interference	5.85	1.80	5.64	2.03	3414.000	.577	.08
VWM	10.84	4.71	14.91	2.19	1722.500	.000	.99
Fluency	26.98	8.61	27.75	7.41	3333.500	.425	.12
Vocabulary	74.66	15.06	86.82	11.77	1837.500	.000	.91
RAN	49.01	12.99	45.86	12.23	3195.500	.220	.18
PA	8.33	3.38	9.79	1.87	2762.500	.009	.40

The analysis of word categories for first- and second-grade students, as shown in Table 2, reveals that second-grade students demonstrated more efficient lexical processing across all word categories for FFD and TFD. For FFD, significant differences were observed for concrete, abstract, high-frequency, and short words. For TFD, significant differences were observed for high-frequency and short words, with additional statistically significant differences for low-frequency, long, and concrete words.

For lexical frequency, moderate to large effect sizes were observed for high-frequency words (FFD: .55; TFD: .87), whereas smaller effects were found for low-frequency words (FFD: .07; TFD: .31). Overall, these patterns suggest differences in processing efficiency across age groups that are more apparent in high-frequency conditions. However, these effects should be interpreted cautiously given the constrained design and limited number of lexical items per condition, which may influence the magnitude and stability of the observed effects.

For word length, moderate to large effect sizes were observed for short words (FFD: .60; TFD: .93) while long words showed a more moderate effect in TFD (.62). These results suggest that between-group differences may be more apparent in later processing stages (TFD) than in early processing stages (FFD), particularly for longer words. As above, these effects should be interpreted in light of the task-specific and stimulus-limited nature of the design.

For word concreteness, large effect sizes were found for concrete words (FFD: .84; TFD: .62), and a large effect was also observed for abstract words in FFD (.87), suggesting between-group differences in the processing of word concreteness, particularly during early stages of visual processing. These differences appeared to be reduced in later processing stages.

The effect sizes suggest that the magnitude of differences between the two age groups varies across word types and processing stages. Larger effects were generally observed for high-frequency, short, and concrete words, whereas smaller effects were found for low-frequenc and long words, and abstract words showed greater variability across processing stages. However, given the constrained design and limited number of lexical items per condition, these patterns should be interpreted as indicative rather than generalizable, and as potentially sensitive to stimulus sampling and task characteristics.

Table 2. FFD and TFD results by grade and word properties

	1st grade		2nd grade		U	p	Effect size
	M	SD	M	SD			
Low-frequency FFD	462.64	476.10	368.48	375.99	3441.50	.645	.07
Low-frequency TFD	2870.68	1994.24	2334.71	1578.03	2930.50	.040	.31
High-Frequency FFD	580.97	498.94	341.50	318.20	2476.00	.000	.55
High-Frequency TFD	1201.78	1112.02	658.25	891.83	1915.50	.000	.87
Short FFD	506.10	443.65	281.38	222.33	2379.50	.000	.60
Short TFD	1021.81	1087.71	476.16	567.63	1812.00	.000	.93
Long FFD	704.22	630.44	522.19	440.80	3154.00	.175	.20
Long TFD	2302.80	2916.29	1454.83	1671.62	2339.00	.000	.62
Concrete FFD	516.53	488.43	270	201.37	1950.50	.000	.84
Concrete TFD	1357.40	1728.15	689.47	1559.55	2345.00	.000	.62
Abstract FFD	540.62	584.09	379.91	252.31	1915.00	.000	.87
Abstract TFD	1250.86	1487.83	668.13	720.39	3030.50	.082	.26

Predictive analysis

For ease of interpretation, the summary values reported below correspond to the standardized difference scores (z-scores) used in the regression analyses, rather than to raw fixation times in milliseconds.

Word frequency

In the hierarchical regressions (Tables 3 and 4), no multicollinearity issues were detected (VIF < 2), and the residuals met the assumptions of linearity, homoscedasticity, and normality. In Table 3, within Block 1, age did not significantly predict FFD ($\beta = -.121$, $p = .115$), explaining only a small proportion of the variance ($R^2 = .015$), indicating limited explanatory contribution of age for this early measure of reading-related processing. When cognitive and linguistic variables were entered in Block 2, the overall model became significant ($F(7,163) = 2.702$, $p = .011$) and explained 10.5% of the variance ($R^2 = .105$), representing an increase of 9 percentage points over the previous model ($\Delta R^2 = .090$). This level of explained variance is modest and consistent with the constrained design and the inherent variability of eye-tracking measures. Age was no longer a significant

predictor (p = .771), whereas fluency (β = .167, p = .043) and PA (β = −.240, p = .005) emerged as significant predictors. Higher fluency was associated with increases in FFD, whereas higher PA was associated with reductions in this measure. The remaining variables (VWM, vocabulary, RAN, and interference) were not significant predictors (p > .05).

Table 3. Hierarchical regression FFD word frequency

Predictor	β	SE	t	p	95% CI
Block 1					
Age	−.121	.20	−1.585	.115	[−.74, .08]
Model 1					
R^2 = .015, F(1,169) = 2.514, p = .115					
Block 2					
Age	-.027	.25	-.291	.771	[−.56, .42]
VWM	-.084	.11	-.979	.329	[−0.34, .11]
Fluency	.167	.11	2.038	.043	[.00, .44]
PA	-.240	.11	-2.870	.005	[−.54, −.10]
Vocabulary	.037	.12	.391	.697	[−.20, .29]
RAN	.129	.11	1.504	.135	[−.05, .40]
Interference	.027	.11	.330	.742	[−.18, .25]
Model 2					
R^2 = .105, ΔR^2 = .090, F (7,163) = 2.702, p = .011					

Note. β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

In Table 4, age alone did not significantly predict differences in TFD (Block 1: F(1,169) = 2.387, p = .124, R^2 = .014, β = −.118). Adding memory, fluency, PA, vocabulary, RAN, and interference in Block 2 resulted in a significant model, F(7,163) = 2.600, p = .014, R^2 = .101, ΔR^2 = .087, indicating a small increase in explained variance beyond age. Overall, this level of explained variance is modest and consistent with the constrained design and limited stimulus sampling. The analysis of the coefficients revealed that only interference was a significant predictor (β = −.162, t = −1.987, p = .049), while the remaining variables were not significant. These results suggest that participant-level variability in TFD was relatively stable, with only limited associations with inhibitory control as measured by interference.

Table 4. Hierarchical regression TFD word frequency

Predictor	β	SE	t	p	95% CI
Block 1					
Age	−.118	.14	−1.545	.124	[−.51, .06]
Model 1					

Predictor	β	SE	t	p	95% CI
R^2 = .014, F (1,169) = 2.387, p = .124					
Block 2					
Age	.003	.17	.030	.976	[−.34, .35]
VWM	−.151	.08	−1.759	.080	[−0.30, .01]
Fluency	-.002	.07	-.021	.984	[-.15, .15]
PA	.000	.08	−.003	.997	[−.15, .15]
Vocabulary	-.108	.08	-1.155	.250	[−.28, .07]
RAN	.114	.08	1.328	.186	[−.05, .27]
Interference	-.162	.07	-1.987	.049	[−.30, -.01]
Model 2					
R^2 = .101, ΔR^2 = .087, F(7,163) = 2.600, p = .014					

Note. β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

Six- and seven-year-old children showed slightly shorter initial (FFD) and total (TFD) fixation durations for high-frequency words compared to low-frequency words (FFD: M = .0048, SD = 1.3; TFD: M = .0059, SD = .95). The results (Tables 3 and 4) are consistent with H1, which predicted more efficient lexical processing for high-frequency words. This pattern was observed across both FFD and TFD and remained evident at the participant level beyond the contribution of age alone. Early efficiency (FFD) showed associations with phonological awareness and reading fluency skills, whereas later efficiency (TFD) showed an association with inhibitory control (interference). These patterns should be interpreted cautiously given the constrained design and limited number of lexical items per condition, and they reflect associations rather than causal effects.

Word length

In the hierarchical regressions (Tables 5 and 6), no multicollinearity issues were detected (VIF < 2), and the residuals met the assumptions of linearity, homoscedasticity, and normality. In Table 5, within Block 1, age did not significantly predict FFD (β = −.092, p = .231), explaining a very small proportion of the variance (R^2 = .009), suggesting limited explanatory contribution of age for this early measure of reading processing. When cognitive and linguistic variables were added in Block 2, the overall model did not reach significance (F(7,163) = .947, p = .472) and explained 3.9% of the variance in FFD (R^2 = .039). The increase over the previous model (ΔR^2 = .030) was not significant, indicating that the inclusion of these skills did not improve model fit in a reliable way. None of the predictors were significant (p > .05). Overall, these results suggest that variability in FFD for short versus long words is not reliably explained by age or by the cognitive and linguistic variables included in the model, and should be interpreted in light of the constrained design and limited stimulus sampling.

Table 5. Hierarchical regression FFD word length

Predictor	β	SE	t	p	95% CI
Block 1					
Age	−.092	.18	−1.202	.231	[−.57, .14]
Model 1 R^2 = .009, F (1,169) = 1.445, p = .231					
Block 2					
Age	-.061	.17	-.635	.526	[−.59, .30]
VWM	−.142	.08	−1.596	.112	[−0.37, .04]
Fluency	-.090	.07	-1.056	.292	[-.30, .09]
PA	.016	.08	.182	.856	[−.18, .22]
Vocabulary	.100	.08	1.027	.306	[−.10, .34]
RAN	-.040	.08	-.451	.653	[−.25, .16]
Interference	-.019	.07	-.231	.818	[−.22, .17]
Model 2 R^2 = .039, ΔR^2 = .030, F (7,163) = .947, p = .472					

Note. β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

In Table 6, within Block 1, age did not significantly predict TFD during sentence reading (β = −.118, p = .125), accounting for a small proportion of the variance (R^2 = .014). When considered in isolation, age showed limited explanatory contribution to variability in TFD. When cognitive and linguistic variables were added in Block 2, the overall model did not reach significance (F(7,163) = .369, p = .919) and explained 1.6% of the variance in TFD (R^2 = .016), with a negligible increase over the previous model (ΔR^2 = .002). None of the variables included were significant predictors (p > .05). Overall, these results indicate that variability in TFD was not reliably explained by age or by the cognitive and linguistic variables included in the model and should be interpreted in light of the constrained design and limited stimulus sampling.

Table 6. Hierarchical regression TFD word length

Predictor	β	SE	t	p	95% CI
Block 1					
Age	−.118	.15	−1.542	.125	[−.55, .06]
Model 1					
R^2 = .014, F(1,169) = 2.379, p = .125					
Block 2					
Age	-.125	.19	-1.291	.199	[−.64, .13]
VWM	.018	.09	.204	.839	[−0.16, .19]
Fluency	.020	.08	.229	.819	[-.15, .19]
PA	.027	.08	.313	.755	[−.14, .20]
Vocabulary	-.014	.09	-.145	.885	[−.21, .18]
RAN	-.032	.09	-.350	.727	[−.21, .14]
Interference	.003	.08	.034	.973	[−.16, .17]
Model 2					
R^2 = .016, ΔR^2 = .002, F (7,163) = .369, p = .919					

Note**.** β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

Six- and seven-year-old children showed highly similar initial (FFD) and total (TFD) fixation durations for short and long words (FFD: M = .0005, SD = 1.1; TFD: M = −.0008, SD = 1.0). The results (Tables 5 and 6) are not consistent with H2, which predicted shorter fixation durations for short words. This pattern was observed in both FFD and TFD, and no predictor reached statistical significance. Overall, these findings suggest that processing differences between short and long words were minimal in this dataset and should be interpreted in light of the constrained design and limited number of lexical items per condition.

Word concreteness

In the hierarchical regressions (Tables 7 and 8), no multicollinearity issues were detected (VIF < 2), and the residuals met the assumptions of linearity, homoscedasticity, and normality. In Table 7, within Block 1, age did not significantly predict FFD (β = −.100, p = .197), accounting for a very small proportion of the variance (R^2 = .010), suggesting limited explanatory contribution of age for this early measure of reading processing. When cognitive and linguistic variables were added in Block 2, the overall model did not reach significance (F(7,163) = .710, p = .664) and explained 3% of the variance in FFD (R^2 = .030). The increase over the previous model (ΔR^2 = .020) was also not significant, indicating that the inclusion of these skills did not improve model fit in a reliable way. None of the predictors were significant (p > .05). Overall, these results suggest that variability in FFD for concrete versus abstract words was not reliably explained

by age or by the cognitive and linguistic variables included in the model and should be interpreted in light of the constrained design and limited stimulus sampling.

Table 7. Hierarchical regression FFD word concreteness

Predictor	β	SE	t	p	95% CI
Block 1					
Age	−.100	.21	−1.295	.197	[−.69, .14]
Model 1					
R^2 = .010, F(1,169) = 1.678, p = .197					
Block 2					
Age	-.102	.26	-1.063	.290	[−.80, .24]
VWM	.016	.12	.184	.854	[−0.21, .26]
Fluency	.085	.11	.990	.323	[-.11, .34]
PA	-.092	.11	-1.055	.293	[−.35, .10]
Vocabulary	.013	.13	.129	.897	[−.24, .28]
RAN	.070	.12	.786	.433	[−.14, .33]
Interference	.019	.11	.221	.825	[−.20, .25]
Model 2					
R^2 = .030, ΔR^2 = .020, F(7,163) = .710, p = .664					

Note. β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

In Table 8, within Block 1, age did not significantly predict TFD during sentence reading (β = .037, p = .633), accounting a negligible proportion of the variance (R^2 = .001). When cognitive and linguistic variables were added in Block 2, the overall model did not reach significance (F(7,163) = 1.041, p = .405) and explained 4.3% of the variance in TFD (R^2 = .042), with a minimal increase over the previous model (ΔR^2 = .042). None of the variables included were significant predictors (p > .05). Overall, these results suggest that variability in TFD for concrete versus abstract words was not reliably explained by age or by the cognitive and linguistic variables included in the model and should be interpreted in light of the constrained design and limited stimulus sampling.

Table 8. Hierarchical regression TFD word concreteness

Predictor	β	SE	t	p	95% CI
Block 1					
Age	.037	.15	.478	.633	[−.22, .37]
Model 1					
R^2 = .001, F(1,169) = .229, p = .633					
Block 2					
Age	.035	.18	.365	.716	[−.30, .44]

Predictor	β	SE	t	p	95% CI
VWM	.061	.08	.686	.494	[−0.11, .23]
Fluency	.045	.08	.534	.594	[-.12, .21]
PA	-.033	.08	-.376	.707	[−.19, .13]
Vocabulary	-.039	.09	-.402	.688	[−.22, .15]
RAN	-.091	.08	-1.020	.309	[−.26, .08]
Interference	-.143	.08	-1.705	.090	[−.30, .02]
Model 2					
$R^2 = .043$, $\Delta R^2 = .042$, $F(7,163) = 1.041$, $p = .405$					

Note. β = standardized coefficients; SE = standard error; CI = confidence interval. All predictors were standardized (z-scores). VIF values were below 2 for all predictors.

Six- and seven-year-old children showed highly similar initial (FFD) and total (TFD) fixation durations for concrete and abstract words (FFD: M = .0091, SD = 1.3; TFD: M = −.0001, SD = .98). The results (Tables 7 and 8) are not consistent with H3, which predicted shorter fixation durations for concrete words. This pattern was observed in both FFD and TFD, and no predictor reached statistical significance. The findings suggest that differences in processing between concrete and abstract words were minimal in this dataset and should be interpreted in light of the constrained design and limited number of lexical items per condition.

Discussion

The aim of this study was to evaluate the role of executive functions and linguistic skills in online lexical processing during reading, considering word properties such as frequency, length, and concreteness. The study involved 171 first- and second-grade students assessed with standardized executive/cognitive-linguistic measures and eye tracking during oral sentence reading, focusing on FFD and TFD across AOIs representing frequency, length, and concreteness contrasts. The findings suggest that frequency effects were more consistently observed across measures and are broadly aligned with theoretical expectations, whereas effects related to word length and concreteness were weaker and less stable. Furthermore, the results indicate that the relationship between executive and linguistic skills and lexical processing varies depending on word category and processing stage (FFD vs. TFD), although these patterns should be interpreted in light of the constrained design and limited number of lexical items per condition. This overall pattern is in line with recent child eye-tracking work showing that word properties and individual reading skills jointly shape online processing, but not always in uniform ways across measures or developmental stages (Hautala et al., 2024; Papadopoulos et al., 2026; Yan & Pan, 2023). More specifically, the findings provide qualified support for H1, whereas H2 and H3 were not supported. Accordingly, the contribution of executive and linguistic skills appears to depend on the lexical property under consideration rather than operating uniformly across all word categories.

Word frequency

Regarding frequency, the findings are generally consistent with H1, suggesting more efficient lexical processing for high-frequency than for low-frequency words in emergent readers aged 6 and 7 years. The moderate to large effect sizes observed for high-frequency words in both FFD (r_{rb} = .55) and TFD (r_{rb} = .87) indicate an advantage in visual and lexical access, which may reflect facilitation due to lexical familiarity at early stages of processing. This pattern is broadly consistent with previous research (Gerth & Festman, 2021; Hautala et al., 2024; Rayner, 1998; Reichle et al., 2003; Yan & Pan, 2023) showing that lexical frequency modulates fixation times during reading, as well as with more recent studies (Hautala et al., 2024; Kim et al., 2022) linking reading fluency to more efficient visual word recognition. In contrast, the smaller effects observed for low-frequency words (FFD = .07; TFD = .31) suggest that the higher processing demands associated with these items may constrain the expression of participant-level variability, indicating that lexical effects are more limited under conditions of increased processing load at this developmental stage.

Multiple regression analyses provide a more fine-grained perspective on the patterns. For FFD, age did not emerge as a significant predictor; however, the inclusion of cognitive and linguistic skills increased the explained variance (R^2 = .105), although this proportion remains modest. Fluency and phonological awareness (PA) emerged as significant predictors, and their effects should be interpreted as associations rather than causal effects. Specifically, higher fluency was associated with longer FFD, which may reflect differences in processing strategies or allocation of attentional resources, whereas higher PA was associated with shorter FFD, suggesting more efficient access to phonological and lexical representations. These findings are consistent with the LQH (Perfetti & Hart, 2002; Perfetti, 2007) and with recent child eye-tracking evidence showing that individual reading skills modulate the impact of word properties on fixation behavior (de-la-Peña, 2024; Yan & Pan, 2023), although they should be interpreted within the constraints of the present design.

For TFD, age did not significantly predict performance, whereas cognitive and linguistic variables accounted for a small proportion of additional variance (R^2 = .101), with inhibitory control (interference) emerging as the only significant predictor. Given recent work linking eye-movement development to reading fluency and local word-processing efficiency in children, this pattern is compatible with the idea that later stages of lexical processing are partly shaped by broader executive control and coordination processes involved in managing competition and contextual integration (Fan & Reilly, 2022; Hautala et al., 2024; Kim et al., 2022; Schotter & Dillon, 2025), although these effects are modest. However, this interpretation should be considered in light of an important methodological constraint: only one item was used for each lexical category. Consequently, the frequency effects reported here may partly reflect idiosyncratic properties of the selected stimuli rather than fully generalizable frequency-based differences. The observed frequency advantage should therefore be interpreted as preliminary and task-specific. Nevertheless, it provides an indication that lexical familiarity may interact with phonological awareness, fluency, and inhibitory control during the early consolidation of reading.

The findings indicate that the observed frequency advantage may reflect an interaction between lexical familiarity and developing cognitive skills (Strandberg, 2022), but recent evidence suggests this interaction is best interpreted as task-sensitive rather than fully uniform across reading contexts. Thus, the findings support H1 within the present stimulus set but do not warrant broad claims about frequency effects across the population of lexical items.

Word length

The results obtained do not support H2, which predicted more efficient processing for short words, and suggest that interpretation of the word-length effect in emergent reading requires qualification. Descriptive analyses showed moderate to large effect sizes for short words in FFD (r_{rb} = .60) and TFD (r_{rb} = .93), broadly consistent with studies (Gerth & Festman, 2021; Yan & Pan, 2023) reporting increased processing demands for longer words. A moderate effect was also observed for long words in TFD (r_{rb} = .62), suggesting that length-related differences may be more apparent during later stages of processing. However, these effects were not consistently reflected in the regression models, which showed that neither age nor cognitive and linguistic skills significantly predicted variability in FFD or TFD ($R^2 \leq .039$), and no systematic advantage for short words emerged.

From a theoretical perspective, these results can be considered in relation to eye-movement control models based on lexical access, such as the E-Z Reader model (Reichle et al., 2003; Rayner, 1998), which assume that lexical variables contribute to fixation behavior while interacting with perceptual and oculomotor constraints. In this sense, the absence of predictive effects in the present data may indicate that, in emergent readers, variability in FFD and TFD is not primarily explained by lexical-level differences in word length, but may instead reflect a combination of early perceptual and oculomotor factors. This interpretation is consistent with recent longitudinal and developmental eye-tracking studies, insofar as they show that changes in fixation behavior during early reading are not reducible to a single lexical variable but reflect broader shifts in coordination, fluency, and visual-verbal processing efficiency (Hautala et al., 2024; Kim et al., 2022) although the present data do not allow strong inferences regarding the relative contribution of these mechanisms. This conclusion should also be interpreted cautiously because the contrast between short and long words was operationalized with a very limited number of stimuli. As a result, the absence of robust predictive effects may reflect restricted variability at the item level rather than a definitive lack of length effects in early reading. Accordingly, the present findings do not demonstrate the absence of a word-length effect; rather, they indicate that a stable short-word advantage was not detected under the specific conditions of this task. Future studies should incorporate a larger and more representative set of short and long words and use mixed-effects models to distinguish participant-level from item-level variability.

The between-group differences observed for children aged 6 and 7 in VWM (r_{rb} = .99), vocabulary (r_{rb} = .91), and PA (r_{rb} = .40) are consistent with differences in cognitive-linguistic resources that may support reading development. These findings are consistent with studies (Hautala et al., 2024; Kim et al., 2022) linking the development of reading fluency to changes in fixation behavior and more efficient coordination of reading processes. In this sense, the patterns observed in TFD may partially reflect developmental improvements in processing

efficiency, although the present data do not allow strong inferences regarding specific mechanisms such as integration or automatization. However, these between-group differences were not clearly reflected in interindividual differences in online processing as measured by FFD and TFD, suggesting a possible dissociation between general cognitive-linguistic development and moment-to-moment eye-movement behavior in this task.

The findings suggest that the word-length effect, while observable at the descriptive level, is not strongly linked to stable individual differences in emergent readers. Rather than reflecting a direct marker of lexical efficiency, length-related differences in this age range may arise from the interaction between task constraints and early-stage processing mechanisms that are not yet fully organized at the level of individual variability. This interpretation is consistent with Grainger and Ziegler (2011), who emphasize the interplay between sublexical processing and system-level constraints in early word recognition. Accordingly, the present results support a more cautious interpretation of length effects in early reading development and highlight the need for models that accommodate both robust experimental effects and the limited emergence of individual differences at this stage. In this respect, H2 was not supported, but its rejection must remain provisional because of the restricted lexical sampling.

Word concreteness

With respect to concreteness, our results indicate that lexical processing efficiency in 6- and 7-year-old children does not show a consistent advantage for concrete over abstract words in either early (FFD) or later (TFD) stages, and therefore H3 is not supported within the present dataset. However, descriptive and inferential findings suggest that any potential effects are weak and not stable across measures, pointing to a context-dependent and limited concreteness effect in early lexical processing. This is broadly consistent with previous studies (Magnabosco & Hauk, 2024) showing that the semantic characteristics of words, including concreteness, can modulate fixation times during reading (Kearney et al., 2025), although such effects tend to be variable and sensitive to task demands. Recent neurocognitive evidence also supports the view that semantic effects unfold dynamically across processing stages rather than as a single, uniform facilitation effect (Farahibozorg et al., 2022). Recent research has continued to document concreteness effects in lexical processing tasks (Kearney et al., 2025), although such effects appear to be more variable in sentence reading contexts and may depend on task demands.

Although developmental improvements in VWM, vocabulary, and PA were observed between ages 6 and 7, these changes were not clearly reflected in differential processing of concrete versus abstract words. This may suggest that the influence of concreteness is not directly modulated by general cognitive or linguistic development in the present task context, in line with eye-tracking evidence reporting relatively weak semantic effects during natural reading (Magnabosco & Hauk, 2024). This pattern is in line with predictions from dual-route models of reading and the E-Z Reader model (Hawelka et al., 2010; Reichle et al., 2013), which allow for general improvements in reading efficiency without necessarily predicting strong or consistent semantic effects at early stages. Complementarily, evidence from recent semantic neuroscience work (Farahibozorg

et al., 2022) indicating that semantic processing unfolds across distributed and temporally dynamic systems further supports the view that concreteness effects may vary across processing stages and contexts.

At the same time, this interpretation is constrained by the fact that only one concrete and one abstract target item were used. Because semantic variables such as concreteness are often sensitive to contextual and item-specific characteristics, the present findings should not be taken as definitive evidence against a concreteness advantage in emergent readers. Rather, they indicate that such an advantage was not robustly observed under the specific task conditions used in this study. The descriptive indication of greater initial processing demands for the abstract item should therefore be treated as item-specific rather than as evidence of a general abstractness disadvantage. More extensive lexical sampling will be needed to determine whether concreteness interacts systematically with cognitive load or linguistic development in early readers.

The present findings suggest that concreteness effects in early reading are subtle, context-dependent, and constrained by task characteristics, and should therefore be interpreted cautiously. Thus, H3 was not supported in the present dataset, but this result cannot be generalized beyond the specific concrete and abstract items examined.

The findings also suggest that lexical processing efficiency in emergent readers is more consistently associated with word frequency and its interaction with specific cognitive and linguistic skills, whereas effects related to word length and concreteness appear to reflect processing demands and contextual constraints rather than stable advantages in efficiency. This pattern points to a complex interaction between lexical properties, cognitive-linguistic development, and early reading processes. However, these relationships should be interpreted cautiously in light of the task-specific nature of the design, the use of a single item per lexical category, and the limited possibility of separating word-level from participant-level variance.

A further methodological limitation concerns the temporal resolution of the eye tracker. Although the 60 Hz device was considered appropriate for fixation durations above 150–200 ms and enabled data collection in a natural classroom setting, its temporal resolution is lower than that of higher-frequency systems typically used in laboratory-based eye-tracking research. This may have constrained the granularity with which small differences in fixation behavior could be detected. The reduced number of experimental sentences also limits ecological validity and inferential strength, while the absence of contextual variables prevents examination of how predictability, sentence-level integration, or discourse context may have influenced the target-word effects.

The unit of analysis was the participant; therefore, the regression findings concern between-participant variability rather than generalization across lexical items. Frequency, length, and concreteness should consequently be understood as properties of the specific items included in the task rather than as fully sampled lexical constructs. Mixed-effects models would be preferable in designs with multiple observations per participant and item, but the present number of lexical items was insufficient to estimate such models reliably. Future studies should therefore use larger stimulus sets, additional experimental sentences, and mixed-effects approaches that include participant- and item-level random effects.

The practical implications should likewise be interpreted as tentative and exploratory because the study was correlational and cross-sectional. The results do not demonstrate that training any of the measured skills would cause improvements in lexical processing. Nevertheless, the frequency findings suggest that exposure to high-frequency words within contextualized sentences may be relevant to the consolidation of efficient word recognition. The associations involving PA, fluency, and inhibitory control also indicate that lexical processing should be considered in relation to broader cognitive-linguistic development rather than as an isolated word-recognition mechanism.

The observed processing demands associated with the long and abstract target items may also support further investigation of contextual, visual, or semantic scaffolding when less familiar or cognitively demanding words are introduced. Similarly, assessment of early reading may benefit from considering not only speed and accuracy but also underlying skills such as phonological awareness, vocabulary, verbal working memory, naming speed, and inhibitory control. These suggestions should not be interpreted as direct pedagogical or clinical prescriptions; rather, they identify plausible directions for intervention studies designed to test whether integrated approaches combining lexical exposure, cognitive-linguistic practice, and meaningful sentence-level reading improve early reading outcomes.

Future research should also examine additional lexical and contextual variables, including imageability, predictability, morphological awareness, attention regulation, and cognitive flexibility. Longitudinal and intervention-based designs could determine whether the associations observed here change as reading becomes more automatized. Including children with neurological, developmental, or learning disorders would also help establish whether the patterns identified in this task differ across populations with diverse literacy-development profiles.

Conclusion

The findings of this study allow us to address the research question in a qualified manner: the extent to which executive and linguistic skills (e.g., interference, VWM, fluency, vocabulary, RAN, and PA) are associated with lexical processing efficiency as a function of word properties (frequency, length, and concreteness) in first- and second-grade children during early reading development. The results suggest that these associations vary depending on the lexical property under consideration and should be interpreted in light of the correlational nature of the design and the constrained stimulus set.

The findings should also be considered in light of several methodological limitations, including the use of a single item per lexical condition, the limited number of experimental sentences, the absence of item-level modeling, and the lower temporal resolution of the 60 Hz eye tracker relative to higher-frequency laboratory systems. Consequently, the results are preliminary, task-specific, and restricted to between-participant variability rather than generalizable lexical effects. Future research should employ larger and more representative stimulus sets, additional sentence contexts, longitudinal or intervention-based designs, and mixed-effects models capable of separating participant- and item-level variance. Further studies could also examine additional lexical and contextual variables, such as imageability, predictability, morphological awareness, attention regulation, and

cognitive flexibility, and extend the analysis to children with neurological, developmental, or learning disorders.

From a pedagogical perspective, the findings tentatively suggest the relevance of combining systematic exposure to high-frequency words with sentence-level reading activities that support lexical recognition in meaningful contexts. The associations observed between fixation measures and phonological awareness, fluency, and inhibitory control also indicate that early reading instruction may benefit from considering lexical processing alongside broader cognitive-linguistic development. In addition, longer, less familiar, or more abstract words may require greater contextual, visual, phonological, or semantic support during instruction. Teachers may therefore consider introducing more demanding lexical items progressively and embedding them in coherent sentences and texts rather than presenting them exclusively in isolation. However, these implications should be treated as provisional and hypothesis-generating, rather than as direct instructional recommendations, because the present study did not test the effectiveness of any pedagogical intervention.

Frequency-related differences were the most consistent across FFD and TFD, with phonological awareness, fluency, and inhibitory control showing modest associations with participant-level variability. By contrast, word length and concreteness were not reliably associated with lexical processing efficiency in the regression models.

The findings suggest that early lexical processing reflects an interaction between properties of the stimulus and developing cognitive-linguistic skills rather than the uniform contribution of a single mechanism. The results therefore provide a basis for further research on how lexical properties and individual competencies jointly contribute to reading development, rather than definitive evidence regarding general frequency, length, or concreteness effects in emergent readers.

Declarations

The authors report that there are no competing interests to declare. The data that support the findings of this study are available on request from the corresponding author [BCY]. The data are not publicly available because they contain information that could compromise the privacy of research participants.

References

Amini, D., Mamizadeh, M., & Pouralvar, K. (2022). Effect of emotional valence, frequency and gender on L2 word recognition time: An eye-tracking study. *Language Related Research, 13*(3), 111-138. http://doi.org/10.52547/LRR.13.3.6

Baron, A., Martins, A., Puggioni, G., & Harwood, V. (2026). Rapid Automatized Naming (RAN) and word reading fluency in early school-aged children: A pilot eye-tracking study. *Journal of Eye Movement Research*, *19*(1), Article 16. https://doi.org/10.3390/jemr19010016

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2014). *Prolec-R. Batería de evaluación de los procesos lectores* (5ª ed.) [Prolec-R. Battery for the evaluation of reading processes]. TEA Ediciones.

de-la-Peña, C. (2024). Eye-tracking contribution on processing of (implicit) reading comprehension. *Journal of New Approaches in Educational Research, 13*, Article 13. https://doi.org/10.1007/s44322-024-00013-w

Dunn, L. M., Dunn, L. M., & Arribas, D. (2010). *PPVT-III Peabody. Test de vocabulario en imágenes* (2ª ed.) [PPVI-III Peabody. Vocabulary test in images]. TEA ediciones.

Fan, X., & Reilly, R. G. (2022). Eye movement control in reading Chinese: A matter of strength of character? *Acta Psychologica, 230*, Article 103711. https://doi.org/10.1016/j.actpsy.2022.103711

Farahibozorg, S. R., Henson, R. N., Woollams, A. M., & Hauk, O. (2022). Distinct roles for the anterior temporal lobe and angular gyrus in the spatiotemporal cortical semantic network. *Cerebral Cortex, 32*(20), 4549-4564. https://doi.org/10.1093/cercor/bhab501

Fawcett, A., & Nicolson, R. (2016). *DST-J. Test para la detección de la dislexia en niños* [DST-J Test for the detection of dyslexia in children]. TEA Ediciones.

Gerth, S., & Festman, J. (2021). Reading development, word length and frequency effects: An eye-tracking study with slow and fast readers. *Frontiers in Communication, 6*, Article 743113. https://doi.org/10.3389/fcomm.2021.743113

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology, 2*, Article 54. https://doi.org/10.3389/fpsyg.2011.00054

Hautala, J., Hawelka, S., & Ronimus, M. (2024). An eye movement study on the mechanisms of reading fluency development. *Cognitive Development, 69*, Article 101395. https://doi.org/10.1016/j.cogdev.2023.101395

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition, 115*(3), 367-379. https://doi.org/10.1016/j.cognition.2009.11.004

Kearney, E., McMahon, K. L., Guenther, F., Arciuli, J., & de Zubicaray, G. I. (2025). Revisiting the concreteness effect: Non-arbitrary mappings between form and concreteness of English words influence lexical processing. *Cognition, 254*, Article 105972. https://doi.org/10.1016/j.cognition.2024.105972

Kim, Y.-S. G., Petscher, Y., & Vorstius, C. (2021). The relations of online reading processes (eye movements) with working memory, emergent literacy skills, and reading proficiency. *Scientific Studies of Reading, 25*(4), 351-369. https://doi.org/10.1080/10888438.2020.1791129

Kim, Y.-S. G., Little, C., Petscher, Y., & Vorstius, C. (2022). Developmental trajectories of eye movements in oral and silent reading for beginning readers: A longitudinal investigation. *Scientific Reports, 12*, Article 18708. https://doi.org/10.1038/s41598-022-23420-5

Korkman, M., Kirk, U., & Kemp, S. (2014). *Batería neuropsicológica infantil NEPSY-II* [NEPSY-II Child Neuropsychological Battery]. Pearson.

López-Martín, E., & Ardura-Martínez, D. (2023). The effect size in scientific publication. *Educación XX1, 26*(1), 9-17. https://doi.org/10.5944/educxx1.36276

Magnabosco, F., & Hauk, O. (2024). An eye on semantics: A study on the influence of concreteness and predictability on early fixation durations. *Language, Cognition and Neuroscience, 39*(3), 302-316. https://doi.org/10.1080/23273798.2023.2274558

Palmović, M., Jäger, L. A., & Hollenstein, N. (2023). Editorial: Eye-tracking while reading for psycholinguistic and computational models of language comprehension. *Frontiers in Psychology, 14*, Article 1326408. https://doi.org/10.3389/fpsyg.2023.1326408

Pan, W., Li, X., Chen, X., & Xu, R. (2025). Textual form features for text readability assessment. *Natural Language Processing, 31*(3), 800-841. https://doi.org/10.1017/nlp.2024.50

Papadopoulos, T. C., Fella, A., Acartürk, C., Bahçekapılı, E., Canaj, K., Cergol, K., Dinçtopal Deniz, N., Hollenstein, N., Hržica, G., Jozipović, M., Kasperė, R., Klimek-Jankowska, D., Søndergaard Knudsen, H. B., Krebs, J., Krejtz, I., Leszko, M., Łockiewicz, M., Motiejūnienė, J., Özkan, A., … Erol Barkana, D. E. (2026). Eye-Tracking in reading research: A systematic review of studies with children of varying reading ability. *European Psychologist, 31*(2), 134-168. https://doi.org/10.1027/1016-9040/a000579

Perfetti, C. A. (2007). Reading ability: Lexical quality and comprehension. *Scientific Studies of Reading, 11*(4), 357-383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). John Benjamins Publishing Company.

Portellano, J. A., Zumárraga Astorqui, L., & Martínez Arias, R. (2009). *Test de interferencia de la Evaluación Neuropsicológica de las Funciones Ejecutivas en Niños* (ENFEN) [Interference test for neuropsychological assessment of executive functions in children]. Tea Ediciones.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review, 33*(2), 110-149. https://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445-526. https://doi.org/10.1017/ S0140525X03000104.

Schotter, E. R., & Dillon, B. (2025). A beginner's guide to eye tracking for psycholinguistic studies of reading. *Behavior Research Methods*, *57*, Article 68, https://doi.org/10.3758/s13428-024-02572-4

Smirnova, A., Chun, K. B., Rothman, W. L., & Sarma, S. (2025). Text simplification for children: Evaluating LLMs vis-à-vis human expert. In *CHI EA '25: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Article 512, 1-10). ACM. https://doi.org/10.1145/3706599.3719889

Strandberg, A., Nilsson, M., Östberg, P., & Seimyr, G. O. (2022). Eye movements during reading and their relationship to reading assessment outcomes in Swedish elementary school children. *Journal of Eye Movement Research, 15*(4), Article 3. https://doi.org/10.16910/jemr.15.4.3

Tobii AB. (2024). *Tobii Pro Lab* [Computer software]. https://connect.tobii.com/

Wechsler, D. (2015). *WISC-V. Escala de inteligencia de Wechsler para niños-V* [WISC-V. Wechsler Intelligence Scale for Children-V]. Pearson.

Yan, M., & Pan, J. (2023). Joint effects of individual reading skills and word properties on Chinese children's eye movements during sentence reading. *Scientific Reports, 13*, Article 14754. https://doi.org/10.1038/s41598-023-41041-4