Journal of Experimental Child Psychology 214 (2022) 105312

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Contextual diversity favors the learning of new words in children regardless of their comprehension skills

Eva Rosa [a,*], Rafael Salom [b], Manuel Perea [c,d,e]

[a] Departamento de Psicología Básica and ERI-Lectura, Universitat de València, Valencia, Spain
[b] Universidad Católica de Valencia San Vicente Mártir, Valencia, Spain
[c] Departamento de Metodología and ERI-Lectura, Universitat de València, Valencia, Spain
[d] Centro de Investigación Nebrija en Cognición, CINC, Universidad Antonio de Nebrija, Madrid, Spain
[e] Basque Center of Cognition, Brain, and Language, Donostia, Spain

ARTICLE INFO

Article history:
Received 14 April 2021
Revised 29 September 2021

Keywords:
Incidental vocabulary learning
Young readers
Contextual diversity
Reading comprehension skills
Semantic distinctiveness model
Naturalistic setting

ABSTRACT

Recent research has shown the benefits of high contextual diversity, defined as the number of different contexts in which a word appears, when incidentally learning new words. These benefits have been found both in laboratory settings and in ecological settings such as the classroom during regular hours. To examine the nature of this effect in young readers aged 11–13 years, we analyzed whether these benefits are modulated by the individuals' reading comprehension scores; that is, would better comprehenders benefit the most from contextual diversity? The manipulation of contextual diversity was done by inserting the novel words into three different contexts/topics, or into only one of them, while keeping constant their frequency of occurrence. Results showed that words encountered in different contexts were learned more effectively than those presented in the same context. More important, the effect of contextual diversity was similar regardless of the participants' comprehension skills. We discuss the implications of these findings for models of word learning and the practical applications in curriculum design.

© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author.
E-mail address: eva.rosa@uv.es (E. Rosa).

https://doi.org/10.1016/j.jecp.2021.105312
0022-0965/© 2021 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

The higher the number of contexts (e.g., films, books) in which a word appears, the faster it is identified on word recognition tasks (Adelman, Brown, & Quesada, 2006; Brysbaert, New, & Keuleers, 2012; see also Perea, Soares, & Comesaña, 2013, for evidence with developing readers) and during sentence reading (Pagán & Nation, 2019; Plummer, Perea, & Rayner, 2014). This phenomenon has been called the *contextual diversity effect.* However, the beneficial impact of the larger number of contexts is not restricted to word recognition and reading; it has also been shown when incidentally learning new words (Hills, Maouene, Riordan, & Smith, 2010; Hoffman & Woollams, 2015; Jones, Johns, & Recchia, 2012; Rosa, Tapia, & Perea, 2017).

Regarding incidental word learning, which was the focus of this experiment, a facilitative effect of contextual diversity has been obtained using various paradigms. Hills et al. (2010) carried out a correlational analysis of the CHILDES (Child Language Data Exchange System) database (MacWhinney, 2000), which contains child-directed speech toward 12- to 60-month-old children. They found that children more rapidly learned words that occurred in many contexts. Using an experimental methodology, Johns, Dye, and Jones (2016) found an effect of contextual diversity in a laboratory word learning experiment with adults. The participants—university students—were presented with small fragments extracted from articles, books, and newspapers. Critically, several low-frequency words were replaced by pronounceable nonwords, which were the target stimuli. Half of these target stimuli were inserted into five highly distinctive texts, whereas the other half were inserted into five redundant texts. During the training phase, participants needed to read the texts and rate the extent to which they understood each passage on a 7-point Likert scale. To test the effect of contextual diversity, Jones, Dye, & Johns (2017) employed a pseudolexical decision task with the novel words from the training phase and a set of nonword foils. They found faster and more accurate responses when the novel words learned were from distinctive texts (i.e., high diversity) than when they were from redundant texts.

Along the same lines, Rosa et al. (2017) examined the role of contextual diversity when acquiring new words with 8- and 9-year-old primary school students in a naturalistic setting. Specifically, they tested the incidental learning of words when reading texts in the classroom. Half of the novel words were presented in three texts related to three different subjects: Spanish language, natural sciences, and mathematics (“high diversity”). The other half were presented in three texts related to only one of these subjects (“low diversity”). Incidental word learning was assessed through two memory tasks (free recall and recognition), a word–pictogram matching task, and a sentence completion task with multiple-choice answers. On this latter task, participants needed to choose the target stimulus from three distractors that were orthographically or phonologically similar to the target. Rosa et al. (2017) found that the novel words were learned better when presented in semantically distinctive texts on all the tasks.

These effects of contextual diversity when learning new words can be easily captured by Jones et al.’s (2012) *semantic distinctiveness model.* In this model, the strongest memory encoding of a novel word occurs with its first encounter. Each subsequent encounter with this word contributes to its consolidation. Specifically, every time we encounter a word, the cognitive system estimates the similarity between the current context and the word’s current memory representation in terms of redundant information (i.e., the proportion of semantic information shared by both). The greater the difference between the current context and the contexts in which the word has been previously experienced, the greater the strength of the memory encoding. As a result, more diverse contexts would produce a more accessible representation than less diverse contexts, thereby capturing the beneficial effect of contextual diversity when learning new words.

The critical issue in the current experiment was whether the mechanisms described by Jones et al. (2012) in their semantic distinctiveness model are affected by the participants’ reading comprehension skills. Clearly, examining the relationships between contextual diversity and reading comprehension skills would help to clarify the mechanisms responsible for the benefits of increasing the number of contexts during word learning; would better comprehenders benefit the most from an increase in the number of contexts when learning novel words? We can consider two different hypotheses. First, one might argue that the better the reader's ability to understand the texts and infer information from them, the more efficient the computation of the semantic similarity of the different contexts where a word appears will appear. Thus, better comprehenders would benefit the most from learning novel words in a high-contextual diversity setting. Second, one might argue that the processes underlying the contextual diversity effect may mainly reflect a general mechanism of encoding variability across different episodes rather than purely semantic variations (see Jones et al., 2017, for a suggestion of this possibility). According to this hypothesis, the modulating role of the participants' reading comprehension skills in the effects of contextual diversity when learning new words would be small or negligible.

Previous research on this issue is quite scarce. Neither Johns et al. (2016) nor Rosa et al. (2017) obtained a measure of the participants' comprehension skills. Notably, Joseph and Nation (2018) conducted an eye movement experiment with 10- and 11-year-old children that analyzed the relationship between contextual diversity and reading comprehension skills. They used low-frequency verbs as target words embedded in 10 semantically redundant sentences (low contextual diversity) and in 10 semantically diverse sentences (high contextual diversity). For example, in the low-diversity condition, the target word *accumulated* was embedded in sentences related to law (e.g., "The police accumulated a lot of strong evidence," "The lawyer accumulated witness statements"); in contrast, in the high-diversity condition, it was embedded in sentences from different semantic fields (e.g., "Lava had accumulated beneath the surface," "His debts accumulated until he had to sell"). To evaluate learning, they measured several dependent variables: the changes in the eye movement pattern relative to the pretest phase and three offline posttests. The offline posttests consisted of a spelling task that tapped the acquisition of the orthographic form and two tasks designed to measure how well the children had learned the meaning (i.e., a task where they needed to complete sentences using the target words and a task where they needed to judge the plausibility of sentences containing the target words). In addition, the participants' comprehension skills were measured through two standardized tests that consisted of answering comprehension questions about two passages they had read.

Joseph and Nation's (2018) results showed evidence of orthographic and semantic incidental learning of the novel words on all the measures described above. The authors also found that those who scored the best on reading comprehension skills performed better on all the assessment tasks. However, contextual diversity effects were not reported on any of the dependent variables. Joseph and Nation (2018) argued that to bring out the effects of contextual diversity, more than 10 exposures might have been necessary. Another explanation is that the lack of effect could have been due to the way in which contextual diversity was manipulated given that they used brief sentences rather than full texts. When using full texts, Rosa et al. (2017) reported sizable effects of contextual diversity after only three exposures to the novel words.

Regarding the interaction between reading comprehension skills and contextual diversity, Joseph and Nation (2018) did not find any significant effects on offline posttests. However, they did find an interaction on several eye movement measures. For the high-diversity condition, reading times were similar for children with good and poor reading comprehension skills. In contrast, for the low-diversity words, better comprehenders spent less time reading. To explain this pattern of results, Joseph and Nation (2018) argued that good comprehenders would spend more time inferring the meanings of the words presented in diverse contexts, whereas bad comprehenders would always need a relatively long reading time. A parallel argument can also explain the discrepancy between the findings with online and offline measurements. The dependent variables measured through eye movements would reflect the time children spend deducing the meanings of words under the various experimental conditions rather than the final level of word learning they achieve. Instead, offline posttests have the added value of measuring word learning because they directly test the acquisition of the orthographic form and the meanings of the target words.

Taking the previous findings into account, we believe that it is necessary to examine the interplay between contextual diversity and comprehension skills in children using an experimental design that has previously been shown to induce robust contextual diversity effects.

Aims of the current study

The main goal of the current study was to examine whether the participants' reading comprehension skills modulated the effects of contextual diversity on incidental word learning in young readers. As in the Rosa et al. (2017) experiment, we manipulated contextual diversity in a naturalistic environment while students read texts during their regular classes. Participants read texts related to three school subjects: Spanish language, natural sciences, and mathematics. Each novel word was presented three times during the learning phase either in three different contexts/subjects (high contextual diversity) or in only one of them (low contextual diversity). The meanings and grammatical forms of the new words were kept constant in all the contexts. The differences compared with the Rosa et al. (2017) experiment were that (a) participants were slightly older (11–13 years vs. 8–9 years), thereby enabling us to generalize the findings to a new age range, and (b) we measured the participants' comprehension skills with a standardized reading test, which allowed us to examine their relationship with the effect of contextual diversity. We chose to examine whether contextual diversity interacts with reading comprehension skills in developing readers instead of adults because variability in reading comprehension skills is greater in this age range than in university students, who are generally good comprehenders. Similar to the Rosa et al. (2017) experiment, we employed a multiple-choice task, a word–picture task, and a recognition task. The multiple-choice task consisted of completing sentences by distinguishing the target words from three lexical distractors that were orthographically or phonologically similar to the target. Thus, this task required the acquisition of the meanings and spellings of the new words and the comprehension of the sentences. In the word–pictogram matching task, participants were shown the 12 target words along with 12 randomly arranged pictures, and they needed to match each word to the picture that represented it. Therefore, this task required the acquisition of the meanings of the target words. On the recognition task, children needed to discriminate the target words from a set of filler words with very low frequency that essentially assessed the strength of memory encoding.

The predictions of the experiment were clear. First, regarding contextual diversity, we expected to obtain better learning in the high-diversity condition than in the low-diversity condition, extending the Rosa et al. (2017) findings to 11- to 13-year-old children. Second, readers with poor reading comprehension skills may find it more difficult to infer the meanings of the novel words from the context. Thus, we expected lower incidental word learning in participants with poor reading comprehension skills than in those with good reading comprehension skills (Bolger, Balass, Landen, & Perfetti, 2008; Cain, Oakhill, & Elbro, 2003; Cain, Oakhill, & Lemmon, 2004; Joseph & Nation, 2018). Third, regarding the main aim of the current experiment, which was to find out whether there is an interaction between reading comprehension and contextual diversity, there are two possible patterns of results. On the one hand, participants with better reading comprehension skills could show a greater effect of contextual diversity. This outcome would show that contextual diversity when learning novel words would be heavily rooted in semantics; good comprehenders would be more efficient in extracting the meanings of novel words from the context. On the other hand, the contextual diversity effect could be approximately the same size regardless of the individuals' reading comprehension skills. This outcome would favor the idea that, when learning new words, the effect of contextual diversity is dependent on a general encoding variability effect, presumably via episodic sources not related to reading comprehension skills.

Method

Participants

A total of 47 children were recruited from a middle-class high school in Spain. Of the initial 47 participants, we excluded 4 children, 3 who answered fewer than two comprehension questions correctly in any training session and 1 who missed at least one of the experimental sessions. This left a final sample of 43 children (22 boys). The mean age was 12.63 years (range = 11–13). Note that Rosa et al. (2017) reported a robust contextual diversity effect using a slightly smaller sample size (N = 33). All the participants spoke Spanish as their first language. They had no reading difficulties that would keep them from taking part in the experiment, as shown by their overall PROLEC-SE battery scores (Ramos & Cuetos, 1999). This study was approved by the experimental research ethics committee of the University of Valencia. Written informed consent was obtained from parents of all the children before they participated in the experiment.

Materials

The target stimuli were 12 concrete words in Spanish with unambiguous meanings (average length = 6.7 letters, range = 5–9). These words did not occur in the LEXIN primary school lexical database in Spanish (Corral, Ferrero, & Goikoetxea, 2009), and they had a very low frequency of use in the EsPal subtitle database (mean = 0.26 per million words, range = 0.01–0.70; Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013). We asked a representative sample of third-grade children (N = 20), who were not part of the experimental sample, about the meanings of these words, and we verified that they did not know them. Appendix A presents the list of target words in Spanish and English and their frequency of use.

We created two counterbalanced sets of materials so that each word could be learned in a high- or low-diversity context. In Set A, six of the target words appeared in a high-diversity context (i.e., inserted into the three types of texts) and the remaining six words appeared in a low-diversity context (i.e., inserted into only one of the three types of text). The opposite procedure was employed for Set B. Each set consisted of nine short texts: three fables, three expository texts with natural or social science contents, and three texts with simple math exercises. All the texts were equal in length (212 words) and difficulty and were appropriate for the reading level of the participants, as assessed by the school's teachers. Four different target words were inserted into each text, making sure that each of them appeared just once in each text—always with the same gender and number—and that the same target words did not ever coincide in different texts in the high-diversity condition. Tables 1 and 2 show how the target words were inserted into the different texts in Experiment Sets A and B, respectively.

The three types of texts—fables, science texts, and math problems—were constructed in such a way as to allow participants to derive the meaning of each new word without the need for an explicit description. For example, for the word *vulpeja* (vixen), the texts from the high- and low-diversity conditions contained similar clues about the characteristics of this animal, namely that it has a reddish tail, it has sharp fangs, it is omnivorous, or it goes hunting at night. To ensure that the students were reading the texts comprehensively, at the end of each text we added two reading comprehension questions with three possible response options, of which only one was correct.

Here is an example of how the word *vulpeja* was inserted into three sentences from the three types of text in the high-diversity condition. Fable: "... una vulpeja de lo más astuta, con los colmillos afilados y la cola rojiza ..." ("... a very cunning vixen with sharp fangs and reddish tail ..."). Science text: "La vulpeja se encuentra en casi cualquier hábitat del hemisferio norte: las praderas, las zonas costeras, la tundra alpina, la taiga, o las mesetas montañosas ..." ("The vixen is found in almost any habitat in the northern hemisphere: grasslands, coastal areas, alpine tundra, taiga, or mountain plateaus ..."). Math problems: "Una astuta vulpeja persigue a su cachorro que se ha escapado corriendo. La madre corre a una velocidad de 4 metros por minuto ..." ("A cunning vixen goes after her cub who has run away. The mother runs at a speed of 4 meters per minute ..."). Appendix B shows another example of the three complete texts in which the word *ualabí* (wallaby) was inserted.

Evaluation instruments

To assess the acquisition of the newly learned words, we employed three tasks (a) a multiple-choice task with lexical distractors that were orthographically or phonologically similar to the target words, (b) a task that required matching words to pictograms, and (c) a recognition task. There was no time limit on any of the tasks.

Multiple-choice task

To create the multiple-choice task, we employed a subset of the Collective Test of Reading Efficiency (*Test Colectivo de Eficacia Lectora* [TECLE]; Marín & Carrillo, 1999). The number of items was reduced to

Table 1
Distribution of target words in the texts from Experimental Set A.

Fable 1	Fable 2	Fable 3
Sargazo	Ualabí	Simientes
Zarceta	Vulpeja	Valvas
Adelfa	Adelfa	Adelfa
Ofidio	Ofidio	Ofidio
Science Text 1	Science Text 2	Science Text 3
Simientes	Sargazo	Ualabí
Vulpeja	Valvas	Zarceta
Cubil	Cubil	Cubil
Piélago	Piélago	Piélago
Math Problems 1	Math Problems 2	Math Problems 3
Ualabí	Simientes	Sargazo
Valvas	Zarceta	Vulpeja
Esqueje	Esqueje	Esqueje
Ponzoña	Ponzoña	Ponzoña

Note. Target words belonging to the high-contextual diversity condition: Sargazo, Simientes, Ualabí, Valvas, Vulpeja, and Zarceta. Target words belonging to the low-contextual diversity condition: Adelfa, Cubil, Esqueje, Ofidio, Piélago, and Ponzoña.

Table 2
Distribution of target words in the texts from Experimental Set B.

Fable 1	Fable 2	Fable 3
Adelfa	Esqueje	Cubil
Ponzoña	Piélago	Ofidio
Sargazo	Sargazo	Sargazo
Valvas	Valvas	Valvas
Science Text 1	Science Text 2	Science Text 3
Cubil	Adelfa	Esqueje
Piélago	Ofidio	Ponzoña
Simientes	Simientes	Simientes
Vulpeja	Vulpeja	Vulpeja
Math Problems 1	Math Problems 2	Math Problems 3
Esqueje	Cubil	Adelfa
Ofidio	Ponzoña	Piélago
Ualabí	Ualabí	Ualabí
Zarceta	Zarceta	Zarceta

Note. Target words belonging to the high-contextual diversity condition: Adelfa, Cubil, Esqueje, Ofidio, Piélago, and Ponzoña. Target words belonging to the low-contextual diversity condition: Sargazo, Simientes, Ualabí, Valvas, Vulpeja, and Zarceta.

12, one for each target word. As in the Marín and Carrillo (1999) test, each item was made up of an incomplete sentence with the last word missing and four possible response options, only one of which was appropriate to finish the sentence. In addition, as in the Marín and Carrillo test, the foils were constructed. Two of them were pseudowords that differed in only one letter from the target word, and the third foil was phonologically similar to but orthographically different from the target word.

Picture–word matching task

To build the picture-word matching task, we conducted a web search of pictograms, drawings, and free distribution images that represented each target word. The test was made up of the 12 selected pictures that were displayed along with the 12 target words in random order. Participants were required to select the image that corresponded to each word.

Recognition task

For the recognition task, we used a total of 54 filler words that did not appear in the LEXIN primary school lexical database in Spanish (Corral et al., 2009). These 54 words were randomly presented to the participants along with the 12 target words, thereby yielding a total of 66 words. The participants' task was to identify the words they had read in the training texts.

Reading comprehension skills assessment

To evaluate participants' reading skills, we used a standardized reading test in Spanish (PROLEC-SE; Ramos & Cuetos, 1999). This instrument provides an overview of the presence of difficulties in each reading process: syntactic processing, the comprehension of explicit and implicit information in expository texts, extracting meaning from expository texts through the elaboration of an outline, the lexical route by reading words with different frequencies and lengths, the phonological route by reading short and long pseudowords with simple and complex syllables, and the use of punctuation marks in syntactic processing. The text comprehension scores were subsequently analyzed to examine their relationship with contextual diversity. The reading comprehension test was administered collectively. Students were asked to read two texts and answer 10 literal and inferential questions about each text in an open format.[1] The scores on the other scales were used only to rule out any potential reading difficulties in the participants.

The experimental materials and data are openly available on the Open Science Framework (https://osf.io/xje5g/?view_only=d3c5dad1b2ba4d49a5dafd88b7dd0479).

Procedure

All the students completed the training phase and the evaluation phase in groups during their regular classes. Before starting the training phase, participants were told that they would need to carefully read simple texts and understand them. They were also told that the text could contain words that they would not know and that they should guess the meanings from the context while trying to understand the general meaning of the text. Then, they were randomly assigned to one of the two experimental sets and were asked to read their corresponding nine texts during the training phase. On each of the three training days, the students read a fable, an expository text with science content, and a text with math problems. As stated above, each text contained 4 of the 12 experimental words. Hence, students read each of the 12 target words three times in their corresponding experimental condition—high versus low contextual diversity. At the end of each text, students needed to answer two reading comprehension questions with three possible answers, of which only one was correct. They had unlimited time to read the texts and answer the comprehension questions. Text presentation was randomized for each student to minimize potential primacy/recency effects in the evaluation phase.

On the fourth day, after completing the training phase, experimental word learning was assessed through the evaluation instruments previously introduced. To equate carryover effects from one task to another, the assessment order was the same for all the participants. It began with the recognition task, which provides less semantic, orthographic, or visual information about the experimental words, and it ended with the pictogram task, which provides more information about the spellings and meanings of the words.

Results

The average number of correct answers per session on the text comprehension questions was 5.23 out of 6 (range = 4–6), thereby showing that the students effectively read for comprehension. Table 3 shows descriptive statistics (means and standard errors) for the high- and low-contextual diversity conditions on the three evaluation instruments.

[1] The test used by Joseph and Nation (2018) to measure reading comprehension was quite similar. In this case, students were asked to read two passages and then answer eight comprehension questions about each one. However, their test was administered individually, and so the students needed to read the texts aloud.

Table 3
Mean percentages correct (and standard deviations) of the three dependent variables for high-contextual and low-contextual diversity along with the contextual diversity effect.

	Multiple-choice task	Picture–word matching task	Recognition task
High contextual diversity	0.63 (0.03)	0.52 (0.03)	0.32 (0.03)
Low contextual diversity	0.50 (0.03)	0.34 (0.03)	0.27 (0.03)
Contextual diversity effect	0.13	0.18	0.05

For the inferential tests, we created Bayesian linear mixed-effects models using the *brms* package (Bürkner, 2018) in R (R Core Team, 2020). The fixed factors were contextual diversity (low [encoded as 0.5] vs. high [encoded as $-$ 0.5]) and the comprehension test scores using standardized values. We employed the most complex random-effects structural model:

$$\text{Dependent variable CD} * \text{zcomp} + (1 + \text{CD} * \text{zcomp}|\text{participant}) + (1 + \text{CD} * \text{zcomp}|\text{item}).$$

Given that the dependent variable was accuracy (1 = correct and 0 = incorrect), it was modeled with the Bernoulli distribution (i.e., family = Bernoulli). Note that family = Bernoulli is the *brms* parallel of family = binomial of generalized linear mixed-effects models in the *lme4* package. We conducted separate analyses for each task: multiple-choice, picture–word matching, and recognition tasks. In all cases, we conducted 5000 iterations with four chains. The fits were successful ($\hat{R}$ = 1.00 in all cases). The output of these models specifies the coefficient, standard error and 95% credible interval (95% CrI) of each effect. An effect was interpreted as significant when its 95% CrI did not cross zero.

Multiple-choice task

We found higher accuracy for words learned under high contextual diversity than under low contextual diversity, b = −0.70, SE = 0.31, 95% CrI [−1.32, −0.10]. In addition, participants with better reading comprehension skills performed better than those with worse reading comprehension skills, b = 0.53, SE = 0.24, 95% CrI [0.08, 1.03]. Importantly, there were no signs of an interaction between contextual diversity and reading comprehension skills, b = 0.10, SE = 0.26, 95% CrI [−0.42, 0.61] (see Fig. 1 for the plot with the predicted probabilities).

Picture–word matching task

Accuracy was higher for words learned with high contextual diversity than for those learned with low contextual diversity, b = −0.86, SE = 0.30, 95% CrI [−1.47, −0.29]. There were no signs of an effect

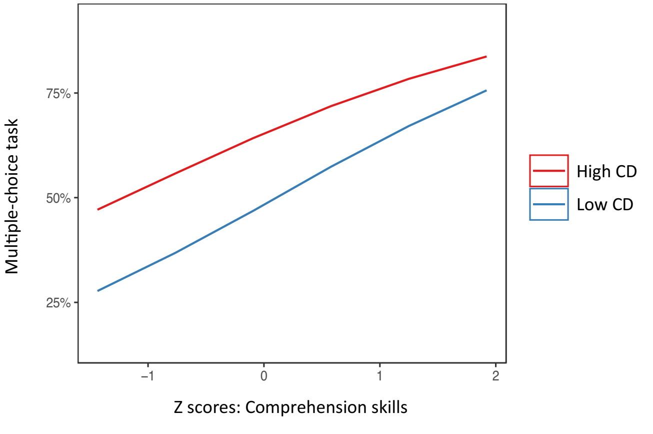

Fig. 1. Predicted probabilities for the multiple-choice task. The upper curve shows high contextual diversity (CD), and the lower curve shows low CD.

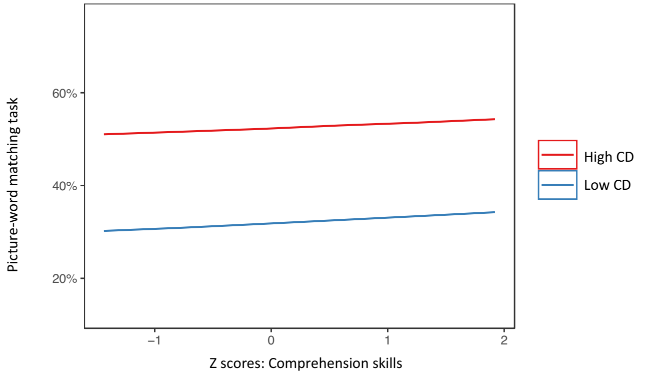

Fig. 2. Predicted probabilities for the picture–word matching task. The upper curve shows high contextual diversity (CD), and the lower curve shows low CD.

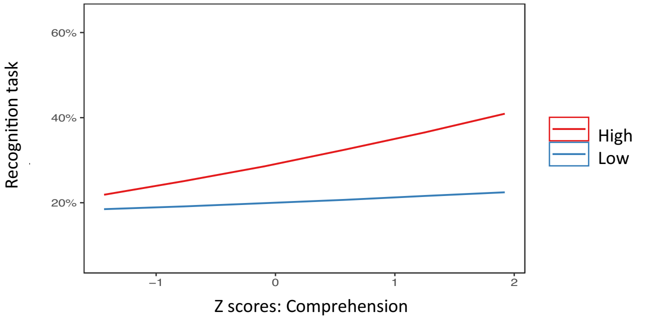

Fig. 3. Predicted probabilities for the recognition task. The upper curve shows high contextual diversity (CD), and the lower curve shows low CD.

of reading comprehension skills, b = 0.04, SE = 0.25, 95% CrI [−0.45, 0.54], or an interaction between the two factors, b = 0.01, SE = 0.28, 95% CrI [−0.54, 0.56] (see Fig. 2 for the plot with the predicted probabilities).

Recognition task

On this task, accuracy on the novel words was not significantly greater in the high-contextual diversity condition than in the low-contextual diversity condition, b = −0.51, SE = 0.37, 95% CrI [−1.27, 0.18]. Neither the effect of reading comprehension skills nor the interaction between the two factors was significant, b = 0.27, SE = 0.21, 95% CrI [−0.15, 0.68] and b = 0.20, SE = 0.38, 95% CrI [−0.96, 0.55], respectively (see Fig. 3 for the plot with the predicted probabilities).

Discussion

The current experiment tested whether, for young readers in a classroom setting, their reading comprehension skills modulated the effect of contextual diversity on incidental word learning. To this end, participants were presented with a set of novel words embedded in three contextually different texts (fable, science, and math) or three contextually similar texts (always fable, science, or math). As expected, the children learned the novel words better in semantically different texts than in semantically similar texts (see Hoffman & Woollams, 2015; Jones et al., 2012; Rosa et al., 2017). Furthermore, on a multiple-choice task, good comprehenders performed better than poor comprehenders. As suggested in the Introduction, the multiple-choice task is probably a better marker of comprehension skills than the other two assessment tasks we employed; it involves not only acquiring the novel words' meanings but also comprehending the sentences to be completed.

More important, we found no signs of an interaction between contextual diversity and the participants' reading comprehension skills on any of the measures (see Figs. 1–3). At a theoretical level, this pattern has important implications for the semantic distinctiveness model (Jones et al., 2012). In the current implementation of this model, the strength of a word's encoding on memory depends on the semantic variability of the different contexts in which we see that word. If the effect of contextual diversity had been purely semantic, one would have expected better comprehenders to benefit more from learning words embedded in different texts; better comprehenders would take more advantage of the semantic dissimilarity of these texts. However, results showed that young readers benefitted equally from contextual diversity regardless of their comprehension skills. These results favor the view that the effect of contextual diversity when learning words is related to a general variability encoding mechanism. Indeed, Johns et al. (2016) anticipated this possibility when they proposed that the effect of contextual diversity "could also be seen as an episodic effect" and that contextual diversity "could be interpreted as an encoding variability manipulation (Bower, 1970), in which distinctive contexts lead to differential encoding, resulting in the observed differences in task performance" (p. 1219). Consistent with this view, we suggest that the mechanism responsible for encoding context variability could also integrate other factors apart from semantic and other linguistic aspects of the context. One illustration is the manipulation of contextual diversity in terms of the physical features of the context. Tapia, Rosa, Rocabado, Vergara-Martínez, and Perea (2021) found that incidental learning of novel words in 8- and 9-year-old children was better when several individuals uttered the fables than when they were narrated by the same individual while keeping the semantic information constant. Taken together, these findings suggest that there is a nonsemantic source in the contextual diversity effect when learning novel words.

At a more general level of theorizing, memory research has repeatedly shown that the retrieval of a given element is more likely when it is involved in a rich network than when it is encoded in isolation (e.g., for early research, see Anderson & Bower, 1972; Craik & Tulving, 1975; Tulving & Thomson, 1973) and that it is not the number of repeated exposures but rather the distinction of these exposures in terms of time and context that affects this recovery (see Glenberg, 1976, 1979). The hypothesis of contextual variability (see Bolger et al., 2008) transfers this idea to the case of vocabulary learning. This hypothesis states that every encounter with a word in a distinct context creates a memory trace with the context's episodic information so that the more memory traces a word has, the more easily it will later be retrieved (Fukkink, Blok, & de Glopper, 2001; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985). Therefore, distributional models of semantic representations that aim to explain the cognitive mechanisms underlying the learning of new words (e.g., semantic distinctiveness model; Jones et al., 2012) may need to reconsider an extended conceptualization of "context." In other words, these models should consider not only semantic factors but also other linguistic and nonlinguistic factors such as perceptual features.

Our findings also have practical implications for incidental vocabulary learning at an applied level by reading in classroom settings. First, in this study the beneficial effects of conducting cross-learning through different topics to optimize the process of acquiring new words in young readers aged 8 and 9 years (Rosa et al., 2017) have been extended to readers aged 11–13 years. Thus, we encourage educators to take contextual diversity into account in the curriculum design of each educational stage by planning the learning of the essential lexicon through different subjects. Providing the right conditions for children to learn new vocabulary in an incidental way by reading materials on different topics would be beneficial in terms of language proficiency and content proficiency in the various academic learning areas.[2]

Second, and most important, *all* readers—even those with poor comprehension skills—can benefit from contextual diversity to learn new words, making it a very effective teaching tool. We acknowledge, however, that to fully understand the relationship between reading comprehension skills and

[2] We should bear in mind that children learn words mainly incidentally through storytelling from a very early age and through their reading experience from mid-childhood onward (Biemiller & Slonim, 2001; Robbins & Ehri, 1994).

contextual diversity, future work should also use other learning assessment tasks such as sentence construction using novel words or comprehension questions about texts. This would provide a more comprehensive way of assessing the meanings of the experimental words and their use in specific contexts. Furthermore, given that the encoding mechanism underlying the contextual diversity effect may be modulated by nonsemantic factors, various other ways of manipulating variability, such as physical features of contexts, may shed more light on the nature of this mechanism (e.g., using virtual reality). Moreover, further experimental research is needed to test the effectiveness of contextual diversity in second-language vocabulary learning, which is the most common type of learning in adults.

Conclusion

Contextual diversity is a powerful enhancer of incidental learning of vocabulary. We have shown that, in young readers, the effect of contextual diversity when learning novel words is independent of their reading comprehension skills. This finding has clear implications for theoretical models such as the semantic distinctiveness model. As indicated above, these models need to extend their conceptualization of what a "context" is, and they need to consider not only semantic factors but also other linguistic and nonlinguistic factors. In practical terms, our findings suggest that it is possible to take advantage of a variety of "contexts" to improve incidental vocabulary learning regardless of the individuals' reading comprehension skills.

Declaration of Competing Interest

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

Acknowledgment

Funding was provided by the Spanish Ministry of Science, Innovation, and Universities (PSI2017-86210-P).

Appendix A

Target words in Spanish (and English) and their frequency of use in the EsPal subtitle database

Target word	Frequency of use	Number of syllables
Adelfa (oleander)	.05	3
Ofidio (ophidian)	.03	3
Valvas (seashells)	.00	2
Simientes (seeds)	.03	3
Ponzoña (venom)	.14	3
Zarceta (green-winged teal)	.00	3
Ulabí (wallaby)	.00	3
Cubil (lair)	.31	2
Piélago (ocean)	.02	3
Vulpeja (vixen)	.00	3
Sargazo (sargasso)	.01	3
Esqueje (cutting)	.03	3

[2] We should bear in mind that children learn words mainly incidentally through storytelling from a very early age and through their reading experience from mid-childhood onward (Biemiller & Slonim, 2001; Robbins & Ehri, 1994).

Note. For some of the experimental words, we did not find an English translation with a frequency of use as low as its equivalent in Spanish.

Appendix B

Shown are examples in Spanish and English of the three types of text in which the target word *ualabí* (the Spanish word for *wallaby*) was inserted. For ease of presentation, the target words are shown here bolded and underlined; they were not bolded or underlined in the experiment.

Fable

En Australia, se encontraban Eloi, un **ualabí** muy grande, que daba los saltos más altos de toda su especie; y Olga, una vulpeja de lo más astuta, con los colmillos afilados y la cola rojiza; discutiendo por ver quién se merecía comerse la adelfa que ese encontraba frente a ellos. Era el arbusto más grande de los alrededores, lleno de flores y sabrosos frutos. Ante la imposibilidad de llegar a un acuerdo amistoso, decidieron medir sus fuerzas en un combate para decidir quién se quedaba definitivamente con tal festín.

Tan concentrados estaban Eloi y Olga en medir sus fuerzas, que no se dieron cuenta de que un escurridizo y alargado ofidio se deslizaba por el suelo hacia la planta. Era tan pequeño que les estaba robando toda la comida delante de sus narices. Un robo del que se enteraron cuando el agotamiento les hizo desistir de su propósito.

Mientras él disfrutaba de su enorme festín, los otros dos casi sin aliento exclamaron,

> ¡Qué tontos que hemos sido! ¡Tanta pelea y energía gastada para que al final todo el fruto de nuestro esfuerzo se lo haya llevado un oportunista que lo único que ha hecho es aprovecharse de la situación y sobre todo de nuestra estupidez!

> Moraleja: Vale más compartir que combatir.

Fable

In Australia, Eloi, a very large **wallaby** that made the highest jumps of its entire species, and Olga, a very cunning vixen with sharp fangs and a reddish tail, were arguing about who deserved to eat the oleander that was in front of them. It was the largest shrub around, full of flowers and tasty fruits. Faced with the impossibility of reaching a friendly agreement, they decided to compare their strength in a fight to decide who would definitely keep this feast.

So focused were Eloi and Olga on measuring their strength that they did not notice a long elusive ophidian sliding across the ground toward the plant. He was so small that he was stealing all their food right under their noses. They learned about this robbery when exhaustion made them give up their goal.

As he enjoyed his huge feast, the other two, almost out of breath, exclaimed,

> What fools we have been! So much fighting and energy expended so that in the end all the fruit of our labor has been taken by an opportunist who has taken advantage of the situation, and especially of our stupidity!

> Moral: It is better to share than to fight.

Science text

En el mundo coexisten unos 7,77 millones de especies animales, de las cuales solo 953.434 se han catalogado. Hay una gran variedad de animales, de diferentes colores y tamaños, que viven en distintos hábitats, y que tienen diferentes tipos de alimentación. Dos ejemplos muy distintos de esta gran diversidad de animales son:

El **ualabí**, que posee grandes y poderosas patas traseras, grandes pies aptos para saltar, una cola larga y musculosa para mantener el equilibrio, y una cabeza pequeña. Es herbívoro, se alimenta de pasto y de raíces durante las tardes y noches frías, generalmente en grupo. Esta especie es nocturna, usualmente pasa el día en quietud dentro de su cubil. Tienen una esperanza de vida de 18 años aproximadamente.

La zarceta, es un animal que mide unos 35 cm de longitud. Tiene un pico plano, redondeado y de un tono pardo o marrón oscuro, a veces moteado o con manchas negras. Sus patas son grisáceas. Se desarrolla en torno a zonas con importantes masas de agua. Las lagunas cercanas a los árboles y con abundante vegetación acuática son un ejemplo de hábitat adecuado, aunque hay variantes de esta especie que abandonan las lagunas y se desplazan al grande y azul piélago, para buscar comida que flota en la superficie del agua.

Science text

Approximately 7.77 million animal species coexist in the world, of which only 953,434 have been cataloged. There are a wide variety of animals of different colors and sizes that live in different habitats and eat different types of food. Two very different examples of this great diversity of animals are:

The **wallaby**, which has large, powerful hind legs, large jumping feet, a long, muscular tail for balance, and a small head. It is herbivorous, feeding on grass and roots on cold afternoons and nights, generally in groups. This species is nocturnal; it usually spends the day quietly inside its lair. They have a life expectancy of approximately 18 years.

The green-winged teal is an animal that measures about 35 cm in length. It has a flat, rounded bill that is brown or dark brown in color, sometimes mottled or with black spots. Its legs are grayish. It develops around areas with large bodies of water. The lagoons near the trees and with abundant aquatic vegetation are an example of a suitable habitat, although there are variants of this species that leave the lagoons and move to the large and blue ocean to look for food that floats on the surface of the water.

Math text

1. Una zarceta recorre 10 metros por minuto volando y 15 metros por minuto desplazándose por el agua, impulsándose con sus patas palmeadas. Si se ha estado moviendo 30 minutos por el agua y 40 minutos volando, ¿cuántos metros ha recorrido en total?
2. Una bióloga está buscando ejemplares de ofidio en la selva tropical de África, para investigar las propiedades de su veneno, y con ello poder elaborar antídotos y medicamentos. Después de varios días, ha atrapado 6 ejemplares de boa, 4 de anaconda y 3 de cobra. Si solo tiene jaulas para transportar 8 ejemplares y necesita el doble de boas que del resto ¿cuántos se podrá llevar de cada tipo?
3. Un agricultor tiene en una caja varios trozos de esqueje, que ha cortado de árboles y arbustos para plantarlos. Tiene para plantar 10 laureles, el triple de higueras y la mitad de plantas de romero. ¿Cuántos árboles y arbustos plantará en total?
4. Hoy han llegado al zoo tres ejemplares de **ualabí** desde Australia. Como se pasan el día dando saltos, consumen mucha agua. El primero consume diariamente 1/2 de lo que consume el segundo. El segundo, el doble que el tercero. ¿Cuál es la cantidad que consumen los dos primeros si sabemos que el tercero consume 10 litros?

Math text

1. A green-winged teal travels 10 meters per minute flying and 15 meters per minute moving through the water, propelling itself with its webbed feet. If it has been moving for 30 minutes through water and 40 minutes flying, how many meters has it traveled in total?

2. A biologist is looking for specimens of ophidian in the tropical forest of Africa to investigate the properties of its venom, in order to elaborate antidotes and medicines. After several days, she has caught 6 boas, 4 anacondas, and 3 cobras. She needs twice as many boas as the rest of the species. Because she only has cages to transport 8 specimens, how many of each type can she take?
3. A farmer has in a box several pieces of cuttings, which he has cut from trees and shrubs to plant them. He has to plant 10 laurels, three times as many fig trees, and half as many rosemary plants. How many trees and shrubs will he plant in all?
4. Today, three specimens of **wallaby** have arrived at the zoo from Australia. Because they jump all day, they consume a lot of water. The first consumes 1/2 of what the second consumes daily. The second consumes twice as much as the third. What is the amount consumed by the first two if we know that the third consumes 10 liters?

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17*(9), 814–823.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*(2), 97–123.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93*(3), 498–520.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122–159.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology, 1*(1), 18–46.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods, 44*(4), 991–997.

Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017.

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.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*(4), 671–681.

Corral, S., Ferrero, M., & Goikoetxea, E. (2009). LEXIN: A lexical database from Spanish kindergarten and first-grade readers. *Behavior Research Methods, 41*(4), 1009–1017.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*(4), 1246–1258.

Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning, 51*, 477–496.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15*(1), 1–16.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*(2), 95–112.

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63*(3), 259–273.

Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 385–402.

Johns, B. T., Dye, M., & Jones, M. N. (2016). The influence of contextual diversity on word learning. *Psychonomic Bulletin & Review, 23*(4), 1214–1220.

Jones, M. N., Dye, M., & Johns, B. T. (2017). Context as an organizing principle of the lexicon. In B. H. Ross (Ed.). *The psychology of learning and motivation* (Vol. 67, pp. 239–283). San Diego: Elsevier Academic Press.

Jones, M. N., Johns, B. T., & Recchia, G. (2012). The role of semantic diversity in lexical organization. *Canadian Journal of Experimental Psychology, 66*(2), 115–124.

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology, 166*, 190–211.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Marín, J., & Carrillo, M. S. (1999). *Test Colectivo de Eficacia Lectora (TECLE) [Collective Test of Reading Efficiency]*. Murcia, Spain: Universidad de Murcia.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*(2), 237–270.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233. https://doi.org/10.2307/747758.

Pagán, A., & Nation, K. (2019). Learning words via reading: Contextual diversity, spacing, and retrieval effects in adults. *Cognitive Science, 43*(1), e12705. https://doi.org/10.1111/cogs.12705.

Perea, M., Soares, A. P., & Comesaña, M. (2013). Contextual diversity is a main determinant of word identification times in young readers. *Journal of Experimental Child Psychology, 116*(1), 37–44.

Plummer, P., Perea, M., & Rayner, K. (2014). The influence of contextual diversity on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 275–283.

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

Ramos, J. L., & Cuetos, F. (1999). *PROLEC-SE: Evaluación de los procesos lectores en alumnos de tercer ciclo de educación primaria y secundaria*. Madrid, Spain: TEA Ediciones.

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86*(1), 54–64.

Rosa, E., Tapia, J. L., Perea, M., & Manalo, E. (2017). Contextual diversity facilitates learning new words in the classroom. *PLoS One, 12*(6), e0179004.

Tapia, J. L., Rosa, E., Rocabado, F., Vergara-Martínez, M., & Perea, M. (2021). Does narrator variability facilitate incidental word learning in the classroom? *Memory & Cognition*. https://doi.org/10.3758/s13421-021-01228-4.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373.