

BRIEF RESEARCH REPORT

Beyond breadth of vocabulary: Vocabulary depth can also be impacted by school's socioeconomic status

Junior Vargas Cuevas, Bruno Facon, Ludivine Javourey-Drevet and Severine Casalis 

University of Lille & CNRS, UMR 9193 - SCALab – Sciences Cognitives et Sciences Affectives, Univ. Lille, Lille, France

Corresponding author: Severine Casalis; Email: severine.casalis@univ-lille.fr

(Received 16 June 2025; revised 16 December 2025; accepted 26 February 2026)

Abstract

This study investigates the impact of socio-economic status (SES) on children's vocabulary, focusing on the distinction between vocabulary breadth (number of words known) and vocabulary depth (quality of word knowledge). It aims to determine whether SES affects both dimensions equally and whether the relationship between SES and vocabulary depth is mediated by vocabulary breadth. Participants were 219 children schooled in France in fourth and fifth grades who had French as their dominant language. Analyses showed that SES significantly influences vocabulary breadth, as predicted by previous research. Importantly, it also impacts vocabulary depth. Mediation analysis revealed that vocabulary breadth can mediate this effect depending on the task used to measure vocabulary depth.

Keywords: socio-economic status; vocabulary; breadth; depth; education

Résumé

Cette étude analyse l'impact du niveau socio-économique (NSE) sur le vocabulaire des enfants, en se concentrant sur la distinction entre l'étendue (nombre de mots connus) et la profondeur (qualité de la connaissance) du vocabulaire. L'objectif est de déterminer si le NSE affecte ces deux dimensions de manière égale, et si la relation entre le NSE et la profondeur du vocabulaire est influencée par l'étendue du vocabulaire. Les participants étaient 219 enfants français scolarisés en CM1 et CM2 dont le français était la langue dominante. Les analyses ont montré que le NSE influence de manière significative l'étendue du vocabulaire, comme prévu par des recherches antérieures. Mais surtout, nous avons également observé un impact sur la profondeur du vocabulaire. L'analyse de médiation a révélé que l'étendue du vocabulaire pouvait jouer un rôle de médiation dans cet effet, en fonction de la tâche utilisée pour mesurer la profondeur du vocabulaire.

1. Introduction

1.1. Vocabulary and socio-economic status

Students from lower socio-economic status (SES) backgrounds often exhibit poorer academic outcomes than their higher SES peers (Schoon *et al.*, 2012). This has been related to limited language exposure and lower academic attainments in low-income families (Hoff, 2006; Nag, 2023; Neuman *et al.*, 2018). Indeed, children from low-SES backgrounds hear fewer, less diverse, and less complex vocabulary than higher SES peers (Hart & Risley, 2003; Hoff, 2003).

Such lack of exposure to rich vocabulary limits knowledge building and literacy development (Hoff, 2006, 2013), leading to impoverished language-related capacities (Rakesh *et al.*, 2024), including vocabulary itself (Cain *et al.*, 2003). The vocabulary–literacy inter-connection and the role of SES in vocabulary development are well documented and their relationship has been studied to understand the underlying mechanisms. For example, Buckingham *et al.* (2013) highlight school and individual factors contributing to the SES–vocabulary growth relationship in English-speaking countries. Individual factors include environmental interactions and early literacy abilities. At school, they identify differences in teaching practices and instructional quality (see also Neuman *et al.*, 2018). Further studies have detailed the mediation of different variables in the SES–literacy relationship, such as parenting styles (Hoff & Laursen, 2019; Pungello *et al.*, 2009), cognitive capacities (Lervåg *et al.*, 2019), language-related abilities (Von Stumm *et al.*, 2020), and the type of vocabulary assessed (Van Der Kleij *et al.*, 2023).

Concerning the impact of children’s environments on vocabulary development, Hoff (2003) demonstrated that the growth of children’s productive vocabulary varies as a function of maternal speech that, in turn, differs according to the SES. Similarly, Ma *et al.* (2024) claim that differences in vocabulary development across SES groups can be explained by the learning environment that parents are able to provide. If we consider linguistic interactions, families with greater economic and educational resources may have more opportunities to offer access to academic vocabulary than low-SES parents that might be at a disadvantage in providing this type of linguistic input, which would lead to differences in school performance.

According to the lexical legacy hypothesis (K. Nation, 2017), both reading and spoken language offer opportunities for encountering words in diverse and meaningful contexts. Such exposure contributes to word acquisition. Perfetti’s Lexical Quality Hypothesis posits that knowing a word involves having robust representations of its orthographic, phonological, and semantic features, as well as strong connections between these dimensions (Perfetti & Hart, 2002). Therefore, someone may have heard the word *serendipity* and know how to write it but not fully grasp its meaning. The distinction between having a formal representation of a word but not the full meaning can be accounted by the description of two features of vocabulary, namely, vocabulary breadth and depth. In the example earlier, *serendipity* may be part of the lexical repertoire, contributing to vocabulary breadth but not depth.

1.2. Breadth and depth of vocabulary

Studies reporting vocabulary measures often focus on the number of lexical items or the vocabulary growth rate (Schmitt, 2014; Webb, 2012; Yanagisawa & Webb, 2019). They assess basically the links between form and word meaning, but not the detailed knowledge of the existing vocabulary. For example, Neuman *et al.* (2018) reported that children from

low-income families were exposed to poorer language in content and quantity, both in families and schools, compared to higher income peers, leading to lower performance in vocabulary tests. The differences were remarkable in the Peabody Picture Vocabulary Test-IV (PPVT-IV; Dunn & Dunn, 2007), and the Woodcock-Johnson III-Picture Vocabulary sub-test (WJIII-PV; Woodcock et al., 2007). These tests require matching images to words, that is, they measure essentially, vocabulary breadth, but not vocabulary depth.

Vocabulary breadth refers to the number of words known by someone while vocabulary depth refers to detailed word knowledge that someone could have of their vocabulary. Initially, this distinction accounted for differences in word knowledge (P. Nation, 2019) and aided vocabulary measurement (Anderson & Freebody, 1981) as well as curriculum design. Lately, it has contributed to studies on L1 and L2 word learning (Read, 2004; Yanagisawa & Webb, 2019) and has been supported by the unique predictive values that both breadth and depth provide in reading assessment (Ouellette, 2006; Tong & Tong, 2022). Alternatively, some authors argue against separating breadth and depth, given their strong correlation (Vermeer, 2001), with some studies providing mixed results on their distinct contribution to reading comprehension (Lawrence, 2018). This debate is relevant, but beyond the scope of this study (for a review, see Webb, 2020). We were interested in the impact of SES on vocabulary and its features rather than their relationship to other linguistic skills.

1.3. *Measuring breadth and depth*

Vocabulary breadth proves easier to assess than depth. According to Schmitt (2014), linking a word form with its meaning reflects vocabulary breadth since this allows to infer the number of words for which there exists a referent in someone's lexicon. Thus, vocabulary-size tests that reflect the capacity of an individual to link word forms to their referents are generally considered measures of breadth (Schmitt, 2014). Classic examples include the PPVT (Dunn & Dunn, 2007) and the WJIII-PV (Woodcock et al., 2007), where participants are required to match a picture with a word presented orally. The resulting score indicates how many words are known by the person taking the test, though the reliability of these scores depends on the quality of test standardization and the comparability of the populations on which the norms were based. This is important to consider, especially in bilingual and low-SES populations, as these kinds of tests are highly decontextualized. They typically require to produce responses for vocabulary that is not common or not related to their experience (see Gibson et al., 2018; Peets & Bialystok, 2015, for discussions).

In contrast, measuring vocabulary depth has proven more difficult to conceptualize. Nation (2001) proposes that assessing knowledge of word-specific features (e.g., grammar, syntax, collocations, and register) is a good way to account for vocabulary depth. However, Meara and Wolter (2009) argue that evaluating different aspects of word knowledge does not necessarily capture overall vocabulary knowledge; instead, it is necessary to evaluate the relationship that exists between words within the lexicon. Thus, measures of depth should consider both the ability to connect words to one another and identify detailed intra-word information.

The synonyms sub-test from the Test of Word Knowledge (Wiig & Secord, 2012) evaluates the ability to retrieve words related to each other, focusing on the connections among words. The Word Associates Test (WAT; Read, 1998) requires participants to link a target word to its meaning and identify relevant word associations (both synonyms and collocates). Another task that could refine such assessment is the attributes sub-test from

the Language Processing Test–Revised (LPT-R; Richard & Hanner, 1995), because it measures the capacity to identify intra-word information. This task requires describing objects and providing information such as colour, size, shape, or category, thereby enabling a more detailed assessment of the extent to which the word is known. The Vocabulary Knowledge Scale (VKS; Wesche & Paribakht, 1996) allows individuals to rate their knowledge of individual words on a five-point scale. Originally developed to measure vocabulary gains from reading, the test is now a general measure of vocabulary depth (Schmitt, 2014). However, because the VKS provides a subjective judgement of knowledge about individual words, it is insufficient on its own to completely capture vocabulary depth. In fact, the more unfamiliar a word is, the more challenging it will be to evaluate the knowledge of it.

All in all, these tasks go beyond the form–meaning association to capture the detailed knowledge of the words already known. Furthermore, some standardized tests blur not only the line between breadth and depth by *demanding* to recall individual word meanings, but also the articulation of nuanced definitions (e.g., WISC-V; Wechsler, 2014). Such overlap may help explain the strong correlations often observed between breadth and depth, as a larger vocabulary facilitates the acquisition of richer meanings and the establishment of relationships among words (Cain *et al.*, 2003).

1.4. Research question

Children from low-SES backgrounds might be at a disadvantage in building quality lexical representations, as they typically receive less linguistic input than their higher SES peers. While research consistently shows that SES negatively affects vocabulary, the tasks used in these studies most often tap on vocabulary size. Given the distinction between *knowing a word* (i.e., breadth) and *knowing it well* (i.e., depth), it remains unclear whether SES affects vocabulary depth to the same extent that it has been shown to impact vocabulary breadth (Neuman *et al.*, 2018). Our study addresses this question in fourth and fifth graders. More precisely, we wondered whether the SES relationship to vocabulary depth, if it exists, is direct or is mediated by vocabulary breadth. We measured breadth through picture designation, while depth was measured as attributes task and a self-assessment task designed for this study as well as a class-adapted synonyms task.

This design was chosen because most SES-related vocabulary studies have been conducted in English and have focused on academic vocabulary or clinical populations (Deane *et al.*, 2014; Schmitt, 2014; Webb, 2012) and because commonly used standardized tests provide information about vocabulary size while measuring vocabulary depth remains challenging. Moreover, standardized tests may amplify SES differences because they often emphasize academic vocabulary that is less familiar to children from low-SES backgrounds (Marulis & Neuman, 2010; Van Der Kleij *et al.*, 2023). As a result, children from low-SES backgrounds may appear to know fewer words not because their vocabularies are smaller, but because the words they do know are under-represented in this kind of assessments.

2. Method

2.1. Participants

Participants were 219 fourth and fifth graders (9–11 years old) from five low- and average-SES schools in France, all of whom spoke French as their primary language. Due to missing data, analyses were conducted on 212 participants ($M = 10$ years, $SD = 0.62$).

In France, schools are designated based on the IPS (Social Position Index; Dauphant et al., 2023), a composite indicator derived from the socio-economic and cultural background of students' families (parents' profession and level of education, number of books at home, etc.), which reflects the socio-economic status of children attending the school.¹ Schools with a low IPS serve areas with significant socio-economic challenges and are designated as REP (Priority Education Network in French).² Since the participating schools in this study differed in IPS values and there was a different number of children per school, it was not appropriate to use IPS as a continuous grouping factor. Instead, the REP classification was used to define the two student groups. Three schools in the REP category comprised the low SES group (IPS = 63.4, 71.2, 77.8; $n = 132$, coded as 0), while two schools outside this category comprised the average SES group (IPS = 104, 126; $n = 80$, coded as 1).

Informed consent was obtained from the children's legal guardians, and oral assent was secured before the assessment, with the option to withdraw at any time. Teachers confirmed that participants had no learning, cognitive, sensory (e.g., uncorrected visual or hearing impairments), or psychiatric disorders, based on their access to student records and regular interactions. This study was approved by the Ethical Committee (Ref. 2023-730-S120) and complied with the EU General Data Protection Regulation (GDPR) under the supervision of the Data Protection Official (DPO, reference 2025-028) at [Lille] University. It adheres to the ethical principles outlined in the Declaration of Helsinki.

2.2. Vocabulary assessment

Children completed tasks measuring breadth and depth:

1. *Attributes*. Adapted from the Lexical Processing Test–Revised (Richard & Hanner, 1995). Children had to read a list of descriptions, including function, colour, size, shape, category, and composition that could correspond or not to words presented in writing. Their task was to indicate the descriptions that corresponded to each word. The words were the same as in the self-assessment task.
2. *Synonyms*. Based on the synonyms sub-test of the Test of Word Knowledge (Wiig & Secord, 2012), children read a text with underlined words. For each underlined word, they had to select a synonym from a list of three possibilities.
3. *Self-assessment of vocabulary knowledge*. Adapted from Wesche and Paribakht (1996), children used a 0–15 scale to indicate how well they believed they knew certain words they heard. The stimuli included 20 words drawn from the Manulex database (Lété et al., 2004). Word frequencies ranged from 3.03 to 493.3 occurrences per million, based on data for 8- to 11-year-olds (Annex 1).
4. *Picture designation task*. Taken from the battery to assess written and oral language (EVALEO 6–15, Launay et al., 2018). Children heard a word and had to select the corresponding image from four options.

¹In 2023, the average IPS in the region where the study was conducted was 99.8 (maximum = 161.9; minimum = 62.5).

²This designation does not imply a different academic curriculum. It is intended solely to identify schools that may require additional resources such as smaller class sizes, further staff, or funding. All schools, whether REP or non-REP, follow the same national curriculum established by the French Ministry of Education.

One minute word recognition task. Children had to cross the words out of a mixed list of words and pseudo-words within a minute. This task was intended to control potential reading-related differences between children.

2.3. Procedure

The session took place in the children’s usual classroom. The main researcher introduced himself, obtained consent from the children to participate in the study and distributed the booklet containing all the tasks. They were administered in a fixed order in pencil-and-paper format to the entire class in a single 45 minutes session.

Self-assessment of vocabulary was the first task, followed by the picture designation task. Next, participants completed the attributes task, and finally, the synonyms task. Once all children had finished, the researcher carefully explained the word recognition task and started the timer. At the end, the researcher collected the booklets and concluded the session.

3. Results

The raw data on which these analyses were carried out will be available on demand at the next address: <https://doi.org/10.57745/LPOH2R>. Descriptive statistics are reported in Table 1. Reliability estimates were good for the attributes task ($\omega = .874$), and acceptable for picture designation ($\omega = .739$) and self-assessment ($\omega = .785$). Internal consistency was not applicable to the synonyms task as scoring was based on the total performance rather than individual item responses. Nonetheless, previous research on similar sub-tests reports reliability coefficients ranging from .84 to .93 (Wiig & Secord, 2012).

The subsequent analyses were conducted using STATA release 16 (Stata Statistical Software, 2019) software. The distribution of the data for the synonyms task did not deviate significantly from normality. Applying log transformations corrected skewness in the attributes, picture designation, and word recognition tasks, with the transformed data meeting the assumption of normality (Shapiro–Wilk, $p > .05$). Self-assessment scores

Table 1. Descriptive statistics for average- and low-SES groups’ performance on vocabulary breadth and depth measures and word recognition

Task	Low-SES		Average-SES		Total	
	<i>M</i> (SD)	Range	<i>M</i> (SD)	Range	<i>M</i> (SD)	Range
Breadth of vocabulary						
Picture designation	33 (3.89)	19–40	35.4 (3.55)	23–41	33.9 (3.93)	19–41
Depth of vocabulary						
Attributes	32.3 (6.41)	14–42	34.8 (6.41)	18–42	33.3 (6.77)	14–42
Self-assessment	13.8 (6.84)	6.20–15	14.1 (1.12)	10.20–15	13.9 (1.22)	6.20–15
Synonyms	3.3 (1.35)	0–6	4.3 (1.18)	2–7	3.7 (1.37)	0–7
Reading						
Word recognition	31.9 (14.53)	9–71	36.2 (15.99)	8–81	33.5 (15.21)	8–81

Note: Picture designation has a maximum score of 42; attributes, 44; self-assessment, 15; synonyms, 7; word recognition has no fixed maximum.

Table 2. Correlation of the standardized variables and SES

Task	SES	Picture designation	Attributes	Synonyms	Word recognition
SES	–				
Picture designation	.32	–			
Attributes	.21	.42	–		
Synonyms	.35	.43	.47	–	
Word recognition	.14	.25	.27	.30	–

Note: All variables were standardized. Correlations are Pearson's *r* coefficients, all of which were significant at $\alpha < .05$.

remained highly skewed, suggesting a near-ceiling effect, and were therefore excluded from further analyses. The remaining scores were standardized ($M = 0$, $SD = 1$) to facilitate interpretation of regression analyses results. Table 2 reports the correlations between SES and the test scores, all of which were significant at $\alpha < .05$.

Model comparisons were conducted to assess the fit of linear, quadratic, and cubic regressions for each pair of variables. Of the 30 comparisons performed (10 relationships $\times$ 3 models), only two yielded statistically significant improvements. However, the additional variance explained by the more complex models was minimal. Thus, to avoid unnecessary statistical complexity, subsequent analyses were conducted within a linear framework (Cohen et al., 2003, pp. 225–227). Hierarchical regression analyses, presented in Table 3, were conducted to examine the effect of SES on both vocabulary breadth and vocabulary depth, as measured by the attributes and synonyms tasks. In addition, we assessed the extent to which each vocabulary dimension predicted variance in the other. In each model, performance on the word-recognition task was entered as a covariate to control for individual differences in reading.

To evaluate the assumptions underlying the regression analyses, residual diagnostics were systematically conducted. Visual inspection of standardized residuals, together with statistical tests, indicated that residuals did not significantly deviate from normality, revealed no severe outliers, and showed no evidence of heteroscedasticity. The variance inflation factors (VIFs) and condition numbers for all predictors were less than 1.16 and 2.30, respectively, indicating no problematic multi-collinearity across analyses. Finally, non-significant results on both the Breusch–Godfrey and Durbin's alternative tests suggested the absence of residual autocorrelation in all analyses.

As expected, SES significantly predicted vocabulary breadth, $F(1, 210) = 25.18$, $p < .001$. Importantly, it also significantly predicted vocabulary depth, $F(1, 210) = 9.74$, $p = .002$ for attributes and $F(1, 210) = 30.22$, $p < .001$ for synonyms. Adding breadth improved depth models, $\Delta R^2 = .109$ for attributes, $\Delta R^2 = .078$ for synonyms, while reducing SES effects which, for the attributes task, became non-significant. This suggested a possible mediation effect. For this reason, mediation analyses were conducted using the “Mediate” algorithm with 1000 bootstrap re-samples to ensure robust confidence intervals (see Hayes, 2017, for a discussion of the advantages of this method over the normal theory approach for making statistical inferences about indirect effects).

When assessed by the attributes task, the direct effect of socio-economic status (SES) on vocabulary depth was no longer significant, indicating that the effect of SES was fully mediated by breadth (Table 4). In contrast, when assessed by the synonyms task, SES remained a significant direct predictor of depth, with a weaker mediation effect by breadth (Table 5). Overall, breadth mediates the relationship between SES and vocabulary

Table 3. Hierarchical regression analyses of breadth and depth variables

Vocabulary breadth										
Variables	<i>b</i>	SE	95% CI for <i>b</i>		<i>t</i>	<i>p</i>	Bootstrapped <i>R</i> ²	ΔR^2	95% CI for bootstrapped <i>R</i> ²	
			LL	UL					LL	UL
Step 1										
SES	0.67	0.13	.41	.94	5.02	<.001	.11	—	.03	.19
Constant	−0.25	0.08	−.42	−.09	−3.08	.002				
Step 2										
SES	0.61	0.13	.35	.88	4.64	<.001	.15	.04	.05	.25
Word recognition	0.21	0.06	.08	.34	3.24	.001				
Constant	−0.23	0.08	.39	−.07	−2.86	.005				
Step 3a										
SES	0.49	0.13	.24	.74	3.88	<.001	.25	.10	.15	.36
Word recognition	0.13	0.06	.00	.25	2.02	.045				
Depth (Att)	0.34	0.06	.21	.46	5.36	<.001				
Constant	−0.18	0.08	−.33	−.03	−2.42	.016				
Step 3b										
SES	0.40	0.13	.14	.67	3.03	.003	.23	.08	.13	.33
Word recognition	0.13	0.06	.00	.25	2.01	.046				
Depth (Syn)	0.32	0.07	.18	.45	4.71	<.001				
Constant	−0.15	0.08	−.31	.00	−1.94	.054				
Vocabulary depth (attributes)										
Variables	<i>b</i>	SE	95% CI for <i>b</i>		<i>t</i>	<i>p</i>	Bootstrapped <i>R</i> ²	ΔR^2	95% CI for bootstrapped <i>R</i> ²	
			LL	UL					LL	UL
Step 1										
SES	0.43	0.14	.16	.71	3.12	.002	.04	-	−.01	.10
Constant	−0.16	0.08	−.33	.00	−1.91	.057				
Step 2										
SES	0.36	0.14	.10	.63	2.68	.008	.10	.06	.02	.18
Word recognition	0.24	0.07	.11	.37	3.68	<.001				
Constant	−0.14	0.08	−.30	.03	−1.65	.099				

Table 3. (Continued)

Vocabulary depth (attributes)										
Variables	<i>b</i>	SE	95% CI for <i>b</i>		<i>t</i>	<i>p</i>	Bootstrapped <i>R</i> ²	ΔR^2	95% CI for bootstrapped <i>R</i> ²	
			LL	UL					LL	UL
Step 3										
SES	0.14	0.13	−.12	.41	1.07	.283	.21	.11	.12	.30
Word recognition	0.17	0.06	.04	.29	2.64	.008				
Breadth	0.36	0.07	.23	.49	5.36	<.001				
Constant	−0.05	0.08	−.21	.10	−0.68	.495				
Vocabulary depth (synonyms)										
Variables	<i>b</i>	SE	95% CI for <i>b</i>		<i>t</i>	<i>p</i>	Bootstrapped <i>R</i> ²	ΔR^2	95% CI for bootstrapped <i>R</i> ²	
			LL	UL					LL	UL
Step 1										
SES	0.73	0.13	.47	.99	5.50	<.001	.10	-	.05	.20
Constant	−0.27	0.08	−.44	−.11	−3.38	<.001				
Step 2										
SES	0.66	0.13	.40	.91	5.09	<.001	.19	.06	.11	.27
Word recognition	0.25	0.06	.13	.38	4.04	<.001				
Constant	−0.25	0.08	−.40	−.09	−3.14	.002				
Step 3										
SES	0.47	0.13	.22	.73	3.65	<.001	.26	.07	.17	.36
Word recognition	0.19	0.06	.07	.31	3.11	.002				
Breadth	0.30	0.06	.18	.43	4.71	<.001				
Constant	−0.18	0.08	−.33	−.03	−2.32	.021				

Note: CI = confidence interval; LL = lower limit; UL = upper limit; bootstrapped *R*² based on 1000 bootstrap re-samples ΔR^2 = change in *R*². Step 3a adds depth as measured by the attributes task; Step 3b adds depth as measured by the synonyms task.

depth, with a stronger direct effect of SES on the synonyms task compared to the attributes task.

4. Discussion

Our analysis revealed that SES predicts vocabulary breadth, in line with previous research (Hart & Risley, 2003; Hoff, 2006; Lervåg et al., 2019). Critically, it also predicted

Table 4. Mediation analysis for vocabulary depth as measured by the attributes task

Effect	Estimate	95% CI	
		LL	UL
Indirect (ACME)	.22	.11	.36
Direct	.15	–.11	.41
Total	.37	.10	.64
Proportion of total effect mediated	.59	.35	1.99

Note: The upper limit of the 95% confidence interval for the proportion mediated here exceeds 100%, a result that stems from the way this index is computed – as the ratio of the indirect effect to the total effect. When the total effect, used as the denominator, is small or unstable in bootstrap samples, the proportion mediated can surpass 100%. This instability limits the interpretability of the proportion of total effect mediated, making the indirect effect (ACME) a more reliable and interesting indicator (Hayes, 2009).

Table 5. Mediation analysis for vocabulary depth as measured by the synonyms task

Effect	Estimate	95% CI	
		LL	UL
Indirect (ACME)	.19	.09	.32
Direct	.48	.22	.73
Total	.67	.40	.91
Proportion of total effect mediated	.28	.21	.47

performance on measures of depth, even after controlling for breadth in the synonyms task, which had not been previously investigated. This suggests that socio-economic disparities extend beyond vocabulary size to include vocabulary depth.

According to the Lexical Legacy Hypothesis (K. Nation, 2017), frequent and diverse encounters with words in meaningful contexts help form robust lexical representations and strengthen their connections. The Lexical Quality Hypothesis (Perfetti & Hart, 2002) further suggests that strong lexical representations are crucial for effective reading and language processing which in turn enhance word knowledge, that is, the depth of vocabulary. Children from lower SES backgrounds may therefore struggle to create stable, inter-connected lexical representations, which is impacting vocabulary depth. One likely explanation is that these children are exposed to less rich and varied language both at home (Hart & Risley, 2003; Hoff, 2006; Nag, 2023) and at school (Buckingham et al., 2013; Neuman et al., 2018).

The variability in mediation outcomes across depth tasks highlights the complexity of the construct itself. In accordance with previous frameworks, we assessed both the depth of intra-word knowledge and the capacity to relate words between them (Meara & Wolter, 2009; Webb, 2012). While a wider breadth of vocabulary can facilitate performance on measures relying on reading, SES continues to yield a direct influence on vocabulary depth when tasks require more inter-lexical processing, such as retrieving semantic links between words. The different patterns observed across tasks may therefore be explained by their specific demands and the capacities they solicit. In this study, the attributes task

required reading and understanding descriptions to determine whether they matched a word, whereas the synonyms tasks requested identifying related words.

It is noteworthy, however, that SES played a significant role in performance on the attributes task before accounting for breadth. Once breadth was included in the model, the impact of SES disappeared, suggesting that general vocabulary knowledge explained children's ability to identify more word characteristics. That is, knowing more words likely improved their understanding of the descriptions used in the task. This underscores the importance of carefully selecting and precisizing what each task assesses when measuring vocabulary depth. Otherwise, the linguistic challenges faced by children from at-risk populations may be under-represented, which may guide to misleading conclusions about their vocabulary knowledge.

The differential impact of SES on denomination and performance on the synonyms task suggests that breadth and depth remain distinct dimensions of vocabulary knowledge, consistent with studies demonstrating their independent contribution to reading (Ouellette, 2006; Tong & Tong, 2022). However, the finding that breadth mediated the SES effect on the attributes task reinforces the idea of carefully selecting measures of depth. Tasks that require producing or recognizing more words than a child actually knows may blur the distinction between these two domains of vocabulary. As Vermeer (2001) argues, they are inherently inter-related, making it difficult to isolate one without engaging the other. Hence, further research on additional facets of vocabulary depth in which SES may have a role might be beneficial.

The present findings also highlight the potential value of interventions that address both vocabulary breadth and depth, particularly for children from low-SES backgrounds. Although many vocabulary-building programmes focus on expanding the number of words children know (Marulis & Neuman, 2010), our results suggest that supporting richer semantic representations and strengthening connections among words may also be important. For instance, strategies that encourage children to explore the meanings, usage, and semantic relationships of words in different contexts could be more effective in fostering deep, long-lasting vocabulary learning (Beck et al., 2013).

Some limitations of this study should be acknowledged. First, the measures of vocabulary depth were tailored for this study and, while they showed acceptable reliability, further validation is needed. Nonetheless, to our knowledge, this is the first attempt to measure breadth and depth separately to analyse the impact of SES. Second, in this study, we could not assess socio-economic status at the individual level. SES was linked to schools rather than to each child, which prevents us from disentangling the effects of children's own SES from potential effects of school-level characteristics. This limitation is relevant because school quality could partly account for the observed differences, as shown in previous work in English-speaking contexts demonstrating SES-related variability in instructional quality (Buckingham et al., 2013; Neuman et al., 2018). Note, however, that in France, every school follows the same national curriculum, and teachers in the low-SES areas are generally experienced. However, one cannot rule out other unmeasured school-level influences. Finally, while we controlled for word recognition in our analyses, other cognitive or environmental variables, such as working memory (Rakesh et al., 2024) or home literacy practices (Neuman et al., 2018), could also influence the relationship between SES and vocabulary development.

Future research may benefit from using different tasks aiming to measure depth in order to identify where the line blurs with breadth. It would also be worthy to evaluate whether interventions that target vocabulary depth in addition to breadth can mitigate

SES-related vocabulary gaps. Clarifying how both domains of vocabulary distinguish could inform more precise vocabulary research and educational practices.

5. Conclusion

In summary, our study highlights SES's impact on both breadth and depth of children's vocabulary. The distinction between breadth and depth of vocabulary may help to design effective educational interventions to reduce the vocabulary gap between children from low- and high-SES backgrounds. In short, efforts to improve children's vocabulary knowledge must focus not only on the number of words they know, but also on the quality and depth of their understanding.

Acknowledgements. The authors acknowledge the schools that participated in this study, as well as the parents and children who generously agreed to take part in the research. Special thanks to the teachers who very kindly facilitated the data collection process. This work was supported by PIA 3 100% IDT Inclusion: un defi, un territoire, which funded the first author's doctoral contract.

Statement of ethical approval. Lille University Behavioral Science Ethics Committee.

Ethics Committee Reference for the Study: 2023-730-S120.

Type of consent: informed written consent from the parents.

Note that a simplified version was provided for parents who might have experienced difficulty understanding the original.

Furthermore, the participating children were informed that they could opt out of the experiment at any time prior to its start.

References

- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In *Comprehension and teaching: Research reviews* (pp. 77–117). International Reading Association. Routledge.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life. Robust vocabulary instruction (second)*. The Guilford Press.
- Buckingham, J., Wheldall, K., & Beaman-Wheldall, R. (2013). Why poor children are more likely to become poor readers: The school years. *Australian Journal of Education*, 57(3), 190–213. <https://doi.org/10.1177/0004944113495500>.
- Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30(3), 681–694. <https://doi.org/10.1017/S0305000903005713>.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203774441>.
- Dauphant, F., Evain, F., Guillermin, M., Simon, C., & Rocher, T. (2023). L'indice de position sociale (IPS): Un outil statistique pour décrire les inégalités sociales entre établissements. *Note d'information de la DEPP*, 1–4. <https://doi.org/10.48464/ni-23-16>.
- Deane, P., Lawless, R. R., Li, C., Sabatini, J., Bejar, I. I., & O'Reilly, T. (2014). Creating vocabulary item types that measure students' depth of semantic knowledge. *ETS Research Report Series*, 2014(1), 1–19. <https://doi.org/10.1002/ets2.12001>.
- Dunn, L., & Dunn, D. (2007). *Peabody picture vocabulary test, fourth edition (PPVT-4)*. NCS Pearson, Inc.
- Gibson, T. A., Oller, D. K., & Jarmulowicz, L. (2018). Difficulties using standardized tests to identify the receptive expressive gap in bilingual children's vocabularies. *Bilingualism (Cambridge, England)*, 21(2), 328–339. <https://doi.org/10.1017/S1366728917000074>.
- Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 27(1), 4–9.
- Hayes, A. F. (2009). Beyond baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408–420. <https://doi.org/10.1080/03637750903310360>.

- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368–1378. <https://doi.org/10.1111/1467-8624.00612>. Guilford publications.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49(1), 4–14. <https://doi.org/10.1037/a0027238>.
- Hoff, E., & Laursen, B. (2019). Socioeconomic status and parenting. In *Handbook of parenting: Biology and ecology of parenting* (Vol. 2, 3rd ed., pp. 421–447). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9780429401459-13>
- Launay, L., Maeder, C., Roustit, J., & Touzin, M. (2018). EVALEO 6-15 – Batterie d'évaluation du langage oral et du langage écrit chez les sujets de 6 à 15 ans [Oral and written language assessment battery for 6- to 15-year-olds] *Rééducation Orthophonique*, 273, 93–135.
- Lawrence, J. F. (2018). Academic vocabulary and reading comprehension: Exploring the relationships across measures of vocabulary knowledge. *Reading and Writing*, 32, 285–306. <https://doi.org/10.1007/s11145-018-9865-2>.
- Lervåg, A., Dolean, D., Tincas, I., & Melby-Lervåg, M. (2019). Socioeconomic background, nonverbal IQ and school absence affects the development of vocabulary and reading comprehension in children living in severe poverty. *Developmental Science*, 22(5), e12858. <https://doi.org/10.1111/desc.12858>.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers*, 36(1), 156–166. <https://doi.org/10.3758/BF03195560>.
- Ma, Y., Fan, H., & Chen, Q. (2024). Socioeconomic status and vocabulary development among children: A meta-analysis. *Current Psychology*, 43(39), 30736–30747. <https://doi.org/10.1007/s12144-024-06666-2>.
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300–335. <https://doi.org/10.3102/0034654310377087>.
- Meara, P., & Wolter, B. (2009). V_Links: Beyond vocabulary depth. In Paul Meara (ed.), *Connected words: Word associations and second language vocabulary acquisition* (pp. 73–84). John Benjamins Publishing Company. <https://benjamins.com/catalog/lllt.24.09vli>
- Nag, S. (2023). Socioeconomic status, sociocultural factors, and literacy development. In C. Perfetti, K. Pugh, L. Verhoeven, & S. Nag (Eds.), *Global variation in literacy development* (pp. 333–352). Cambridge University Press. <https://doi.org/10.1017/9781009242585.015>.
- Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *npj Science of Learning*, 2(1), 3. <https://doi.org/10.1038/s41539-017-0004-7>.
- Nation, P. (2001). Knowing a word. In ISP Nation, *Learning Vocabulary in Another Language* (pp. 23–59). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759.004>.
- Nation, P. (2019). The different aspects of vocabulary knowledge. In Stuart Webb (ed.), *The Routledge handbook of vocabulary studies*. Routledge.
- Neuman, S. B., Kaefer, T., & Pinkham, A. M. (2018). A double dose of disadvantage: Language experiences for low-income children in home and school. *Journal of Educational Psychology*, 110(1), 102–118. <https://doi.org/10.1037/edu0000201>.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554–566. <https://doi.org/10.1037/0022-0663.98.3.554>.
- Peets, K. F., & Bialystok, E. (2015). Academic discourse: Dissociating standardized and conversational measures of language proficiency in bilingual kindergarteners. *Applied PsychoLinguistics*, 36(2), 437–461. <https://doi.org/10.1017/S0142716413000301>.
- Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11, pp. 189–213). J. Benjamins Pub.
- Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental Psychology*, 45(2), 544–557. <https://doi.org/10.1037/a0013917>.

Rakesh, D., Lee, P. A., Gaikwad, A., & McLaughlin, K. A. (2024). Annual Research Review: Associations of socioeconomic status with cognitive function, language ability, and academic achievement in youth: A systematic review of mechanisms and protective factors. *Journal of Child Psychology and Psychiatry*, **66**(4), 417–439. <https://doi.org/10.1111/jcpp.14082>.

Read, J. (1998). *Validating a test to measure depth of vocabulary knowledge* (p. 60). Lawrence Erlbaum Associates Publishers.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 209–227). John Benjamins Publishing Company. <https://doi.org/10.1075/llt.10.15rea>.

Richard, G. J., & Hanner, M. A. (1995). *Language processing test-revised*. LinguSystems.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, **64**(4), 913–951. <https://doi.org/10.1111/lang.12077>.

Schoon, I., Jones, E., Cheng, H., & Maughan, B. (2012). Family hardship, family instability, and cognitive development. *Journal of Epidemiology and Community Health*, **66**(8), 716–722. <https://doi.org/10.1136/jech.2010.121228>.

Stata Statistical Software. (2019). *Stata Statistical Software: Release 16 (Version 16) Computer software*. StataCorp LLC.

Tong, X., & Tong, S. X. (2022). How vocabulary breadth and depth influence bilingual reading comprehension: Direct and indirect pathways. *Learning and Individual Differences*, **100**, 102227. <https://doi.org/10.1016/j.lindif.2022.102227>.

Van Der Kleij, S. W., Burgess, A. P., Ricketts, J., & Shapiro, L. R. (2023). Tracking vocabulary and reading growth in children from lower and higher socioeconomic backgrounds during the transition from primary to secondary education. *Child Development*, **94**(1), 1–10. <https://doi.org/10.1111/cdev.13862>.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, **22**(2), 217–234. <https://doi.org/10.1017/S0142716401002041>.

Von Stumm, S., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Preschool verbal and nonverbal ability mediate the association between socioeconomic status and school performance. *Child Development*, **91**(3), 705–714. <https://doi.org/10.1111/cdev.13364>.

Webb, S. (2012). Depth of vocabulary knowledge. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (1st ed.). Wiley. <https://doi.org/10.1002/9781405198431.wbeal1325>.

Webb, S. (Ed.) (2020). *The Routledge handbook of vocabulary studies*. Routledge, Taylor & Francis Group.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children (WISC-V)*. <https://psycnet.apa.org/record/2018-36604-011>

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, **53**(1), 13–40. <https://doi.org/10.3138/cmlr.53.1.13>.

Wiig, E. H., & Secord, W. (2012). *Test of Word Knowledge [Dataset]*. <https://doi.org/10.1037/t15161-000>

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock- Johnson psychoeducational battery* (3rd ed.). Riverside Publishing.

Yanagisawa, A., & Webb, S. (2019). Measuring depth of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., pp. 371–386). Routledge.

Annexes

Annex 1. List of words used in the attributes and the self-assessment tasks

GRAM	ORTHO	NBLET	FREQ
NC	Chou-fleur [cauliflower]	10	3
NC	Carrelage [tiles]	9	3
NC	Ventilateur [fan]	11	3.1
NC	Kimono [kimono]	6	3.5
NC	Gobelet [cup]	7	7.9

Annex 1. (Continued)

GRAM	ORTHO	NBLET	FREQ
NC	Loutre [otter]	6	8.9
NC	Cravate [tie]	7	10.2
NC	Ruche [beehive]	5	13.3
NC	Poubelle [trash can]	8	21.5
NC	Charrette [trolley]	9	28.9
NC	Poire [pear]	5	29.9
NC	Marteau [hammer]	7	34.3
NC	Coquille [shell]	8	35.4
NC	Haricot [bean]	7	36.7
NC	Bonnet [beanie]	6	37
NC	Crocodile [crocodile]	9	56
NC	Montre [watch]	6	78.1
NC	Oreille [ear]	7	289.7
NC	Train [train]	5	326.9
NC	Cheval [horse]	6	493.3

Note: GRAM: grammatical category; ORTHO: orthographic form; NBLET: number of letters; FREQ: word frequency.

Cite this article: Vargas Cuevas, J., Facon, B., Javourey-Drevet, L., & Casalis, S. (2026). Beyond breadth of vocabulary: Vocabulary depth can also be impacted by school's socioeconomic status. *Journal of Child Language* 1–15, <https://doi.org/10.1017/S0305000926100579>