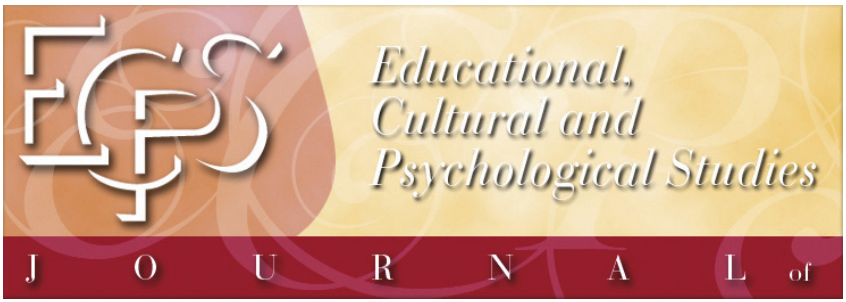

31
June 2025

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Note di Ricerca
Research Notes

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Improving Vocabulary Skills: What Strategies to Be Applied in Primary School? *

Marta De Angelis [1] **- Antonio Calvani** [2]

[1] *Università degli Studi del Molise - Department of Humanities, Education and Social Sciences (Campobasso, Italy)*

[2] *Società per l'Apprendimento e l'Istruzione Informati da Evidenza – S.Ap.I.E. (Italy)*

DOI: https://doi.org/10.7358/ecps-2025-031-anca

marta.deangelis@unimol.it
antonio@calvani.it

MIGLIORARE LE ABILITÀ LESSICALI: QUALI STRATEGIE APPLICARE NELLA SCUOLA PRIMARIA?

Abstract

Evidence-based research indicates that intentional vocabulary enhancement programmes are more effective than extemporaneous vocabulary teaching practices, with effect size values exceeding 0.60. However, it remains to determine which programme is most suitable for use, considering its efficacy, sustainability and adaptability to various contexts. In this study, we developed and tested a lexical enhancement programme that integrates students' weekly activity of independently collecting new terms with reflection activities on terms extracted from short texts during class lessons. The experiment conducted with a large sample of fourth and fifth primary school classes shows that the experimental group (EG) achieves a significantly greater improvement than the control group (CG), not only on the target words addressed in the texts but also in general lexical knowledge and, to some extent, in synthesis skills. The significance of such programmes is emphasised, as they can now be easily implemented and adapted to different contexts, aided by AI.

* Even though the authors have jointly conceived this paper, the elaboration of the paragraphs is attributed as follows: number 1 and number 2 to A. Calvani; numbers 3, 4 and 5 to M. De Angelis; the Conclusions to both authors.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Keywords: Lexical enhancement; Primary school; Quasi-experimental design; Vocabulary development; Vocabulary skills.

1. Introduction

The important actions that every school system should undertake to reduce the lexical gap between students are among the most relevant educational recommendations, particularly in a country like Italy, which has a significant history of illiteracy. Since the 1960s, this issue has been the subject of analysis and denunciation due to its profound socio-cultural implications (Milani, 1957; De Mauro, 1963; Barbagli, 1972). In subsequent years, attention to vocabulary has been developed in the field of language teaching (Casadei & Basile, 2019), although this analysis has not been accompanied by specific experimental interventions.

Currently, the National Guidelines for the Curriculum of preschool and the first cycle of education state that «[…] pupils must expand their oral vocabulary and learn to read and write correctly, with an increasing enrichment of lexicon, […] a task that is all the more important, the more there is evidence today of a progressive impoverishment» (MIUR, 2012, pp. 29-30). The INVALSI Italian Framework underscores the importance of both the quantitative and qualitative dimensions of vocabulary, the latter being understood as the «ability to recognise, implicitly or explicitly, the relationships of meaning between words (synonymy, opposition, inclusion), their morphological characterisation and possibly their etymology» (INVALSI, 2018, p. 5).

At an international level, the specific problem of vocabulary has been the subject of extensive literature from various perspectives: the age of learners and the optimal conditions for lexical development; the study of the relationship between vocabulary and text comprehension; teaching methods and their effectiveness.

Regarding contextual conditions, the significance of linguistic interactions in early childhood has long been emphasised. This aspect, which has prompted numerous early intervention programmes since the 1960s (Camaioni, 1983), is today corroborated by recent contributions from neuroscience on the role of dialogic narration between adults and children (Wolf, 2009; Panza, 2015).

Concerning the relationship between vocabulary and text comprehension, vocabulary is not the sole factor influencing text understanding; other factors such as the ability to make lexical and semantic inferences as well as monitor one's comprehension processes and knowledge related to

the structure of the text are also crucial (Oakhill & Cain, 2007; Montesano, 2020). Nevertheless, lexical knowledge remains a strong predictor of text comprehension (Laufer, 1989).

There are numerous teaching methods which may be difficult to classify due to their variations and integrations (NICHD, 2000; Fisher *et al.*, 2016). The National Institutes of Children's Health and Development (2000) identifies over 20 available methods, including keyword research, associations, semantic maps, context analysis, root/affix analysis, verbal pre-instructions, the use of flashcards and dictionaries/glossaries, extensive reading and listening/reading stories, with or without prior explanation of the target words.

These methods have evolved into various forms with improvement in the Italian context, developed by creative writers and teachers. Notably, the significant contributions of Gianni Rodari and his *Grammatica della fantasia* (1973) and, more recently, those of Nicola Zuccherini (2018, 2020), who integrates different teaching strategies such as semantic organisations, derivations through suffixes and prefixes, idioms and usage.

Regarding the effectiveness of these methods, numerous systematic interventions have demonstrated a high impact on student learning, with effect size values greater than 0.60 (Marulis & Neuman, 2010; Hattie, 2012, 2023). However, a recent umbrella review by Pannone and Pellegrini (2023) confirms these findings while also highlighting the significant differences in results when measured with tools designed by researchers compared to standardised external tools. Strikingly, the results are considerably lower in the latter case; thus, while it is relatively easy to achieve effects in semantic areas closely related to application, obtaining them in the broader universe of general vocabulary appropriate for the age group in question is much more challenging.

To briefly summarise what works in lexical learning within evidence-based education (Fisher *et al.*, 2017; Bilton & Duff, 2021), we can highlight the following aspects:

- The role of explicit vocabulary teaching, either alone or in combination with implicit vocabulary teaching strategies, which should be predominant from primary school onwards.
- Actively engaging students in word searches and analyses, ensuring that vocabulary development results from meaningful and intentional learning.
- Tracking learned words; a new term must be contextualised within pre-existing lexical and semantic structures. It is beneficial for students to use a specific workbook as a «treasure chest» or «bank» of words, allowing easy retrieval when necessary.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

- The integration of passive lexical competence concerns the ability to understand words, and active lexical competence and involves using the lexical corpus appropriately and consciously through personal written production.
- Linguistic processing and organisation of the lexicon (differentiations, synonymies, antonyms and ordering in semantic fields).
- The importance of the dialogic aspect requires that children be given opportunities to learn through discussion, comparison and sharing with others.
- A metacognitive approach involves providing moments of reflection on the lexicon and the mechanisms that regulate its use, enabling the student to become a «strategic reader».

2. The vocabulary enhancement programme «PL – dal Testo alla Parola» (PL-TP)

This study aimed to identify a systematic programme of lexical enhancement that is not only effective but also extremely easy to implement and highly adaptable to various contexts.

The research methodology employs the Evidence-Based Improvement Design (EBID) model (Calvani & Marzano, 2020), which aims to implement experimental pathways based on the already acquired evidence. The selected programme is then applied to a large sample of teachers who must adopt the shared protocol without modification; only after the experimentation and evaluation of its effectiveness are the participating teachers involved in improvement proposals.

The teaching methodology integrates two approaches: the children's weekly autonomous search for unknown words, culminating in the construction of a «word chest» notebook and a structured procedure for classroom lessons. Among the various possible choices, a top-down approach was favoured, which begins with reading a text and subsequently performing specific linguistic elaborations on the identified terms.

The theoretical and applicative components of the programme are based on the scientific evidence of the Reciprocal Teaching (RT) model, which encompasses fundamental strategies (questioning, searching for key information; summarising, synthesising the passage in a few words; anticipating, predicting what might happen in the story; clarifying, understanding the meaning of unknown terms). This model has already been tested with excellent results in Italy (Calvani & Chiappetta Cajola, 2019; Calvani

et al., 2023). The programme, named *Potenziamento Lessicale – dal Testo alla Parola* (PL-TP), utilises shorter texts with categories of words belonging to the same semantic scope, reducing the strategies of predicting and questioning to allow more focus on summarising and clarifying.

Starting from the target vocabulary in the texts, students are encouraged to use some of these terms to construct additional sentences and establish elementary semantic relationships between them (synonymy, antonymy and hypernymy). These semantic relationships were chosen for their relevance and cognitive implications: the hypothesis is that «[…] between the semantic relations of synonymy, opposition, hyponymy/hypernymy, etc., and certain cognitive modalities […] there exists a close link, in the sense that just as words are related to each other in various ways within the mental lexicon, it is reasonable to assume that conceptual structures are connected in a similar manner» (Casadei & Basile, 2019, p. 45).

The programme aimed at students in the fourth and fifth grades of primary school, employed 18 short texts of approximately 400 characters each, specifically created by the researchers and characterised by simple syntax and familiar content for the children, allowing them to focus on the words and their semantic differentiations. Each text contains a list of target terms belonging to the same semantic area: the decision to group target words into semantic fields is based on the understanding that while the semantic system develops independently of the lexical label, the ability to learn new words is facilitated by categorising them within a common semantic category (Goodrich & Lonigan, 2017). Each text contains eight target terms for class IV and nine for class V. The terms were selected using the *Nuovo vocabolario di base della lingua italiana* by Tullio De Mauro (2016), employing a repertoire of basic, high-use and high-availability terms, while avoiding specialist terminology.

Table 1 shows an example of a text from the first unit, «Emotions» (fifth-grade repertoire), while *Table 2* provides a descriptive summary of the teaching procedure.

Table 1. – Example from the fifth-class repertoire (target words in bold).

Today is Greta's birthday. She wakes up **excited** because she is sure to find a surprise from her parents. **Impatient**, she gets out of bed but does not see any present in her room. What a **disappointment**! She goes to the kitchen **perplexed** and **melancholic**. But suddenly, the sight of a box on the table interrupts her **anguish** and fills her with **enthusiasm**. There, in fact, is the doll she had dreamt of. **Satisfied** and **comforted**, she goes to hug her parents.

Between lessons, students are tasked with collecting new words according to a pre-arranged grid called «The Word Detective» (*My new word is …*

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Where did I find it … What should it mean … What does it go with … Definition … I write a sentence with the word …). Each week, a few children share and present the new word they have learnt to the entire class.

Table 2. – Description of a typical teaching procedure of a work unit.

The procedure of the lesson / work unit

Lesson Start
At the beginning of each unit, once a week, a maximum of 2-3 students are listened to in rotation. They illustrate a new word found during their weekly independent research. *Did you also know the new word discovered by Carlo? Could you suggest another sentence using this word?*

Text Dictation
The teacher first dictates the text of the work unit planned for the day. Each time a target word is read, the teacher pauses and asks the children to underline it, as these words will be discussed later.

Modelling: Summarising
The teacher thinks aloud, posing key questions from RT: *Who is the protagonist of the story? How does she feel at the beginning? What happens when she doesn't find her gift? What do you think will happen next? What is the gist of the text? Could you summarise it in a few words? What title would you give to this story?*

Reflection on Target Terms
The teacher asks questions to uncover the common characteristics of some of the target words. *Let's examine these words [perplexed - impatient - disappointment - excited]. What do they mean? What do they have in common? Can we categorise them under a single term? What does perplexed mean, to be perplexed … What could I substitute for perplexed? What could be the opposite of this word? To determine if this word has a similar meaning to «perplexed», I can try substituting it in a sentence and see if the meaning remains intact … For example, Greta is doubtful … would that still work in your opinion?*

Individual Work on Target Terms
Now, children, choose another word from the underlined ones and write a sentence using it. Try replacing it with another from the list, as I did. How would the meaning change? Once you have finished, compare your sentence with that of your partner.

Metacognitive Work in Pairs
Now, working in pairs, find a synonym and an antonym for the two chosen words … Is there also a word that encompasses them?

Final Feedback from the Teacher
When the pairs have completed their work, the teacher asks some of them to share their findings, providing feedback to the class. It is important that during this phase, the children explain their choices.
The responses of some pairs are discussed and compared in class. *So, children, let's see which words you have chosen … Did you agree …? Who selected different words? Which word can encompass them all?*

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

3. Methods

3.1. *Hypothesis and design*

The research design is quasi-experimental, featuring a non-random experimental and control group. The plan included a pre-test for both the experimental and control groups regarding the children's lexical abilities; the application of the PL-TP programme exclusively in the experimental groups; a new evaluation (post-test) for both groups at the experimentation and evaluation of the training intervention.

The phases of the intervention were as follows:

1. Entry questionnaire for teachers to collect data related to the participating classes (personal data, learning levels, presence of students with special needs and vocabulary teaching experiences adopted in class by teachers) (beginning of September 2023).
2. Training intervention for teachers of experimental groups to learn how to work with students through synchronous meetings and asynchronous demonstration materials on an e-learning platform (approximately 15 hours, mid-September 2023).
3. Administration of pre-tests to the experimental and control groups (end of September 2023).
4. Application of the experimental programme in the experimental groups twice a week, for a total of 18 sessions (October / November / early December 2023).
5. Administration of post-tests to the experimental and control groups (mid-December 2023).
6. Exit questionnaire for teachers of the experimental groups only (December 2023).

The intervention aims to verify the following hypotheses regarding the potential impact of the vocabulary enhancement programme:

- H1. Improvement in the acquisition of specific target words contained in the proposed texts (specific vocabulary).
- H2. Improvement in the acquisition of words not contained in the proposed texts (general vocabulary).
- H3. Improvement in the processes implemented when elaborating the meaning of words.
- H4. Improvement in the children's ability to understand and summarise texts not included in the programme.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

3.2. *Measures*

In order to verify the effectiveness of the programme, the following tests were administered at the beginning and end of the classes involved (experimental and control):

1. *Summarising Test* (ST). This test consists of five short texts accompanied by three multiple-choice questions, each with four answer alternatives. The questions assess the ability to identify the most important information in a text, find the most suitable title and choose the words that best express the meaning in a text. There are two versions for the age group in question, both at entry and exit (Calvani & Menichetti, 2019). It is used to verify H4.
2. *Multidimensional Vocabulary Test* (MVT) (Aprile, 2012). This test consists of a short text accompanied by 20 multiple-choice items with four answer alternatives. It evaluates various lexical processes (synonymic, antonymic, categorical, functional, decontextualisation and contextualisation). The vocabulary of the test was not included in the experimental programme. There are two different tests for classes IV and V (one at the entrance and one at the exit). It is used to verify H3.
3. *General Lexical Test* (GLT). This is a vocabulary test with closed answers, in which children must identify the synonym from a set of four alternatives. It is a revision/expansion of the Verbal Meaning Test n.v., which was found to be generally easy (Montesano, 2020). The terms used are different from the target terms. The same version is used for IV and V at both entry and exit. It is used to verify H2 [2].
4. *Specific Lexical Test* (SLT). This is a vocabulary test designed ad hoc in which children are asked to identify synonyms (40 items), intruders (20 items) and hypernyms (10 items) from a set of four alternatives. The terms are selected from the target terms of the experimental programme, in the two versions for IV and V. It is applied in the same format at both input and output. It is used to verify H1.

Additionally, a semi-structured knowledge questionnaire was administered to the participating teachers in the initial phase to gather useful information regarding the class and the methodologies used for lexical learning.

An additional questionnaire was administered to the teachers of the experimental groups only at the experimentation and evaluation stage of

[2] Both the SLT and GLT tests are available on the website https://www.lessico.sapie.it. Some statistical analyses have been conducted to evaluate their psychometric characteristics starting from the data available as input.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

the activities, asking them to evaluate the overall experience and provide suggestions for improvement.

The teachers of the experimental groups also maintained a logbook in which they recorded the details of the interventions (date of the intervention, passage worked on in class, any identified critical issues/problems as well as the number of children absent for each work unit).

3.3. *Participants*

The programme was tested in 10 schools in the Campania region, involving 38 fourth-grade classes (EG = 16; CG = 22) and 37 fifth-grade classes (EG = 20; CG = 17). The specific characteristics of the sample examined at the beginning of the test for both fourth and fifth-grade classes are reported in *Tables 3* and *4*.

Table 3. – Composition of the sample related to the fourth grade.

Characteristics	Experimental	Control
N	307	335
Age (mean)	8.65	8.70
Male	158 (51.47%)	163 (48.66%)
Female	149 (48.53%)	172 (51.34%)
Disabilities	21 (6.84%)	21 (6.84%)
Specific learning disorders	11 (3.58%)	7 (2.09%)
Special educational needs [3]	37 (12.05%)	13 (3.88%)
Normal time [4]	261 (85.02%)	273 (81.49%)
Full time [5]	46 (14.98%)	62 (18.51%)

[3] Children with socioeconomic, linguistic or cultural disadvantages have been classified in the SEN category.

[4] Normal time means a time spent in school of 24, 37 or 30 hours per week.

[5] Full-time means attending school for 40 hours per week.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table 4. – Composition of the sample related to the fifth grade.

Characteristics	Experimental	Control
N	385	306
Age (mean)	9.68	9.79
Male	189 (49.09%)	137 (44.77%)
Female	197 (51.17%)	169 (55.23%)
Disabilities	12 (3.12%)	18 (5.88%)
Specific learning disorders	13 (3.38%)	6 (1.96%)
Special educational needs	28 (7.27%)	24 (7.84%)
Normal time	306 (79.48%)	265 (86.60%)
Full time	79 (20.52%)	41 (13.40%)

The teachers involved in the research for the fourth-grade classes included 11 in the experimental groups (*M* years of teaching = 30.3; *SD* = 8.34) and 18 in the control groups (*M* years of teaching = 26.6; *SD* = 12.9); for the fifth-grade classes, the teachers of the experimental groups numbered 14 (*M* years of teaching = 26.6; *SD* = 12.2), while the control groups comprised 12 teachers (*M* years of teaching = 27.5; *SD* = 8.0).

4. Data analysis

In line with the quasi-experimental pre-post test design, baseline equivalence between the experimental and control groups was first assessed using Hedges' *g*. Intervention effects were then evaluated through ANCOVA on the four literacy measures (ST, MVT, GLT, and SLT), with centered pre-test scores entered as covariates to control for initial differences, using the statistical software *Jamovi* (version 2.6.2). These analyses are reported in each results subsection in accordance with the study's hypotheses and outcome measures.

4.1. *Baseline equivalence*

Initially, the baseline equivalence of the experimental and control conditions was analysed, as required by the standards of the What Works Clearinghouse (2022). For each test, the Hedges' *g* value was calculated, verifying that it was at least ≤ 0.25. This analysis revealed no significant differences between the two groups in the pre-tests for both classes (*Tables 5* and *6*); thus, they were considered equivalent.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table 5. – Results obtained at pre-test (fourth grade).

Test*	M (SD)		Baseline ES
	EG	CG	*g*
ST	15.6 (5.26)	15.4 (5.57)	0.03
MVT	12.4 (3.38)	11.9 (3.73)	0.13
GLT	21.2 (6.62)	19.9 (6.95)	0.19
SLT	44.5 (18.2)	44.3 (16.7)	0.01

Table 6. – Results obtained at pre-test (fifth grade).

Test	M (SD)		Baseline ES
	EG	CG	*g*
ST	18.4 (5.58)	18.1 (5.84)	0.05
MVT	11.8 (3.56)	11.3 (3.22)	0.14
GLT	25.4 (6.58)	26.0 (6.98)	0.08
SLT	51.9 (15.4)	52.0 (15.9)	0.00

Note: * Legend of the test acronyms: Summarizing Test (ST), Multidimensional Vocabulary Test (MVT), General Lexical Test (GLT) and Specific Lexical Test (SLT).

4.2. *Results at post-test*

Table 7 presents the descriptive and inferential statistics for each outcome in both fourth and fifth grade. For each test, unadjusted post-test means with standard deviations (SD) are reported alongside adjusted post-test means with standard errors (SE), derived through ANCOVA controlling for pre-test performance. To offer a comprehensive understanding of effect magnitude, multiple standardized indicators were reported. Cohen's *d* was used as a conventional effect size metric, while Morris's d_{ppc2} was employed to account for pre-test variability, providing a more accurate estimate of the intervention's impact.

Controlling for pre-test scores, ANCOVA results revealed that, in fourth grade, the intervention led to significant improvements in three out of four domains. The SLT produced the strongest effect [$F(1, 501) = 48.8$, $p < .001$; $d_{ppc2} = 0.36$], providing robust support for hypothesis 1 on the acquisition of specific lexical knowledge. The GLT also showed a significant effect [$F(1, 482) = 16.8$, $p < .001$; $d_{ppc2} = 0.22$], aligning with hypothesis 2 regarding generalisation to untaught vocabulary. A smaller but statistically significant gain was observed in the ST [$F(1, 494) = 7.39$, $p < 0.05$; $d_{ppc2} = 0.20$], consistent with hypothesis 4 on improved summarising ability. No significant effect was found for the MVT [$F(1, 481) = 0.33$, $p = .568$], offering no support for hypothesis 3.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Table 7. – Means (SD), Adjusted Means (SE), and Effect Sizes by group, test, and grade level.

Grade	Test	Group (N) [6]	Unadjusted Mean (SD)	Adjusted Mean (SE)	*d*	d_{ppc2}
4th	ST	EG (247)	19.5 (5.73)	19.4 (0.33)	0.23	0.22
		CG (250)	18.1 (6.23)	18.2 (0.32)		
	MVT	EG (222)	11.9 (2.98)	11.9 (0.20)	0.04	-0.08
		CG (262)	11.7 (3.19)	11.7 (0.19)		
	GLT	EG (243)	25.3 (6.64)	24.9 (0.34)	0.41	0.22
		CG (242)	22.5 (6.95)	22.9 (0.34)		
	SLT	EG (239)	57.3 (13.8)	57.3 (0.67)	0.44	0.36
		CG (265)	50.8 (15.8)	50.8 (0.63)		
5th	ST	EG (317)	22.0 (6.05)	21.9 (0.27)	0.11	0.07
		CG (276)	21.3 (5.75)	21.4 (0.29)		
	MVT	EG (319)	11.2 (3.38)	11.1 (0.17)	0.28	0.12
		CG (282)	10.3 (3.30)	10.4 (0.18)		
	GLT	EG (321)	27.9 (6.48)	28.1 (0.26)	0.04	0.13
		CG (283)	27.6 (7.08)	27.3 (0.27)		
	SLT	EG (320)	61.3 (13.2)	61.3 (0.56)	0.28	0.26
		CG (289)	57.3 (14.9)	57.2 (0.59)		

In fifth grade, the only clearly significant effect emerged in the SLT [$F(1, 606) = 24.4$, $p < .001$; $d_{ppc2} = 0.26$], again confirming hypothesis 1. In line with hypothesis 2, the GLT reached statistical significance [$F(1, 601) = 4.16$, $p < .05$], although the effect sizes were minimal ($d_{ppc2} = 0.13$), suggesting limited educational relevance. While the MVT showed a modest effect ($d_{ppc2} = 0.12$), raw scores declined in both groups, indicating a possible protective function rather than actual improvement. No significant differences emerged for the ST [$F(1, 590) = 1.76$, $p = .186$], providing no support for hypothesis 4 in this grade.

Across all measures – except for the MVT in fourth grade – pre-test scores were strong predictors of post-test performance, highlighting the importance of controlling for baseline ability in assessing intervention effects.

[6] The sample shows a slight reduction compared to the initial number. Data relating to students with certified disabilities and those with linguistic disadvantages requiring customised tests and participation in the programme with simplified activities in addition to those who did not complete both the pre and post-test were excluded from the overall analysis of the results.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

4.2.1. Teachers' satisfaction

As previously mentioned, in the evaluation phase of the experimentation, the teachers of the experimental classes provided an overall evaluation of the programme, suggesting improvements for its optimisation. Twenty-three teachers from the experimental classes responded to the questionnaire: 87% expressed a desire to apply the experimented method in their future teaching activities. The programme's perceived effectiveness was rated highly, with 78.2% of responding teachers evaluating it at the highest levels (4-5) on a scale from 1 to 5 (1 = not at all effective; 5 = very effective).

Teachers were also asked to report on the children's appreciation of the proposed activities: the activities were found to be highly appreciated in almost all cases (91.3%), confirming the programme's applicability in the classroom.

Among the factors that could hinder the full implementation of the programme, the most significant were the overload of school commitments (43.5%) and potential interference with other school programmes and activities (39.1%). Teachers' suggestions for future improvements primarily concerned the timing of the programme (69.6%) and the duration of work on each individual text (13%). The analysis of open-ended responses revealed a desire for more relaxed times on the application of the programme: «The time to dedicate to this program should be longer» (G:2); «More time should be given to the project because each text should be assimilated better […]» (G:10); «Considering possible unforeseen events, it would be appropriate to have more time» (G:11); «Each text could have a greater in-depth study, but this would also require longer times which are often difficult to find due to too many school commitments» (G:14).

5. Discussion

Although lasting only 18 biweekly sessions over approximately three months, the program yielded significant advantages for the EG compared to the CG across three of the four tests used, albeit with varying impact between the fourth and fifth-grade classes. While the sample size, statistical significance, and effect sizes suggest that the findings are promising, limitations related to the sample size and quasi-experimental design should be considered when interpreting the results. Thus, it can be concluded that PL-TP outperforms the outcomes generally achieved through common

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

teaching practices implemented in classes of the same level. This further confirms that a systematic programme of lexical enhancement is preferable to practices based on extemporaneous teaching, even if the effect sizes obtained in this study are lower than those reported in Hattie's meta-analyses (2012, 2023).

It was anticipated that a greater impact on the learning of target words would be observed (Cohen's *d* fourth grade = 0.44; Cohen's *d* fifth grade = 0.28); however, the impact on the summarising tests was less predictable (given that the reference programme on RT had a duration twice as long). This effect was even less so for the general vocabulary in the fourth grades, an aspect that, as previously mentioned, occurs less frequently in the international literature. On the contrary, no differences were found using the MVT test, which evaluated some lexical processes (explanation of meaning, synonyms, antonyms, hypernyms and functional uses).

It is also essential to acknowledge the limitations of this investigation. The use of two different strategies leaves us uncertain about the extent to which each contributed to the outcomes; in particular, the impact of the vocabulary extension carried out through the autonomous collections remains unclear.

Moreover, the freedom of time left in the various phases may have hindered a more precise evaluation of the time dedicated to processing the terms, which may have been compromised in some instances. Furthermore, we cannot ascertain whether this may have influenced the lack of improvement in the MVT test. It remains to be verified in future experimental interventions whether and to what extent the semantic association activities employed can translate into lasting linguistic competence. In addition, further considerations can be made. It is possible to know a word at different levels of complexity; the more complex aspects reported by lexical semantics (Casadei & Basile, 2019) and the ability to integrate the term into its social and communicative context, articulating it with precision and flexibility (Perfetti, 2007), was not included in the research scope. Furthermore, the survey does not account for long-term effects. The assessment was conducted only during the experimentation and evaluation of the experiment, providing no insight into whether the acquisitions are retained or whether the programme can promote the autonomous development of further enrichment. Despite these limitations, the experimental model has proven effective, allowing a clear distinction compared to classes of the same level that followed the standard curriculum. Its simplicity favours the sustainability and further dissemination of the approach, as confirmed by the appreciation of the teachers (and indirectly the students) in the experimental groups.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Beyond the general linguistic programme tested here, an intriguing area for extension emerges for interventions focused on lexical anticipation with specific terms in particular disciplinary fields (for example, in lessons or teaching units of science, history, etc.).

For its potential extension, the greatest challenge lies in the time required to prepare texts suited to the specificity of the learning contexts (content, age of the students and regulation of linguistic complexity), an aspect that becomes even more crucial if one wishes to consider the adaptations needed to address students' specific difficulties.

While the material used here to create suitable texts required considerable time investment from the authors who designed them, advances in artificial intelligence (e.g., tools such as ChatGPT and other large language models) now offer new possibilities for generating programmes of this nature. These improvements provide appropriate indications regarding content and specify semantic areas and desired levels of linguistic complexity. Although the results obtained may not be optimal, they will certainly assist in the preliminary phase of preparing a substantial number of drafts of texts and usable terms.

6. Conclusion

Every teacher, at any school level, understands the importance of enhancing their students' vocabulary and typically endeavours to do so by focusing on any challenging terms encountered during classwork. However, beyond this ordinary practice, it is evident that systematic teaching strategies can be implemented to promote lexical enrichment.

A significant area for intervention is the final years of primary school when students can employ intentional strategies to enrich their lexical repertoire through readings integrated with linguistic production and metalinguistic reflections.

In the absence of experiments with systematic interventions, we proposed the PL-TP programme which integrates autonomous term searches with lessons based on text comprehension (according to the RT model), followed by active elaborations and reflections on elementary semantic relations, both individually and in pairs. The results of the programme, applied over 18 sessions to a large sample of fourth and fifth-grade students, demonstrated a significant advantage compared to the control group in terms of lexical enrichment, the ability to understand and summarise a text, and, to some extent, in terms of general external vocabulary.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

Although this does not imply that there are no other approaches that can yield positive results, we can assert that the programme, if reapplied and adapted, for instance, to specific disciplinary areas, has a high probability of facilitating a notable advancement compared to the common practices currently employed in schools. The major criticality remains in the time required for its preparation, but in this regard, AI offers valuable support.

Acknowledgements

We sincerely thank Antonio Marzano for his great support in the recruitment phase of the participating schools and the students in Primary Education Sciences: Adua Galise, Olimpia Meo, Assunta Nunziata, Jennifer Palo, Teresa Pellegrino, Tea Melania Pesa, Viviana Retta, Rowena Rovere, Ilaria Schettino and Enza Scognamiglio for their collaboration and support to the teachers during the experimental phase. We also extend our gratitude to Teresa Serafini for her valuable suggestions during the programme development phase.

References

Aprile, L. (2012). *Sviluppo del vocabolario. Teorie, ricerche, strumenti operativi*. Torino: Utet Università.

Barbagli, M. (1972). *Scuola, potere e ideologia*. Bologna: il Mulino.

Bilton, C., & Duff, A. (2021). *Improving literacy in key stage 2: Guidance report*. London: Education Endowment Foundation.

Calvani, A., & Marzano, A. (2020). Progettare per un miglioramento basato su evidenze. Quale metodologia? *Giornale Italiano della Ricerca Educativa*, *13*(24), 67-83.

Calvani, A., Marzano, A., Montesano, L., Pellegrini, M., Rizzo, A.L., Traversetti, M., & Vivanet, G. (2023). Improving reading comprehension and summarising skills in primary school: A quasi-experimental study. *Journal of Educational, Cultural and Psychological Studies* (*ECPS*), *28*, 81-100.

Calvani, A., & Menichetti, L. (2019). La prova quantitativa di sintesi. Il Summarizing Test (ST). In A. Calvani & L. Chiappetta Cajola (a cura di), *Strategie efficaci per la comprensione del testo. Il Reciprocal Teaching* (pp. 411-430). Firenze: S.Ap.I.E. Scientifica.

Camaioni, L. (1983). Lo sviluppo del linguaggio in una prospettiva internazionale. In L. Tornatore, L. Camaioni, V. Volterra, & S. Mantovani (a cura di), *Imparare a parlare* (pp. 68-109). Torino: Loescher.

Casadei, F., & Basile, G. (2019). *Lessico ed educazione linguistica*. Roma: Carocci.

Clearinghouse W. W. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0* (National Center for Education Evaluation and Regional Assistance, NCEE). US Department of Education, Institute of Education Sciences.
https://ies.ed.gov/ncee/wwc/Handbooks

De Mauro, T. (1963). *Storia linguistica dell'Italia unita*. Bari: Laterza.

De Mauro, T. (2016). *Il nuovo vocabolario di base della lingua italiana – Internazionale*.
https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana

Fisher, D., Frey, N., & Hattie, J. (2016). *Visible learning for literacy, grades K-12: Implementing the practices that work best to accelerate student learning*. Thousand Oaks, CA: Corwin Press.

Fisher, D., Frey, N., & Hattie, J. (2017). *Teaching literacy improvement in the visible learning classroom, grades K-5*. Thousand Oaks, CA: Corwin Press.

Goodrich, J.M., & Lonigan, C.J. (2017). Language-independent and language-specific aspects of early literacy: An evaluation of the common underlying proficiency model. *Journal of Educational Psychology*, *109*(6), 782.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.

Hattie, J. (2023). *Visible learning: The sequel. A synthesis of over 2,100 meta-analyses relating to achievement*. New York: Routledge.

INVALSI (2018). *Quadro di riferimento delle prove INVALSI di Italiano*.
https://invalsi-areaprove.cineca.it/docs/file/QdR_ITALIANO.pdf

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.

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.

Milani, L. (1957). *Esperienze pastorali*. Firenze: Libreria Editrice Fiorentina.

MIUR (2012). *Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione del 2012*. Firenze: Le Monnier.
https://www.miur.gov.it/documents/20182/51310/DM+254_2012.pdf

Montesano, L. (2020). *Vocabolario e comprensione del testo. Uno strumento per la valutazione del lessico nella scuola primaria*. Firenze: S.Ap.I.E. Scientifica.

Online ISSN 2037-7924 - Print ISSN 2037-7932 - ISBN 978-88-5513-208-4

NICHD – National Institutes of Children's Health and Development (2000). *Report of the national reading panel: Teaching students to read. An evidenced-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups*. Bethesda, MD: National Institutes of Health.

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Pannone, C., & Pellegrini, M. (2023). Cosa sappiamo sull'efficacia dei programmi di potenziamento lessicale. Una sintesi delle evidenze. *Formazione & insegnamento*, *21*(2), 125-135.

Panza, C. (2015). Nati per leggere e lettura dialogica: a chi e come. *Quaderni ACP*, *2*, 95-101.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383.

Rodari, G. (1973). *Grammatica della fantasia*. Torino: Einaudi.

Wolf, M. (2009). *Proust e il calamaro. Storia e scienza del cervello che legge*. Milano: Vita e Pensiero.

Zuccherini, N. (2018). L'ora di lessico. Dall'indagine sulle competenze degli studenti al curricolo delle parole. *Italiano LinguaDue*, *10*(2), 167-184.

Zuccherini, N. (2020). *Quante parole conosci? Attività e test per imparare il lessico italiano*. Bologna: La Linea.

Riassunto

La ricerca evidence-based ci mostra che programmi di potenziamento intenzionale del vocabolario sono più efficaci di pratiche estemporanee di insegnamento del vocabolario, con valori di effect size superiori a 0.60. Rimane tuttavia da scegliere quale programma sia più conveniente utilizzare, considerando accanto alla sua efficienza la valutazione della sua sostenibilità e adattabilità ad altri contesti. In questo lavoro abbiamo allestito e sperimentato un programma di potenziamento lessicale che integra un'attività settimanale di raccolta di termini nuovi compiuta autonomamente dagli alunni, con attività di riflessione compiute su termini ricavati da brevi testi durante lezioni in classe. La sperimentazione, compiuta su un ampio campione di classi di IV e V di scuola primaria, mostra che il gruppo sperimentale (GS) consegue un miglioramento significativamente più elevato rispetto a quello di controllo (GC) non solo sulle parole target trattate nei testi ma anche su termini esterni attinti dal vocabolario generale e, in parte, sulla capacità di sintesi. Si sottolinea l'importanza di programmi di questo tipo, che possono essere oggi facilmente implementati e adattati a diversi contesti, anche con l'aiuto della I.A.

Parole chiave: Abilità lessicali; Disegno quasi-sperimentale; Potenziamento lessicale; Scuola primaria; Sviluppo del vocabolario.

How to cite this paper: De Angelis, M., & Calvani, A. (2025). Improving vocabulary skills: What strategies to be applied in primary school? [Migliorare le abilità lessicali: quali strategie applicare nella scuola primaria?]. *Journal of Educational, Cultural and Psychological Studies (ECPS)*, *31*, 51-69. https://doi.org/10.7358/ecps-2025-031-anca