



Università
Ca' Foscari
Venezia

Corso di Dottorato di ricerca
in Scienze del Linguaggio
ciclo XXXIII

Tesi di Ricerca

**Development of preliteracy and literacy in monolingual and
mono-literate bilingual children in an Italian school setting.**

A longitudinal study with metaphonological, RAN, and graphic intervention

SSD: L-LIN/01

Coordinatore del Dottorato
ch. prof. Enric Bou Maqueda

Supervisore
ch. prof. Anna Cardinaletti

Co-Supervisore
ch. prof. Giulia Bencini

Dottorando
Emanuele Casani
Matricola: 861105

Emanuele Casani

**Development of preliteracy and literacy
in monolingual and mono-literate bilingual children
in an Italian school setting.**

A longitudinal study with metaphonological, RAN, and graphic intervention

Ph.D. Thesis

2021

ABSTRACT

For the last fifty years, the study of reading processes, in particular in subjects with acquired (e.g., Denes, Cipollotti, & Zorzi, 1996) and developmental deficits (e.g., Ellis, 1984), has received increasing attention by the cognitive and neuropsychological literature, which led to identifying different models of reading and writing development (Frith, 1985; Coltheart et al., 2001). On the other hand, particular attention has been devoted to some abilities that have been taken as essential for literacy development, such as phonological awareness (Liberman, 1971) (or metaphonology), and Rapid Automatized Naming (RAN) (Denckla & Rudel, 1976). Although the role of these abilities is widely shared among scholars, there is a shortage of longitudinal studies on metaphonology development (Dehaene, 2009), and on the relation between metaphonology and writing skills (Zoccolotti et al., 2005). This is even more true when we take into consideration languages other than English, such as Italian.

The present research longitudinally analyzed the development of preliteracy and literacy in a sample of Italian monolingual children as compared to a sample of mono-literate bilingual children with minority languages attending the same school. The aims were to investigate how diverse preliteracy skills contribute to the development of the first steps of literacy in a language with shallow orthography like Italian; whether this development can be enhanced through specific training of some preliteracy skills, namely phonological awareness, RAN, and graphic abilities; and whether such training has any effects in case of atypical development. Children were monitored from the last year of kindergarten to the 3rd grade of primary school through four studies.

Study 1 allowed to describe the typical development of language, metalanguage, RAN, and short-term memory prerequisites of literacy, and the first steps of acquisition of reading and writing in monolingual and bilingual children. We found similar phonological skills, but a disadvantage in the other abilities of bilinguals, including RAN and verbal working memory. An area of particular difficulty, not only for bilinguals, was the syntactic production of direct object clitic pronouns, which revealed to be particularly sensitive to the elicitation conditions. A test was implemented (Casani, Cardinaletti, & Volpato, in preparation) with balanced morphosyntactic elicitation conditions.

Study 2 investigated how the above-mentioned abilities intervene in the development of reading and writing in the same sample of typically developing children. Reading and writing revealed their multicomponential nature, in which different abilities contribute to different extents in different developmental times. RAN was assumed as a measure of phonological automaticity, intervening in every index of reading, including accuracy and comprehension, and writing. Different measures of RAN and phonological awareness predicted different portions of literacy.

Study 3 showed the effects of three different training interventions (metaphonological, RAN, and graphic) administered in two different cycles (in the last year of kindergarten and 1st grade of primary school). The META group showed positive effects on several metaphonological measures. The RAN group revealed a positive effect on syllable blending, thus showing a relationship of RAN with phonology. The META group and, unexpectedly, the GRAPH group, improved significantly on several measures of literacy, thus showing that training phonological awareness and visual skills has positive effects on reading and writing.

Study 4 analyzed a small group of children who developed some literacy difficulties during the course of the project. The individual, social, linguistic, metalinguistic, and memory predictors of the deficits were explored, showing significant effects of social deprivation (with and without language disorder), working memory, vocabulary, metaphonological skills, and RAN on different deficits.

A single training administered in the 1st grade to children reported by teachers for difficulties in the acquisition of literacy did not show any significant effect, as the group developed different literacy deficits in the 3rd grade.

Children who attended the double metaphonological training as in Study 3 had positive effects on several measures of literacy. Children who attended the RAN training showed positive effects on non-word writing, which were attributed to the acquisition of automaticity in the phonological route of writing.

We conclude that interventions tapping preliteracy skills should be planned timely, starting from kindergarten, and reinforced in 1st grade, during early literacy acquisition. Such interventions should be multicomponential, to face the different abilities involved in the multifaceted tasks of reading and writing acquisition.

Table of contents

1. Introduction	11
1.1. Development of reading and writing in shallow and deep orthographies	13
1.2. Development of literacy. Stage theories	16
1.2.1. Stage theories and atypical reading development	19
1.3. The dual-route model of reading	20
1.3.1. Dual-route model of reading and atypical reading development	22
1.4. Computational models of reading: the self-teaching hypothesis	26
1.4.1. Self-teaching hypothesis and atypical reading development	29
1.5. Language prerequisites of literacy	30
1.6. Metaphonological prerequisites of literacy	35
1.6.1. Metaphonological skills and atypical reading development: the core phonological deficit hypothesis	38
1.6.2. Intervention on metaphonological skills	42
1.7. RAN	45
1.8. (Pre)literacy and literacy acquisition. Is there a bilingual advantage?	49
2. The research project: General method	53
2.1. Participants	54
2.2. Materials and general procedures	54
2.3. Language skills	56
2.4. Metaphnological skills	58
2.5. Short-term memory	58
2.6. RAN	58
2.7. Visual skills	59
2.8. Word decoding	59
2.9. Sublexical decoding	60
2.10. Text decoding	61
2.11. Reading comprehension	61
2.12. Word writing	62
2.13. Non-word writing	62
2.14. Text writing	63
3. Study 1. Development of pre-literacy and literacy in typically developing monolingual and mono-literate bilingual children in an Italian school setting	64

3.1. Aims 1	64
3.2. Method	64
3.2.1. Participants	64
3.2.2. Materials	65
3.2.3. Procedures	66
3.2.4. Analyses	69
3.3. Results 1	69
3.3.1. Language skills	70
3.3.2. Metaphonological skills	73
<i>Global metaphonological skills</i>	
<i>Analytical metaphonological skills</i>	
3.3.3. RAN	80
3.3.4. Short-term memory	83
3.3.5. Visual memory and discrimination	85
3.3.6. Literacy skills	87
<i>Sublexical decoding</i>	
<i>Word and non-word decoding Text decoding</i>	
<i>Reading comprehension</i>	
<i>Word and non-word writing Text writing</i>	
3.4. Discussion 1	97
3.4.1. Language skills	97
3.4.2. Metaphonological skills	103
<i>Global metaphonological skills</i>	
<i>Analytical metaphonological skills</i>	
3.4.3. RAN	108
3.4.4. Short-term memory	110
3.4.5. Visual memory and discrimination	112
3.4.6. Literacy skills	112
<i>Sublexical decoding</i>	
<i>Word and non-word decoding Text decoding</i>	
<i>Reading comprehension</i>	

Word and non-word writing

Text writing

3.5. Interim summary 1_____	119

4. Study 2. Effects of short-term memory, language, metaphonology, and RAN on the development of literacy in typically developing monolingual and monoliterate bilingual children in an Italian school setting _____121

4.1. Aims 2_____	121

4.2. Method 2_____	122

4.2.1. Participants, materials, and procedures_____	122

4.2.2. Analyses_____	122

4.3. Results 2_____	123

4.3.1. Effects of preliteracy at the end of kindergarten on literacy in the middle of the 1st grade_____	128

4.3.2. Longitudinal effects of preliteracy skills on word writing between the middle of the 1st grade and the middle of the 2nd grade_____	130

4.3.3. Longitudinal effects of preliteracy skills on text reading and word writing between the end of the 1st grade and the middle of the 2nd grade_	131

4.3.4. Longitudinal effects of 1st and 2nd grade preliteracy skills on literacy in the middle of the 2nd grade_____	132

4.3.5. Effects of language skills and spoonerism on literacy in the middle of the 2nd grade_____	135

4.4. Discussion 2_____	136

4.4.1. Effects of preliteracy skills at the end of kindergarten on literacy in the middle of the 1st grade_____	136

4.4.2. Longitudinal effects of preliteracy skills on word writing between the middle of the 1st grade and the middle of the 2nd grade_____	140

4.4.3. Longitudinal effects of preliteracy skills on text reading and word writing between the end of the 1st grade and the middle of the 2nd grade_	142

4.4.4. Longitudinal effects of 1st and 2nd-grade preliteracy skills on literacy in the middle of the 2nd grade_____	144

4.4.5. Effects of language skills and spoonerism on literacy in the middle of the 2nd grade_____	149

4.5. Interim summary 2_____	151

5. Study 3. Effects of metaphonological, RAN, and graphic interventions on the development of preliteracy and literacy in monolingual and mono-literate bilingual children in an Italian school setting	154
5.1. Aims 3	154
5.2. Method 3	155
5.2.1. Participants	155
5.2.2. Materials	156
5.2.3. Procedures	157
5.2.4. Analyses	161
5.3. Results 3	162
5.3.1. Effects of the training on pre-literacy skills	162
<i>Global metaphonological skills</i>	
<i>Analytical metaphonological skills</i>	
<i>RAN</i>	
<i>Visual skills</i>	
<i>Non-word repetition</i>	
<i>Syntactic skills</i>	
<i>Receptive and productive vocabulary</i>	
<i>Digit and word span</i>	
5.3.2. Effects of the training on literacy skills	178
<i>Word and non-word decoding</i>	
<i>Sublexical decoding</i>	
<i>Text decoding</i>	
<i>Reading comprehension</i>	
<i>Word writing</i>	
<i>Non-word writing</i>	
<i>Text writing</i>	
5.4. Discussion 3	191
5.4.1. Effects of the training on pre-literacy skills	191
<i>Global metaphonological skills</i>	
<i>Analytical metaphonological skills</i>	
<i>RAN</i>	
<i>Visual skills</i>	
<i>Non-word repetition</i>	

<i>Syntactic skills and vocabulary</i>	
<i>Digit and word span</i>	
5.4.2. Effects of the training on literacy skills	199
<i>Word and non-word decoding</i>	
<i>Sublexical decoding</i>	
<i>Text decoding</i>	
<i>Reading comprehension</i>	
<i>Word/non-word writing</i>	
<i>Text writing</i>	
5.5. Some methodological remarks	204
5.6. Interim summary 3	206
6. Study 4. Individual, social, and (meta)linguistic predictors of literacy difficulties	208
6.1. Aims 4	208
6.2. Method 4	209
6.2.1. Participants	209
<i>Description of the impaired sample</i>	
6.2.2. Materials	211
6.2.3. Procedures	211
<i>Testing</i>	
<i>Interventions</i>	
6.2.4. Analyses	212
6.3. Results 4	214
6.3.1. Individual and social predictors of literacy difficulties	215
6.3.2. Memory, linguistic, and metalinguistic predictors of literacy difficulties	218
<i>Short-term memory, phonological, and visual skills</i>	
<i>Language skills</i>	
<i>Global metaphonological skills</i>	
<i>Analytical metaphonological skills</i>	
<i>RAN & rhyme recognition</i>	
6.3.3. Effects of the intervention on reading and writing difficulties	241
6.4. Discussion 4	245
6.4.1. Individual predictors of literacy difficulties	245
6.4.2. Memory, linguistic and metalinguistic predictors of literacy difficulties	247

Short-term memory, phonological, and visual skills

Language skills

Global metaphonological skills

Analytical metaphonological skills

RAN and rhyme recognition

6.4.3. Effects of the intervention on reading and writing difficulties	264
6.5. Interim Summary 4	268
7. General discussion	272
8. Conclusion	281
References	284
Appendixes	326
Appendix A	326
Appendix B	332
Appendix C	343
Appendix D	364
Acknowledgments	379
Errata Corrige	381

1. Introduction

“Children are wired for sound, but print is an optional accessory that must be painstakingly bolted on” (Pinker, 1997). In this statement, there is both the naturalness and effort involved in adapting neural systems that evolved for managing sound, vision, and attention (Dehaene, 2009) to the complex tasks of reading and writing. These neural systems are genetically equipped to support visual processes, working memory, attention, motor movements, higher-level cognition mechanisms, as well as different language layers, namely phonology, morphology, syntax, and semantics (Norton & Wolf, 2012). Smooth functioning and integration of these processes is then a pre-requisite for effective acquisition of the written language, which is not genetically supported by our brain for being a “recent” conquest of humanity in evolutionary terms.

For the last fifty years, the study of reading processes, in particular in subjects with acquired (e.g., Denes, Cipollotti, & Zorzi, 1996) and developmental deficits (e.g., Ellis, 1984), has received increasing attention by the cognitive and neuropsychological literature, which led to identifying different models of reading and writing development (Frith, 1985; Coltheart et al., 2001). On the other hand, particular attention has been devoted to some abilities that have been taken as essential for literacy development, such as phonological awareness (Liberman, 1971) (or metaphonology), and Rapid Automatized Naming (RAN) (Denckla & Rudel, 1976).

Although the role of these abilities is widely shared among scholars, there is a shortage of longitudinal studies on metaphonology development (Dehaene, 2009), and on the relation between metaphonology and writing skills (Zoccolotti et al., 2005). This is even more true when we take into consideration languages other than English, such as Italian.

Mainstream literature is essentially based on native speakers of English, whose orthography presents a high degree of opacity, so findings might not be necessarily transferrable to

languages with shallow orthographies (see Wimmer, 1993), such as Italian. One practice consequence is that, in Italy, standard clinical protocols aiming at preventing reading and writing difficulties between kindergarten and primary school are almost completely absent (Marotta et al. 2008). A certain preference for training executive (e.g., graphism) rather constructive (e.g., metaphonology) preliteracy skills has also been observed (Perrotta & Brignola, 2010). Furthermore, mainstream literature has mainly concentrated on reading, whereas research on writing has received less attention (Zoccolotti et al., 2005). To our knowledge, research on literacy prerequisites in bilingual populations with minority languages in Italy is even more scarce.

The present study tries to fill these gaps by proposing a longitudinal analysis of reading and writing prerequisites and abilities in Italian monolingual children as compared to bilingual children with minority languages. A comprehensive analysis of language, metaphonology, and RAN abilities, besides working memory and some visual skills, is conducted in the period of transition from kindergarten to primary school.

The general aims are to investigate how diverse preliteracy skills contribute to the development of the first steps of literacy in a language with shallow orthography like Italian; whether this development can be enhanced through a specific training of some preliteracy skills, namely phonological awareness and RAN; and if such training has some effects in case of atypical development. Possible differences between monolingual and bilingual children in the development of these patterns are also explored.

This chapter describes the typical and atypical development of literacy according to recent models of cognitive neuropsychology, and the development of preliteracy skills in the presence and absence of training interventions, in light of some recent literature dealing with monolingual and bilingual children.

Chapter 2 describes the general method of the project (full sample, materials, and general procedures).

Chapters 3, 4, 5, and 6 describe and discuss the 4 studies constituting the research.

Chapters 7 and 8 summarize and conclude the research.

1.1. Development of reading and writing in shallow and deep orthographies.

Different from oral language, written language must be voluntarily learned through long and “painstakingly” learning processes and based on cognitive abilities that were not originally devoted to the orthographic elaboration¹. These processes formally begin by matching sounds to graphic symbols, an initially difficult operation which, in typical development, leads to decoding brief familiar words in a relatively short time. Parallely, children learn to trace orthographic signs. In the first school months, they acquire one by one phoneme-grapheme correspondence in their language and the orthographic equivalence between capital letters, lowercase, and uppercase. Later, they learn the main syllable transcription rules and irregular word transcription. In typical development, decoding gets more and more fluent and writing gets more and more effective during primary school and first secondary school (Zoccolotti et al., 2005). These procedures refer to the phonic or alphabetic method for learning to read (Chall, 1967), which is the most practiced in Italy and is coherent with the framework adopted by the present study, as opposed to the whole-word method (for a study comparing the two methods, see Alegria et al., 1982). It is also the most appropriate for shallow orthographies, such as the Italian one, in which it allows to read and correctly write the vast majority of words based on grapheme-phoneme conversion rules. Some exceptions are however possible.

¹ There is strong evidence obtained through functional magnetic resonance that reading acquisition partially recycles cortical areas evolved for the visual recognition of objects and faces, the so called *visual word form area* (Dehaene & Cohen, 2011).

In Italian reading, exceptions are mainly constituted by (Zoccolotti et al., 2005)

- few irregular words, e.g., *glicine*, where the string *GLI* is uttered as [gli] and not as [li];
- words with more than 3 syllables, which can be paroxytone (e.g., li[']mone), or proparoxytone (e.g., [']gondola), and sometimes generate semantic minimal pairs (an[']cora – [']ancora). The tonic stress position seems not to follow definite rules (Colombo, 1992), even though a certain degree of regularity is associated with some final sequences, such as –oro, which tends more frequently to generate paroxytones (e.g., al[']loro), or –ola, which tends more frequently to generate proparoxytones (e.g., [']bambola) (Burani & Arduino, 2004).

In Italian writing, some difficulties can depend on (Zoccolotti et al., 2005)

- non-homograph homophones, (e.g., *anno* – *hanno*);
- words including the homophone sequences *quo* / *cuo* (e.g., *liquore* / *cuore*);
- established loanwords (e.g., *jeans*; *mouse*), which in Italian maintain the original pronunciation.

Other frequent difficulties in the oral-writing direction depend on different regional pronunciation, for instance (Zoccolotti et al., 2005)

- in most North-Western varieties, words like *aglio* ['aλλo] and *olio* ['oljo] are pronounced with the same sound (λ);
- in Northern varieties, words like *cena* ['tʃena] and *cielo* ['tʃelo] are pronounced with the same initial syllable (tʃe), as well as words like *scelta* ['ʃelta] and *scienza* ['ʃentsa] are both pronounced with the initial syllable [ʃe];
- in Northern varieties, words like *genio* ['dʒɛnjo] and *segno* ['seɲno] are pronounced with the same sound [ɲ]

- in Northern and Southern varieties, the pronunciation of words including occlusive consonants followed by [l] or [r] does not allow to discriminate simple or geminate consonants, as happens in words like *libro* ['libro] and *febbre* ['fɛbbre], or *biblico* ['bibliko] and *pubblico* ['pubbliko], where the sound [b] is uttered in both cases as shortened in Northern varieties, and geminate in Southern varieties.

Angelelli et al. (2004) as well as Wimmer & Mayringer (2002) note that the higher frequency of ambiguities in phoneme-grapheme conversion, as compared to limited irregularities in the grapheme-phoneme conversion, has implications for the comparison of reading and writing performances in shallow orthographies, such as German, Finnish, and Serbo-Croatian, besides Italian.

Segmental reading in languages with deep orthographies, such as English (but also French and Hebrew), is much more difficult because the correspondence between phoneme and respective grapheme or graphemic string is often opaque. This does not occur with low-frequency word subsets but with high-frequency words.

In reading, this generates (Zoccolotti et al., 2005)

- graphemes or graphemic strings corresponding to different phonemes, such as the string *EA*, which can be uttered as [i:] in words like *veal*, [ɛ] in words like *head*, [ɑ:] in words like *heart*, [eɪ] in words like *steak*;
- exceptions to relatively regular pronunciations, such as for the grapheme *i*, which is generally uttered as [aɪ], except when it is followed by double consonants (e.g., *mint* [mɪnt], *hint* [hɪnt]), with possible exceptions to the exception (e.g., *kind* [kaɪnd], *find* [faɪnd], *pint* [paɪnt]); or for the grapheme *o*, which is generally uttered as [əʊ], except in words like *some* [sʌm], *blood* [blʌd];
- idiosyncratic uttering, such as in words like *yacht* [jɒt], or *colonel* ['kɜ:n(ə)l].

Ambiguities in English writing are generated by non-homograph homophone sequences like (Zoccolotti et al., 2005)

- [eɪn], which can be written as *EI* (e.g., *sane*), or *AI* (e.g., *brain*);
- [i:], which can be written as *EE* (e.g., *week*) or *EA* (e.g., *weak*);
- ['mɔ:r.nɪŋ], which can correspond to the words *morning* or *mourning*;
- [sel], which can correspond to the words *cell* or *sell*;
- [saɪt], which can correspond to the words *site*, *sight*, or *cite*;
- ['ɜ:nɪst], which can notoriously correspond to the word *earnest* and the name *Ernest*.

These few examples can give an idea of some difficulties implied in learning to read in languages with different orthographies. Such difficulties require the reader to adopt diverse reading modes, which have been described by the neuropsychological literature through cognitive models of reading elaboration.

1.2. Development of literacy. Stage theories

The literature on literacy acquisition has mainly focused on English. The models proposed by Frith (1985), Marsh et al. (1981), and Seymour & MacGregor (1984) are based on the common idea that literacy is acquired according to sequential stages that are relatively independent of each other (for a review of the three theories, see Stuart & Coltheart, 1988).

Frith's model (1985) distinguishes three implicational main phases (see Figure 1):

(1) a *logographic* stage, in which written strings are treated holistically, similar to what happens for the recognition of pictures, based on graphic salient features and contextual and pragmatic strategies. Recognition is instantaneous, without decomposing letter sequences, so that phonology is retrieved after whole word recognition (Stuart & Coltheart,

1988). Children can reach a remarkable competence through this elementary strategy, even in the range of hundreds of words (Zoccolotti et al., 2005);

(2) an *alphabetic* stage, in which written strings are sequentially alphabetically decomposed into grapheme sequences, and correspondences between individual letters and phonemes are used. In this phase, children can read (not necessarily correctly) unfamiliar words and non-words.

(3) an *orthographic* stage, in which orthographic units are instantly analyzed without phonological conversion as parts of the learner's orthographic lexicon.

Frith (1985) does not explain how the transition from one to the next phase takes place, so we could assume that the following strategy is adopted when the previous one is not sufficient anymore. Previous strategies should remain available to the reader once they are acquired, as the decoding of lexical units is not sufficient, per se, to decode non-words.

Each phase is characterized by the acquisition of new procedures and the consolidation and automation of the ones that are already acquired, alternately for reading and writing, as shown in Figure 1.

<i>Step</i>	<i>Reading</i>	<i>Writing</i>
1a	<i>logographic</i> ₁	(symbolic)
1b	<i>logographic</i> ₂	<i>logographic</i> ₂
2a	<i>logographic</i> ₃	<i>alphabetic</i> ₁
2b	<i>alphabetic</i> ₂	<i>alphabetic</i> ₂
3a	<i>orthographic</i> ₁	<i>alphabetic</i> ₃
3b	<i>orthographic</i> ₂	<i>orthographic</i> ₂

Figure 1. Frith's model of the acquisition of reading and writing (1985).

The (1) logographic and the (3) orthographic phase are firstly acquired in reading, whereas the (2) alphabetic step is firstly made in writing. Once the respective strategies are consolidated enough, they transfer to the other ability.

The cognitive-developmental model proposed by Marsh et al. (1981) distinguishes four stages: (1) linguistic guessing, (2) discrimination net guessing, (3) sequential decoding, and (4) hierarchical decoding. Stages (1) and (2) approximately include processes that are covered by Frith's logographic phase, with the difference that, in stage (1), only rote learning and linguistic guessing are available to learners to recognize unfamiliar words, without regard to graphic features, whereas, in stage (2), guessing is based on both visual and linguistic cues. Stage (3) is similar to Frith's alphabetic phase, with the difference that the former is limited to sequential decoding of regular CVC words containing a short vowel (e.g., *men*). Stage (4) covers the abilities involved in Frith's advanced alphabetic phase and orthographic phase. Learners can make use of analogy and higher-order rules, such as conditional rules allowing them to pronounce the same letter differently according to neighboring letters.

Seymour & McGregor (1984) proposed a model based on Frith (1985), which distinguishes three broad and partially overlapping phases. A (1) logographic stage, an (2) alphabetic stage, and an (3) orthographic stage.

Stuart & Coltheart (1988) showed that mere stage theories are not sufficiently accurate in describing reading acquisition, as they do not take into consideration the support of phonological skills and knowledge of letter-sound correspondences. Learners who can use phonological insights will do so from the very beginning of learning to read and overtake the logographic stage, whereas learners with insufficient phonological skills will resort to a visual mediation.

Significant differences are reported according to the orthographic depth of learners' language. At the end of primary school, speakers of languages with shallow orthographies (Italian, German, Finnish, or Greek) reach 95% accuracy in decoding short words, whereas

speakers of languages with deeper orthographies are much less accurate (French = 79%; Danish = 71%; English = 35%), as shown by Seymour et al. (2003). The authors argue that learners of deep orthographies need both a logographic and alphabetic foundation, whereas learners of shallow orthographies need only an alphabetic foundation. Research on Austrian children did not find evidence of definite acquisitional phases for reading and writing, as described above (Wimmer & Hummer, 1990).

In a cross-sectional study, Orsolini et al. (2003) documented four different strategies in the very early steps of reading acquisition by 20 Italian children, namely (1) guessing through orthographic cues (initial letter or syllable); (2) segmental alphabetic decoding without sound blending; (3) an alphabetic strategy like grapheme-to-syllable conversion; and (4) proficient alphabetic strategies. Neither less proficient ((1) and (2)) nor more proficient ((3) and (4)) strategies were associated with phonological skills.

The above-mentioned analyses of shallow-orthography languages, in line with the analyses conducted in English by Stuart & Coltheart (1988), confirm great individual variability in the first steps of reading acquisition and suggest that stage theories might not be enough to explain the complexity of literacy acquisition.

Recent research tried to explain this complexity through dual-route models (section 1.3) and computational models (section 1.4) of reading and writing.

1.2.1. Stage theories and atypical reading development

In the framework of stage theories of literacy acquisition, atypical development might be due to a *delay* or *deviance* in the acquisition of one of the phases (Frith, 1985). In case of delay, the learner is expected to develop the normal sequence of stages at a slower pace, whereas, in case of deviance, compensatory strategies will be adopted after the point of failure.

Since a characteristic of typical development is the possibility of merging different strategies belonging to different phases (see Figure 1), atypical development might provoke the

impossibility of putting together different strategies rather than the inability to acquire new ones, thus resulting in six subtypes of developmental disorders. In any case, Frith's theory predicts that children with dyslexia/dysorthography can use logographic strategies.

According to the model of Seymour & McGregor (1984), phonological dyslexia can be interpreted as either a delay or a deviation in the acquisition of the alphabetic and orthographic lexicons, forcing the learner to rely on the logographic lexicon, whereas morphemic dyslexia as a disturbance in both the logographic and the orthographic lexicons. Seymour et al. (2003) suggest that languages with opaque spelling need both a logographic and an alphabetic foundation, whereas those with transparent spelling only need an alphabetic foundation.

1.3. The dual-route model of reading

There is wide consensus on the fact that an accurate model to describe typical reading aloud processes should include both a sublexical and a lexical route² (Coltheart et al., 2001; for review, see Castles, 2006). The lexical route allows reading by accessing a lexicon or memory store of previously seen written words, whereas the sublexical (or non-lexical) route uses a set of translation rules to convert graphemes into phonemes. The lexical route allows to successfully process familiar words, but not unfamiliar words or non-words. The core of the lexical route consists of two lexicon storages containing, respectively, the orthographic and the phonological lexicon. A known written word activates the correspondent entry in the

² A third route of reading involving a direct connection between lexical orthographic representations and lexical phonological representations has been hypothesized to account for the preserved oral reading of some subjects with severe semantic impairments in combination with a disruption of the sublexical phonological route. The reading output would result in partial lexical semantic information combined with partial sublexical phonological information. Word reading could be preserved even for objects that cannot be named, along with impaired non-word reading. For a case study, see Wu et al. (2002).

input (visual) orthographic lexicon. This lexicon is organized by written word frequency, hence, compared to words with similar orthographic (and phonological) properties, the more frequent, the more accessible words are. At this point, the activated lexical representation in the input orthographic storage can either be processed through the semantic system assigning meaning to the read word or be directly sent to the phonological output lexicon assigning phonological information to the read known word. The direct connection between the two lexicon storages, possibly mediated by the semantic system, allows for both accurate and fast conversion.

Unknown words must be necessarily processed through the sublexical route, which applies a set of conversion rules without resorting to lexical resources, thus allowing to read every kind of regular words and non-words. This process guarantees accurate decoding of strings that adhere to orthographic and phonotactic rules of the language but requires a longer time for conversion.

At the top of the dual-route model, there is a common stage of orthographic visual analysis, which is responsible for letter identification, the encoding of letter position within the word, and binding of letters to words. In the last stage, the phonological string generated through either the former or the latter central route is sent to the phonological output buffer, a short-term interface between phonological representations and articulatory motor programming, having the function to keep the information until full production and, possibly, assembling phonological strings into larger units (Castles, 2006; Friedmann & Coltheart, 2018; Zoccolotti et al., 2005).

Similar to reading aloud, writing under dictation in alphabetized subjects with typical development is implemented along a dual-route model of writing. At the top of both routes, there is a common level of acoustic analysis of acoustic stimuli. In skilled readers, after the recognition of a known acoustic stimulus, this is sent to the input phonological lexicon, where lexical entries are stored; at this point, the entry may either activate the semantic system

assigning meaning to the heard word or be directly sent to the output orthographic lexicon, where words' orthographic knowledge is stored. This procedure allows writing regular and irregular known words but not non-words and words whose orthography has not been learned before. If the peripheral acoustic analysis stage does not recognize the heard stimulus as known, this is sent to the acoustic phonological converter, where it is translated into the correspondent phoneme sequence; this is temporarily stored in the phonological buffer, which rules the sequential elaboration through the phoneme-grapheme conversion mechanism. This process allows writing every kind of regular words and non-words, but not words with irregular orthography, which must be necessarily processed through the lexical route. The orthographic strings elaborated through either the former or the latter route are sent to the grapheme buffer, a short-term interface between the orthographic representation and writing motor programming (Zoccolotti et al., 2005).

1.3.1. Dual-route model of reading and atypical reading development

According to the framework described above, reading and writing deficits can be caused by atypical functioning or non-functioning of specific components of the dual-route model. This view, which is common in the realm of acquired dyslexia, has been recently introduced also in the study of developmental dyslexia (Friedmann & Haddad-Hanna, 2014), thus allowing the identification of 19 different types of developmental dyslexias, depending not only on the central routes (namely the lexical and sublexical route) but also on the peripheral stage of reading, which is involved in the orthographic visual analysis of written strings (Friedmann & Coltheart, 2018). A deficit in this area would be responsible for peripheral dyslexias. Subjects can identify letters correctly but fail encoding letter order within words (letter position dyslexia). Transpositions generally occur in adjacent middle letters, in particular

when migrations generate an existing and more frequent word (e.g., (in English³) *loin* → *lion*; (in Italian⁴) *ovile* → *olive*). Letters can migrate between neighboring words (attentional dyslexia), both horizontally and vertically, generally but not necessarily maintaining their original relative position (e.g., (in English) *fight late* → *light fate*; (in Italian) *vena telo* → *vena velo*). Even in this case, the lexicality of the target word affects migrations, namely migrations are more frequent when target words are existent. Word length also affects migrations (longer strings are more affected), as well as similarity (migrations occur more often between similar words), word order (migrations are more frequent from the first to the second word than vice versa), and letter position (final letters are more likely to migrate).

A deficit in the function of creating abstract letter identities can be responsible for incorrect identification, omissions and substitutions of letters in isolation and within words, and missing responses (letter identity dyslexia). Omissions can involve only one side of the word or text (neglect dyslexia or neglexia), typically the left side, thus resulting in a deficit in the prelexical stage leading to omit portions of words (especially when those are affixes), entire words, or entire phrases.

A deficit in the output of the orthographic visual-analyzer (visual dyslexia) can affect all functions of the orthographic-visual analysis, leading to all the above-mentioned errors, namely letter identification, letter position, migrations, and letter-to-word binding, even in atypical fashions.

A deficit in the lexical or sublexical route of reading is responsible for central dyslexias. A deficit in the lexical route (surface dyslexia) determines reading only through grapheme-to-phoneme conversion, thus leading to regularization of irregular graphemes (e.g. sonorization of silent letters as in *comb*); errors in reading ambi-phonetic graphemes (e.g., *beer* – *bear*), homophones (e.g., *weather* → *whether*), and potentiophones, namely words

³ English examples in this section stem from Friedmann & Coltheart (2018).

⁴ Italian examples in this section stem from Traficante et al. (2017) and Casani (2019).

that can result into another existing word when they are read through the sublexical route (e.g., *now* → *know*). In shallow orthographies like Italian, surface dyslexia causes slower than normal reading (Spinelli et al., 1997; Zoccolotti et al., 1999) as well as errors in polysyllabic words with unpredictable stress (Zoccolotti et al., 1999), such as proparoxytones (e.g. [']*lirica* → *l[']*rica*).

A deficit in the sublexical route (phonological dyslexia) is characterized by difficulty in reading pseudowords and unknown words alongside the correct reading of words that are stored in the input lexicon. Other problems might lie in the conversion of single letters into phonemes (letter-to-phoneme conversion phonological dyslexia) or in the conversion of graphemes that include more than one letter (e.g. *chip* → *kip*) (multi-letter phonological dyslexia). A selective impairment in the sublexical route can lead to omission, substitution, transposition, and addition of vowel letters (vowel dyslexia), in particular within pseudowords and unknown words where errors result in an existing word (e.g., *bit* → *but*). A deficit in the output buffer, which is dedicated to storing phonemes for the brief time that is necessary to conversion (phonological-output-buffer phonological dyslexia), is responsible for errors in reading longer words, especially if these are morphologically complex, and longer nonwords, with omissions, substitutions, and transpositions of some phonemes. Problems are also present in the repetition and spontaneous uttering of morphologically complex words and longer nonwords (e.g., *ringraziamenti* → *rangrizzamenti*; *sentenziava* → *sensiva*).

A deficit in both the lexical and the sublexical route (deep dyslexia) forces the reader to read via meaning, thus making semantic errors (e.g., *lime* → *lemon*), morphological errors (*played* → *play*), and some visual errors (e.g., *gum* → *game*). Function words and morphologically complex words are particularly problematic: the former are generally substituted for visually similar concrete words, for other function words, or they are just omitted; the latter are simplified through omissions and substitutions of morphological affixes (*smiled* → *smile*). Subjects with deep dyslexia can show syntactic deficits.

A deficit in the connection from the orthographic input lexicon to the semantic lexicon alongside the preserved sublexical route, as well as preserved lexical route between the orthographic input lexicon and the phonological output, causes correct reading aloud of every word type but no comprehension at all (direct dyslexia). In this case, the problem is not a semantic deficit or a deficit in the conceptual system in itself, as this would be evident from semantic errors in speaking in the former case (Howard & Gatehouse, 2006), and from the production of unrelated words in speaking in the latter case (Friedmann et al., 2013). An assessment word decoding battery implemented in this framework by Friedmann & Gvion (2003) has been recently adapted to the Italian language by Traficante et al. (2017)⁵. Their pilot study found 6 different types of developmental dyslexias in 52 Italian poor readers compared to 210 normal readers from the 2nd to the 5th grade of primary school, thus showing that it is possible to discriminate several types of reading impairments due to selective segments of the dual-route model of reading even in a shallow orthography like the Italian. Casani (2019) applied a coding scheme based on Friedmann & Coltheart (2018) to the text reading (Cornoldi & Colpo, 2011) of 21 children with developmental dyslexia, four of which with a developmental language disorder, compared to 32 typically developing children from the 1st to the 2nd grade. He could detect 11 different error types. Figure 2 shows the error types detected in Italian children.

⁵ The still unpublished test was kindly provided by Daniela Traficante and Claudio Giuseppe Luzzatti to be used in the experimental section of the present study.

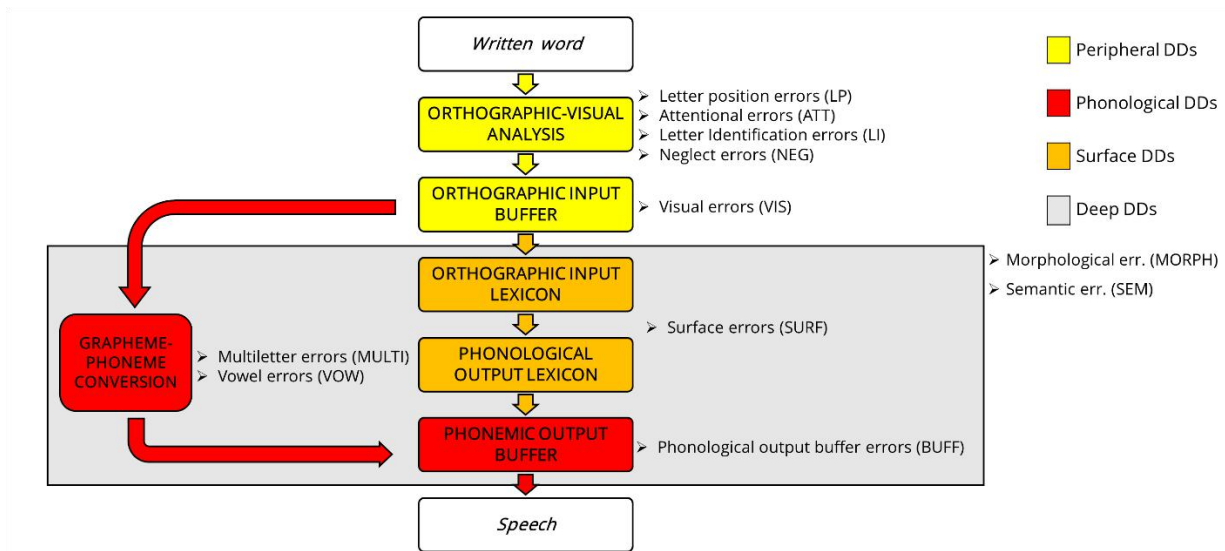


Figure 2. Main decoding errors found in Italian children and types of developmental dyslexias associated with different components of the dual-route model of reading aloud (Casani, 2019; adapted after Friedmann & Coltheart, 2018).

Children with language disorders presented the most compromised situation, with significant error proportions due to a combined deficit in the sublexical route and the orthographic-visual analysis stage. The analyses of individual performances confirmed an impairment in both the lexical and sublexical route in 3 out of 4 children with language disorders. Interestingly, both Traficante et al. (2017) through the word lists and Casani (2019) through the text reading found, respectively, 11 and 2 cases of poor readers that had not been detected through standard measures. This encourages a fine-grained error analysis in reading assessment, even in shallow-orthography languages like Italian.

1.4. Computational models of reading: the *self-teaching hypothesis*

Besides stage-models and dual-route models of literacy, the building of connectionist computational models simulating detailed aspects of human behavior has been used since

the late 1980s to investigate the development of children's reading (for an overview, see Seidenberg, 2007).

In this framework, Perry et al. (2013) proposed the self-teaching hypothesis of phonological decoding. They discuss previous connectionist models by arguing that it is highly implausible that an "external teacher" might supply the pronunciation of all words through exposure to a very large corpus of spelling-sound pairs. Conversely, they suggest that reading development starts from the explicit teaching of a small number of spelling-to-sound mappings. These early rudimentary decoding skills, combined with phonological representations of spoken words that are part of the learners' lexicon before learning to read, provide an internally generated signal, which gradually and interactively improves decoding, orthography, and lexicon. This is a connectionist dual-process model (Perry et al., 2007; 2010; 2013) that includes a non-lexical process mapping orthography to phonology in a two-layer associative network⁶, and a lexical process connecting orthography to phonology in a hard-wired interactive activation network. The non-lexical network quickly learns linear relationships between strings of graphemes and strings of phonemes, allowing for decoding also non-words, with possible phonological errors in ambiguous or complex words. The lexical process, in turn, links the orthographic entries of words to their phonological correspondences, allowing them to read any type of word except non-words. In typical conditions, reading aloud takes place thanks to the reciprocally integrated output of the two processes.

Figure 3 (from Ziegler et al., 2020) shows the three main phases implementing the self-teaching hypothesis of phonological decoding. The initial small network (1st phase) is formed through the explicit teaching of a small set of grapheme-phoneme correspondences. The

⁶ The two-layer associative network takes graphemes (letters or simple letter combinations) as input and uses phonemes as output. Grapheme-phoneme correspondences are acquired, yet at a slower pace, even though the teaching is based on syllables, morphemes, or whole words (Hutzel et al., 2004).

initial network allows for decoding words having a preexisting representation in the phonological lexicon (2nd phase). The activation of a preexisting word in the phonological lexicon allows for the creation of an orthographic entry (3rd phase) so that phonology works as a self-teaching device (or “built-in teacher”) refining and strengthening the network of letter-sound connections. The teaching signal that is internally generated by phonology contributes to increasing the decoding network.

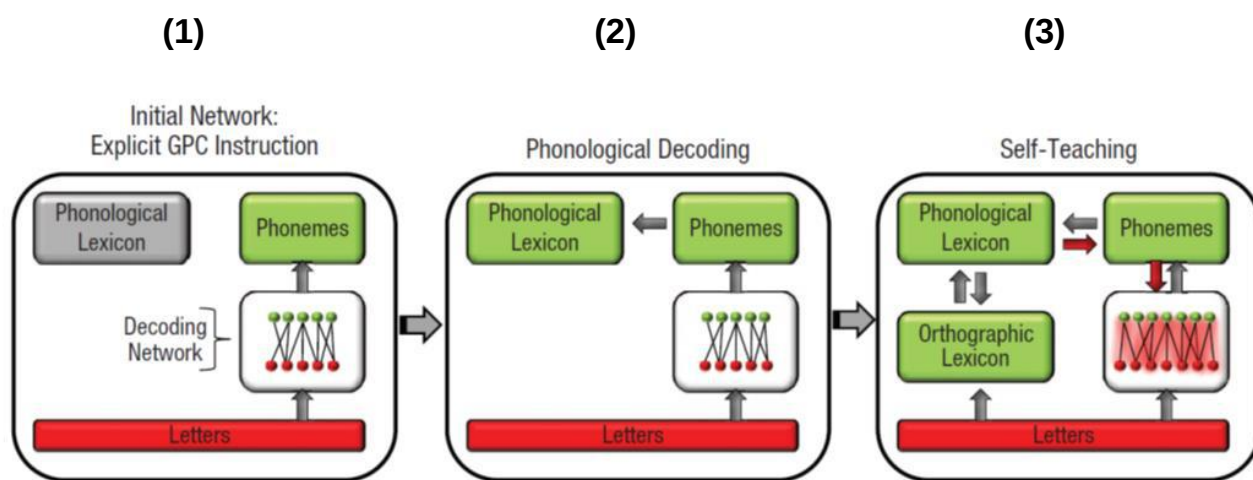


Figure 3. Self-teaching hypothesis of phonological decoding (Ziegler et al., 2020)

- (1) A small decoding network is formed through explicit teaching of a restricted number of grapheme-phoneme correspondences.
- (2) The initial network activates the decoding of preexisting words in the phonological lexicon.
- (3) An orthographic entry is created, which increases the decoding network.

Starting from a reduced set of grapheme-phoneme correspondences (1), the model can acquire word-specific orthographic representations for more than 25'000 words. The most plausible candidate among the activated phonological competitors is chosen by resorting to language cues, namely semantic, morphology, and lexicon information. Learners' vocabulary seems to have a relevant role in the choice of the right candidate. In a 3-year longitudinal study conducted on 140 children, Tunmer & Chapman (2012) showed that vocabulary influenced reading comprehension directly, and future decoding and word recognition indirectly through a *set of variability*, namely the ability to determine the correct

approximation in the pronunciation of spoken English words (see Venezky, 1999). Set of variability, in turn, influenced future reading comprehension indirectly through both decoding and word recognition. Learners' vocabulary showed to have a relevant role in the correction of partial decoding attempts by reference to word meaning and pronunciation (Dyson et al., 2017). This is particularly relevant to the 20% (English) words with irregular pronunciation, as self-teaching was revealed to be efficient for learning more than 80% of the words (Ziegler et al., 2020).

1.4.1. Self-teaching hypothesis and atypical reading development

Perry et al. (2019) and Ziegler et al. (2020) showed that computational models allow simulating individual learning trajectories and intervention outcomes based on the five main component skills of the self-teaching-hypothesis model: letters, phonemes, decoding network, phonological lexicon, and orthographic lexicon (see Figure 3). The simulation was conducted on one of the biggest samples presented in the English literature (Peterson et al., 2012), which describes decoding accuracy in 622 children, of which 388 with developmental dyslexia. The performances in orthographic choice, phoneme deletion, and vocabulary were taken, respectively, as measures of processing efficiency in the orthographic lexicon, correct activation of phonemes, and size of the phonological lexicon. A series of models were analyzed which took alternatively into exam some of the main deficits documented by the literature for developmental dyslexia, namely a phonological deficit (Hulme et al., 2015; Snowling, 2001), a visual-deficit (Stein & Walsh, 1997), and a global noise model assuming general processing inefficiency due to noisy computations

(Hancock et al., 2017)⁷. All three models fit behavioral results, with the best fitting for the phonological model, followed by the visual-deficit, and the noisy-computation model, but the best performance was observed in the multi-deficit model.

These studies advocate in favor of a multifactorial (computational) approach to reading. This has practical implications for dyslexia and intervention.

Dyson et al. (2017), through a longitudinal intervention design, showed that children can be taught to correct errors in the pronunciation of irregular words, and this can produce generalized effects on learning to read. The above-mentioned computation designs, in turn, show that increasing vocabulary is more beneficial for reading irregular words than nonwords, whereas increasing the efficiency of phonological processing determines the opposite pattern, and increasing orthographic efficiency is beneficial for every word type. The presence of marked interindividual differences suggests that the choice of fitting interventions largely depends on subjects' initial conditions (Perry et al., 2019).

1.5. Language prerequisites of literacy

The studies cited so far show that there is a strong interplay between decoding and oral language, in particular phonology and lexicon.

Nation & Snowling (2004) showed that children's oral language proficiency (age 8;6 – 13;0) influences the course of reading development. They found that three measures of non-phonological oral language tapping vocabulary knowledge and listening comprehension

⁷Research on subjects with atypical reading development has documented the presence of phonological and language deficits (Hulme et al., 2015; Snowling, 2001), visual deficits (Stein & Walsh, 1997), attentional deficits (Vidyasagar & Pammer, 2010), or auditory and temporal processes deficits (Vandermosten et al., 2010). This led to hypothesize different reasons for developmental dyslexia.

predicted individual differences in reading comprehension, word recognition, and irregular word reading 4 ½ years later, after controlling for autoregressive effects.

According to Frith's model (1985), functional strings like morphemes can be read as a whole as children acquire the orthographic phase. This can let us hypothesize that higher morphosyntactic competence might ease written decoding.

Casani (2019) found a significant presence of morphology errors and semantic errors in text decoding⁸. This led to deepening the possible effects of children's language competence on decoding. Strong predictive effects of syntactic comprehension and production⁹ on decoding accuracy emerged, along with 3 error categories, namely morphology errors (which showed the strongest effect), semantic errors, and phonological-output-buffer errors (i.e., errors in long and morphologically complex words), that have to do with morphology to different extents (Casani, in preparation). The role of morphological information on both decoding speed and accuracy has been emphasized in different studies. Morphology is reported to provide reading strategies for correctly decoding and spelling unfamiliar words (Verhoeven & Perfetti, 2003, 2011). Morphological awareness was found to be more and more predictive of both reading and spelling achievement in children exposed to different orthographies as their grade increases (Mann and Singson, 2003; Berninger et al., 2010; Casalis et al., 2011), and to have different degrees of association with word recognition and spelling in different scripts, by predicting unique variance in addition to phonological awareness (McBride-Chang et al., 2003, 2005). Reading and spelling performances seem to be fostered by access to word morphemic constituents in different languages varying for morphological

⁸ The significant presence of a combination between morphology and semantic errors in reading might indicate the presence of a deep dyslexia (see 1.3.1), which depends on a deficit in both the lexical and sublexical route of reading (Friedmann & Coltheart, 2018). Along with this possibility, we hypothesized a possible delay in the development of the orthographic step of reading (see Frith, 1985), which could hamper the recognition of functional strings.

⁹ Syntactic comprehension was measured through a standard sentence - picture matching test (Bisiacchi et al., 2005), whereas syntactic production through the first version of the clitic-production test used in the present study (Arosio et al., 2014).

richness and orthographic transparency (Verhoeven & Perfetti, 2011). In Italian, in particular, morphology was found to affect the reading fluency of typically developing children (2nd to 7th grade), and the decoding speed of children with dyslexia in decoding non-words made up of a root and a possible derivational suffix compared to non-words that do not include existing morphemes (Burani et al., 2002, 2008; Traficante et al., 2011). Morphological facilitation of word reading speed was also found in 2nd graders and children with dyslexia, whereas in older skilled readers it was limited to low-frequency words (Burani et al., 2008; Marcolini et al., 2011). Angelelli et al. (2014) found that in Italian, morphological information is a useful resource for both reading and spelling, as typically developing children benefit from the presence of morphological structure when they read and spell non-words. In processing low-frequency words, however, morphology facilitates reading but not spelling. They attribute their results to successful cooperation between lexical and sublexical processes in reading and spelling, which facilitate morpho-lexical access. The mentioned studies converge on the idea that access to morphemes can facilitate lexical reading for low-frequency words that otherwise would not probably be represented as a whole in the mental lexicon. This also happens in Italian and in other shallow orthographies (for Spanish, see Defior et al., 2008) where the orthography-phonology mapping might be expected as sufficient for correct decoding and spelling. The predictive effects of syntactic comprehension and production on decoding accuracy found by Casani (2019) might be expected for the production of clitic pronouns, as their scarce phonological salience matched to their high load of morphosyntactic information might expose them to be either overlooked or avoided during reading. Nevertheless, the effect of overall syntactic comprehension, which was even stronger, suggests that more general reasons related to (morpho)syntactic competence might be plausible. Disentangling the specific grammatic structures that contribute to decoding accuracy might shed light on this aspect.

What has been said so far can help delineate the role that language competence plays in learning to read effectively. Following the dual-route model of reading, access to the lexicon and its structural properties is a pre-requisite for the activation of the lexical route of reading, which allows fluent decoding. Moreover, according to the self-teaching hypothesis, vocabulary and knowledge of word meanings are important to correct, incorrect, or partial decoding attempts. For these reasons, vocabulary is expected to play a central role in reading.

“For adequate reading comprehension from grade three on, children require both fluent word recognition skills, and an average or greater vocabulary” even though “the presence of these two accomplishments does not guarantee a high level of reading comprehension but the absence of either word recognition or adequate vocabulary insures a low level of reading comprehension” (Biemiller, 2005). Vocabulary size per se does not however seem to be a condition for 1st and 2nd-grade success, but its effects can be appreciated in the long term. A relationship between oral receptive vocabulary in 1st grade and reading comprehension in 11th grade, accounting for 30% of the variance ($r = 0.55$), was stressed by Cunningham & Stanovich (1997). Age-normal vocabulary demands typically appear in 3rd or 4th-grade books (Chall et al., 1990; Chall et al., 1991). This might result in underrating the importance of vocabulary between kindergarten and the early years of primary school. This could also be favored by the fact that direct intervention on vocabulary is not likely to show immediate evidence, as revealed by research involving vocabulary intervention programs with upper elementary children and elder groups, which reported small albeit significant effects on general vocabulary or reading (Stahl & Fairbanks, 1986; Edwards et al, 2003). There is some evidence, however, that “children with small vocabularies do acquire new words during instruction as well as children with larger vocabularies, suggesting that much of the difference is a matter of opportunity rather than ability” (Biemiller, 2005).

Studies conducted on December- and January-born children attending the same class, namely children who differ in age by one month but in schooling by one year, showed similar vocabulary sizes, with a non-significant impact of schooling by the end of 2nd grade, which might correspond to the end of pre-literacy for several children (Cantalini, 1987). Although there are attempts to directly speed up vocabulary acquisition (for the “fast mapping” theory, see Carey, 1978), there is little evidence for significant improvement in the acquisition of word meanings in early primary years (up to 2nd grade), whereas there is evidence that differences in 2nd grade between vocabulary size of average children (about 6000 root word meanings) and children in the lowest quartile (about 4000 root word meanings) are related to marked variations in home vocabulary and support (Biemiller, 2005).

Hart & Risley (cited in Hart & Risley, 2003) designed a half-day intervention in Arkansas preschools and found that they could easily increase children's vocabulary size by teaching them new words, but they could not accelerate their vocabulary growth rate because the effects were transitory. Namely, they could not intervene in the developmental trajectory. Through a longitudinal observation of 42 American children in their family context from age 7 months to 3 years, Hart & Risley (2003) found that in the first 4 years of life, an average child in a professional family would accumulate experience with almost 45 million words, an average child in a working-class family with 26 million words, and an average child in a welfare family with 13 million words. They also calculated that in the first 4 years of life, the average child in a professional family would have accumulated 560'000 more instances of encouraging than discouraging feedback, the average child in the working-class family 100'000 more encouragements than discouragements, but an average child in a welfare family would have accumulated 125'000 more instances of prohibitions than encouragements. By estimating the number of intervention hours needed to equalize children's early experience, they conclude that the problem of skill differences among children at the time of school entry is crucial and intractable because an intervention aiming

at increasing vocabulary acquisition cannot address just one knowledge or skill, but an entire approach to experience. The effect of interventions on vocabulary conducted in different contexts confirms these difficulties. A meta-analysis conducted on 37 studies to explore the impact of vocabulary on comprehension found a modest relationship ($r = 0.43$). Custom comprehension measures ($d = 0.50$) but not standard measures ($d = 0.10$) increased. An intervention study conducted on 1'232 children in 4th and 5th grade in 24 American schools found small average distal treatment effects (*approx.* $d = 0.15$) compared with proximal effects (*approx.* $d = 1.24$), then confirming modest transfer to global vocabulary and comprehension, in contrast with significant intensity and depth of instruction for learning the taught corpus of words.

These data confirm the difficulty to directly intervene on vocabulary and, what is interesting to our aims, to transfer possible effects to literacy. This is particularly evident because vocabulary has been mainly treated concerning reading comprehension, which is not reliably measurable in the very early steps of literacy when children are learning to decode and write single words or short texts. In this regard, other measures might be more indicative prerequisites of literacy rather than vocabulary size. We refer to the abilities to face structural properties of the lexicon, as well as to quickly retrieve lexicon from memory, which can be measured through metalinguistic skills (in particular metaphonological skills in the phase of preliteracy), and rapid naming of visual stimuli.

1.6. Metaphonological prerequisites of literacy

Metaphonological skills (or phonological awareness, or, at a finer level, phonemic awareness) are a set of abilities allowing to perceive, recognize, and manipulate the sounds (syllables and phonemes) of the spoken language (Bortolini, 1995; Cheung et al., 2001).

They are considered prerequisites for literacy acquisition (Ball & Blachmann, 1988; Bradley, 1988; Frith, 1985; Hulme & Snowling, 1988; Lundberg et al., 1988), especially in alphabetic languages (Morais, 1991). The literature reports relations among all the major domains of language awareness (phonological, word, and structural) and between language awareness and overall linguistic skills, highlighting that literacy knowledge correlates with metalinguistic skills and, specifically, with phonological awareness (Chaney, 1992).

Alegria et al. (1982) showed that children learning to read according to a whole-word method are specifically unable to manipulate phonemes, contrasting with children learning to read according to a phonic method¹⁰. In their study, the phonic group performed significantly better than the whole-word group in a phonic reversal task, but similarly in the syllable reversal task, thus indicating that 6-year-old children are ready to discover that speech consists of a sequence of phones and that the time in which they do so is influenced by the way they learn to read. This is the transition from a global to an analytical awareness, namely from the syllable level to the phoneme level (Bradley & Bryant, 1983; Morais, 1991). The global awareness allows for discrimination (e.g., recognition of word minimal pairs), classification (e.g., rhyme and syllable recognition), blending, and segmentation of syllables. It is also present in subjects with non-alphabetic languages, illiterate adults, and preschool children (Martini, 1995; Martini et al., 2003). Syllables are acoustically more marked than phonemes and can individually be uttered with clear perceptual autonomy. For this reason, children spontaneously develop syllable knowledge before phoneme knowledge,

¹⁰ This position was better specified by Morais et al. (1987), who showed that “first-graders taught to read according to a pure whole - word method (it was carefully checked that no analysis of the words into segments was provided by the teachers) made, on average, little progress from the fourth to the ninth month of school (6% and 18% of correct responses, respectively) in the task of initial consonant deletion, with large interindividual differences. About 13% of the subjects made considerable progress, reaching a high level of performance in the retest. Either these children had received supplementary instruction at home or they had abstracted the segmental structure of speech from their experience with the alphabetic material. Anyway, such experience should not be assimilated to mere exposure.” They concluded that “the ability to read new words in the first stages of learning to read depends critically on segmental ability, even when the method of instruction does not specify the segments” and “receiving reading instruction is not sufficient to develop segmental analysis ability but alphabetic literacy is (almost) a sufficient indication of segmental skill” (Morais et al., 1987: 423-424).

independently of the acquisition of written language (Zuccheromaglio, 1985). Tasks involving minimal units like phonemes, conversely, are more challenging because they require abstraction of the phonemic speech unit, and "the increase in ability to segment phonetically might result from the reading instruction that begins between five and six. Alternatively, it might be a manifestation of some kind of intellectual maturation" (Lieberman et al., 1977: 213). This maturation in second childhood involves the development of general awareness of their intellectual processes, including metalinguistic activities (Bortolini, 1995).

Analytical or phonemic awareness is not present in subjects with non-alphabetic languages and illiterate subjects. The abilities involved concern phonemic blending and segmentation, manipulation (e.g., syllable or letter deletion; spoonerism, i.e. reversal of initial phonemes), and classification (e.g., rhyme recognition and production, and phonemic verbal fluency).

Some cross-sectional studies tried to describe the developmental patterns of metaphonological skills in Italian. Devescovi et al. (1989) asked four groups of Italian children aged 4;7, 5;5, 6;2, and 6;7, respectively, to detect phonemes, syllables, and morphemes within words uttered by some puppets. Children, including the youngest, detected 14% of phonemes and 20% of syllables and morphemes, with a significant increase between each age range. Pinto & Condilera (2000) showed the presence of the rhyme recognition ability (through a multiple-choice test) already in 4-year-old children, with an increase of accuracy in elder children. Martini et al. (2003) administered several metaphonological tasks to 77 children in kindergarten (none of whom had any sort of orthographic skills) divided into 4 groups (age 4;0 – 4;6 – 5;0 – 5;6). They obtained good correctness judgments by every group without significant differences, and high syllabic segmentation scores already in 4-year-old children, with a significant increase from one age to the other. Syllable blending and segmentation turned to be quite spontaneous operations, due to syllables' higher perceptual salience, whereas the initial-syllable recognition task, which probably requires finer skills, revealed some difficulties for younger children. Rhyme

recognition revealed a progressive increase, which was significant between the first two and the second two ages. Recognition of the initial phoneme arose between ages 4;0 – 4;6 and increased significantly between ages 5;0 – 5;6. Phoneme blending and segmentation, as expectable, had lower scores at every age but synthesis seemed to be a bit easier because a significant increase was observed between 4;6 and 5;0, whereas segmentation showed no significant increase. Verbal fluency tasks with syllable/phoneme hint are difficult in this age range, as they require autonomous retrieval abilities. Nevertheless, a significant increase was observed between ages 4;0 – 4;6 and 5;0 – 5;6. Syllable verbal fluency turned out to be easier compared to phonemic verbal fluency. Both Devescovi et al. (1989) and Martini et al. (2003) found that for younger children (4;0 years old), initial syllable recognition is easier than final syllable recognition, but this difference disappears in the following age. Elhassan et al. (2017) found phonemic awareness not to contribute to any reading measure in a group of 38 fluent readers, as compared to 39 moderate and 47 dysfluent readers. They argue that fluency is characterized by a shift from conscious decoding to rapid and accurate visual recognition of words and conclude that phonemic awareness is influential but not sufficient for fluent reading. Cárnio et al. (2017) found a correlation between phonological awareness and reading comprehension of sentences and texts.

1.6.1. Metaphonological skills and atypical reading development: the core phonological deficit hypothesis

One dominant view suggests that developmental dyslexia is caused by a core phonological deficit (Ramus et al., 2013; Snowling & Hulme, 2011; Stanovich & Siegel, 1994; Vellutino et al., 2004). The vast literature concerning this point (for a review, see Snowling, 2000) agrees in considering English-speaking children with reading deficits to be poor in tasks of both phonological discrimination and phonological awareness. This issue, however, is vividly

debated (Blomert & Willems, 2010; Castles & Coltheart, 2004; Harm & Seidenberg, 1999). In particular, the nature of the causal link between phonological awareness and decoding skills is bidirectional, and phonological awareness develops with increased decoding skills, as shown by studies conducted on adults who had or had not learned to read (see Morais et al., 1979). Moreover, it has been recently observed that phonological awareness varies according to the orthographic depth of the respective language (Castles et al., 2003), as it is a more sensitive predictor of reading in languages with deep orthographies.

Vulchanova & Farukh (2018) recently confirmed these points by recurring to the *orthographic depth hypothesis* and the *grain size theory* (Frost, 2006; Frost et al., 1987). According to the *orthographic depth hypothesis*, the opaque relationship between phonemes and graphemes in deep orthographies is tackled by recurring to lexical mediation. The extent of involvement of lexical mediation is determined by the orthographic depth of the language. Before learning to read, words are stored as holistic phonological units. As literacy is acquired, these bigger phonological representations gradually give way to syllable and then phoneme representations and determine a restructuring of the learner's lexicon granularity. According to the *grain size theory*, phonology offers a bigger scale and orthography a smaller scale of granularity, which are represented by phonological units and letters, respectively. Hence, the degree of consistency between phonemes and letters determines the speed of reading development (see Vulchanova & Farukh, 2018).

Developmental language disorders frequently overlap with developmental dyslexia and phonological deficits are frequently found in both disorders (Bishop et al., 2009; Catts et al., 2005; Joanisse et al., 2000; Kamhi & Catts, 1986; McArthur et al., 2000; Ramus et al., 2013). This has been explained through different hypotheses. The *severity hypothesis* (Kamhi & Catts, 1986) suggests that developmental dyslexia and developmental language disorders are two stages of the same disorder depending on the same etiology and differing in terms of severity. According to this model, a phonological processing deficit determines a reading

deficit in less severe cases and an oral language deficit in more severe cases. The *additional deficit hypothesis* (Bishop & Snowling, 2004) suggests that a phonological processing deficit underlies both developmental dyslexia and developmental language disorders, and the addition of a non-phonological disorder to the reading deficit determines the oral language deficit. The *comorbidity hypothesis* (Catts et al., 2005) considers the phonological processing deficit as strictly associated with developmental dyslexia and the developmental language disorder as a different disorder that frequently co-occurs due to comorbidity. Both the *severity* and the *additional hypothesis* have at least two limitations. Both predict that language disorders and dyslexia always co-occur. Moreover, the severity hypothesis assumes that language skills of dyslexics and language impaired children vary on a single dimension, whereas we know that different language skills can be independently impaired (for different types of developmental language disorders, see Friedmann & Novogrodsky, 2008; for different language deficits in developmental dyslexia, see, for instance, Guasti, 2013), The additional hypothesis, in turn, assumes that phonological skills of dyslexic and language impaired children vary along one single dimension.

Ramus et al. (2013) showed that developmental language disorders and developmental dyslexia do not always co-occur, phonological deficits are not always present in developmental language disorders, and different aspects of phonology can be differently impaired. By analyzing four groups of children (with reading deficit, language deficit, both deficits, and typical development), they found three distinct sources of variance for language skills, one for non-phonological language skills, and two for different aspects of phonology, namely phonological skills (including abilities known to be relevant to reading acquisition and to be impaired in dyslexia, namely phonological awareness, verbal working memory, and rapid naming), and phonological representations (including tasks that attempt to tap phonological representations more directly with minimal involvement of additional cognitive

skills (non-word repetition, non-word discrimination, picture-word matching, articulation)¹¹. Children with dyslexia-only appeared to be significantly more impaired in phonological skills than phonological representations, whereas language-impaired children were equally impaired in phonological skills and phonological representations. Children with dyslexia plus language disorders were impaired in phonological skills, phonological representations, and non-phonological language skills, and their deficits were typically more severe than those of children with either language disorders only or dyslexia only¹².

The authors explained these outcomes through a multiple-component model of language abilities distinguishing into pure dyslexia, pure language disorders, and comorbid dyslexia and language disorders (Figure 4).

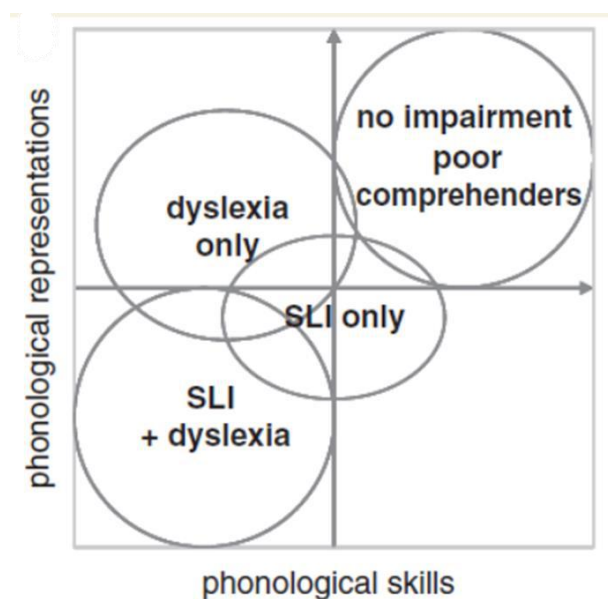


Figure 4. Two-phonological-dimension model of developmental language disorders (here SLI) and developmental dyslexia. Axes correspond to the threshold of $-1.5 SD$ (Ramus et al., 2013)

¹¹ The distinction between phonological skills and phonological representations is taken by Ramus et al. (2013) on the basis of a data-driven factor analysis that identified two different sources of variance for phonological tests. The authors support this distinction by citing several studies (including Hickok & Poeppel's (2007) distinction between dorsal and ventral routes for speech processing). The possibility to tap phonological representations without involving other levels of representation is however arguable, as acknowledged by the authors. Furthermore, including activities like picture-word matching in this category seems to be arguable.

¹² With regard to language tests, children with dyslexia only performed under average in vocabulary but not grammar.

1.6.2. Intervention on metaphonological skills

The interest in phonological awareness has led to the experimentation of several interventions aiming at exploring the possible effects on reading.

Schuele & Boudreau (2008) reviewed the literature to provide speech pathologists with data on the variables to consider in the design and implementation of phonological awareness intervention. Although their study suffers from the bias of specifically referring to the USA educational context¹³, some general parameters can be taken into account. Firstly, to have an impact on reading and spelling, metaphonological activities should be targeted. Secondly, when choosing or designing a phonological awareness intervention program, it is important to consider what is taught in the classroom. Thirdly, letter-sound representations should be introduced in phonological awareness intervention after segmentation and blending skills are established (Spector, 1995).

Concerning the intervention time, they observe that greater change in phonological awareness (i.e., larger effect size) is realized when intervention is provided in preschool or kindergarten rather than in the early primary school (Blachman, 1997; Bus & Van Ijzendoorn, 1999; Troia, 1999).

Concerning results, reporting group effects might neglect that some children show minimal growth, as phonological awareness is a relatively stable skill (Blachman, 1994). With children who fail to attain a sufficient foundation of phonological awareness, it is critical to explore the factors, including the timing of intervention (Blachman, 1994; Torgesen, 2000).

¹³ Willig (1985) warns that many studies from the educational research, especially in the domain of bilingualism, are methodologically compromised.

As for children with cognitive deficits, hearing impairments, or severe communication impairments, the timing of phonological awareness intervention could perhaps be later than for age-matched typically developing peers (Blachman, 1997).

Concerning duration, 5 to 18 hours of intervention provided substantial benefit, with longer programs not necessarily leading to greater benefit. Typical intervention programs (Ball & Blachman, 1991) were spread over 7 to 12 weeks, with 3 to 5 sessions per week, 15 to 30 minutes in length.

Concerning the procedure, there is consensus on the fact that the sequence of phonological awareness development proceeds from rhyme recognition and word segmentation into syllables to the awareness of individual sounds, with the highest level of phonological awareness being the deletion and manipulation of phonemes, as shown in Figure 5.

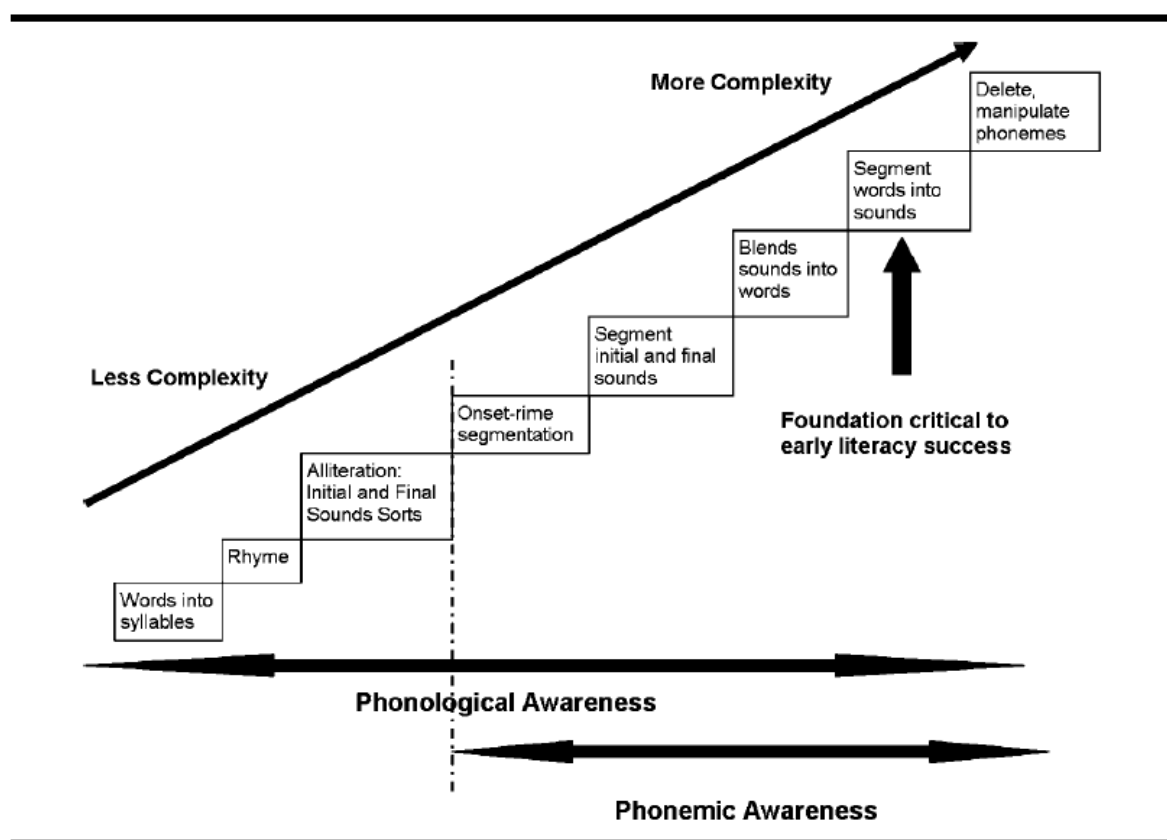


Figure 5. Phonological intervention sequence (Schuele & Boudreau, 2008)

Pape-Neumann et al. (2015) investigated whether a phonological-awareness training, compared to a phonology-based or a visually-based reading training, improved reading in 30 German dyslexic children (age $M = 9.8$) with a phonological awareness deficit in 3rd and 4th grade, who were matched by forming triplets based on IQ, reading quotient and phonological awareness. The phonological awareness training directly improved reading comprehension in dyslexic children with a phonological awareness deficit, who equally benefitted from the visually based reading training. In contrast, the phonology-based reading training had a direct selective effect on decoding but not on reading comprehension. Long-term improvement of reading comprehension and decoding was similar across all training groups, irrespective of the training method. The authors concluded that phonological awareness may but does not need to be part of reading remediation for dyslexic children with a phonological deficit learning to read a consistent orthography and that a visually based reading strategy might compensate for the phonological deficit in dyslexic children after the initial stage of reading acquisition.

Furthermore, a syllable-based reading intervention (Müller et al., 2020) aiming at fostering word reading via high-frequency syllable segmentation improved standard measures of phonological recoding, direct word recognition, and text-based reading comprehension in 79 German poor readers in 2nd grade. A similar intervention (Müller et al., 2017) on German 4th-grade poor readers improved word reading fluency but not reading comprehension. A syllable repeated reading in 2nd and 3rd-grade Finnish poor readers showed an improvement in reading fluency of trained syllables and a transfer to the word level only for longer and less frequent syllables.

Concerning intervention on children with atypical development, Al Otaiba et al. (2009) reviewed the literature on children with language disorders from preschool to the 3rd grade. Besides generally positive outcomes, she found large individual differences in response to

interventions, which showed poor evidence to enable language-impaired children to match their typically developing fellows.

These outcomes could be interpreted in light of the Simple View of Reading (Gough & Tunmer, 1986), according to which reading is the product of decoding and language comprehension. If one of these two abilities is impaired, reading comprehension turns to be also impaired. In this vein, reading comprehension is predictable from a measure of mere decoding (e.g., pseudoword decoding) and a measure of language comprehension (e.g., listening). In this vein, reading disabilities can result in selectively impaired decoding (*dyslexia*), selectively impaired comprehension (*hyperlexia*), or in the combination of both (*garden variety reading disability*).

Socioeconomic status and maternal education have been eventually reported as significantly related to phonological awareness, with wide gaps between lower and middle-income children (Fernandez-Fein & Baker, 1997; Lonigan et al., 1998).

1.7. RAN

Rapid Automatized Naming (RAN) involves a coordinated system of perception, memory, lexis, and articulation (Wolf, 2001), which implement the following processes (Wolf & Denckla 2005: 2):

- 1) attentional processes to the stimulus;
- 2) bi-hemispheric visual processes responsible for initial feature detection, visual discrimination, and pattern identification;
- 3) integration of visual features and pattern information with stored orthographic representations;
- 4) integration of visual and orthographic information with stored phonological representations;

- 5) access and retrieval of phonological labels;
- 6) activation and integration of semantic and conceptual information with all other input; and
- 7) motoric activation leading to articulation.

For about 40 years, a series of studies have suggested that the RAN performance can be considered a reliable predictor of reading fluency, as opposed to phonological measures, which would be better predictors of reading accuracy (for recent data by 8/9-year-old Urdu-speaking children, see Farukh & Vulchanova (2014) for a contrast between RAN and non-word repetition; Vulchanova & Farukh (2018) for a contrast between RAN and phonological awareness). In the beginning, Denckla and Rudel (1976), through a tool used in the treatment of adult dysphasia (Oldfield & Wingfield, 1965), showed that dyslexic children named fewer objects than non-dyslexic children, and, even when correct on words with less than 30-per-million frequency of occurrence, they performed slower than controls, even those suffering from minimal brain dysfunction. However, unlike dysphasic subjects, dyslexic subjects rarely made structural errors. A well-known longitudinal study on 8 dyslexic cases (Wolf & Obregon, 1992) revealed that problems in lexical retrieval persist across primary school and correlate with the onset of a proportional reading disorder. In particular, slowness in naming tasks would be a predictor of difficulties in written-text decoding rather than in written-text comprehension (Wolf & Obregon, 1992), and it would be related more to reading speed than to reading accuracy (Ziegler et al., 2010). Wolf & Bowers (1999), by studying large samples of children with reading disorders in the United States and Canada, found that phonological awareness and RAN contributed separately to reading. Swanson et al. (2003) showed a correlation, albeit weak, between RAN and metaphonological skills ($r = 0.38$). It seems that phonological awareness is important earlier in reading acquisition but, as children reach good decoding accuracy, the relationship between RAN and reading becomes stronger. The orthographic depth of the language seems to determine this

moment, as speakers of languages with shallow orthographies shift away from phonology earlier (Vaessen et al. 2010). Recent studies on RAN performances found that articulation time is not as strongly associated with reading as overall RAN scores (Clarke et al. 2005, Cutting & Denckla 2001, Georgiou et al. 2006, Neuhaus et al. 2001). Instead, it seems that the intermediate pause time may be the RAN component associated with reading. Neuhaus et al. (2001) found that the pause and articulation time were strongly related, and pause time, especially on the letter RAN task, predicted both single-word reading and reading comprehension.

Cross-linguistic studies conducted on 18 languages with both opaque and shallow orthographies confirm similar results as in English, namely that RAN predicts reading both longitudinally and concurrently and in both reading-impaired and typically developing populations (for a review, see Norton & Wolf, 2012). Landerl et al. (2019) conducted a longitudinal study on 1'120 children in 1st and 2nd grade with 5 alphabetic languages (English, French, German, Dutch, and Greek) with different levels of orthographic complexity. RAN was a consistent predictor of reading fluency in all orthographies, whereas the association between phonological awareness and reading was complex and mostly interactive. The authors conclude that RAN has to do with a language-universal cognitive mechanism that is involved in reading alphabetic orthographies irrespective of complexity, whereas the relationship between phonemic awareness and reading depends on many factors like task characteristics, developmental status, and orthographic complexity.

Non-alphabetic languages are also affected by RAN. Concerning Chinese-English bilinguals, those who have difficulty reading both Chinese and English seem significantly slower in naming than peers struggling in just one of their languages (McBride-Chang et al., 2012).

Sullivan et al. (2018) revealed that proficient bilinguals demonstrate slower lexical retrieval in naming tasks than comparable monolinguals. They explain this through the competition

account, according to which slower word retrieval in bilinguals is caused by the need to resolve the conflict between cross-language competitors that are jointly activated by the two languages in the bilingual mind.

Poarch & van Hell (2012) compared children (age 5 – 8) that were, respectively, second language learners of English, (German + English) bilinguals, (German + English + another language) trilinguals, and adult (German + English) bilinguals, in 5 different naming experiments. They revealed that, in bilingual children, lexical processing in the non-dominant language is affected by knowledge of the dominant language, which provides evidence for the coactivation of languages in picture naming in the non-dominant language. Only bilingual and trilingual children with relatively high proficiency in the non-dominant language showed cross-language activation during picture naming in both their non-dominant and dominant languages. Moreover, highly proficient adult bilinguals displayed a significant cognate facilitation effect in both the dominant and non-dominant languages, with a greater effect when naming in the non-dominant language, thus replicating previous results obtained in Spanish–Catalan bilinguals (Costa et al., 2000). The cognate facilitation was not present in monolinguals. The authors explained this by higher activation that cognate words receive from both their dominant and non-dominant language nodes compared with noncognates, which receive activation from only their dominant language node, and by the fact that in bilinguals, the links between the semantic system and the lexical nodes are stronger in the dominant than in the non-dominant language (Kroll & Stewart, 1994).

To our knowledge, only a few experiments involving RAN intervention reported significant effects. Among those that are published (Conrad & Levy, 2011; de Jong & Vrielink, 2004; Fugate, 1997; van der Stappen & van Heybroeck, 2018), the first three trained only letter naming, whereas the last one trained picture RAN besides phonological awareness, confirming the independence between RAN and phonological awareness and the influence

of each measure on different aspects of literacy, namely of RAN on word decoding speed, and of phonological awareness on sublexical spelling.

The experimental section of the present study will use 3 different measures of RAN (pictures, colors, and digits) besides several measures of phonological awareness.

1.8. Preliteracy and literacy acquisition. Is there a bilingual advantage?

Some studies argue that a sort of bilingual executive advantage is reliably associated with several cognitive outcomes, including increased attentional control, working memory, metalinguistic awareness, and abstract and symbolic representation skills (for a review, see Adesope et al., 2010). The issue is, however, highly debated (see Lukasik et al., 2018). The area of inhibitory control is reported to be particularly enhanced in bilinguals (for a review, see Bialystok et al., 2012). Recent neuro-imaging studies suggest increased neuroplasticity in the adult brain exposed to more than one language, with a degree of alteration that is proportional to language experience and depending on different modes of immersive language experience (Kuhl et al., 2016).

By studying fixation times, Weikum et al. (2007) documented a more extended window in bilingual than monolingual infants (8 months vs. 6 months) for the visual discrimination of two different languages. Very early language experiences (before age 3) can have a lasting influence on phonological processing, suggesting that representations established from first-language input persist over time to influence the processing of second-language phonology (Pierce et al., 2015).

An area of particular interest is the metalanguage, which is reported to develop more in bilinguals than in monolinguals (Bialystok, 2001). It is plausible that the necessity to switch between different codes might develop early metalinguistic skills in bilinguals, but when it

comes to explicit metaphonological competence, the issue is more complex. One crucial factor is the type and distance of children's mother tongues, as specifics of metaphonological skills vary according to the more salient phonological aspects of a language. Italian children, for instance, seem to be more adept at syllable and phoneme detection than English ones, and Czech children have higher phoneme awareness but lower onset-rime detection than English children (Bruck & Genesee, 1995). In this regard, Bruck & Genesee (1995) showed that English kindergarteners involved in a French immersion program developed slightly higher onset-rime manipulation skills in English than their English monolingual peers. The difference disappeared in 1st grade but bilinguals had developed higher syllable awareness, whereas monolinguals slightly higher phonemic awareness. The exposure to L2 French, which has higher syllabic salience than English, had determined this difference. Monolinguals' higher phonemic awareness, instead, might be due to their schooling in English reading, which bilinguals had not received. Since both groups were tested in English, this suggests a cross-language transfer of metalinguistic skills. Cross-language transfer of phonological awareness is reported in several experiments (for Spanish - English, see Cisero & Royer, 1995, and Durgunoglu et al., 1993; for Punjabi - English, see Stuart - Smith & Martin, 1997; for Italian - English transfer of phonological skills thanks to a shallower orthography, see Campbell and Sais, 1995; for the same point in Urdu - English, see Mumtaz & Humphreys, 2001).

These data open the question of whether literacy skills besides phonological skills are transferrable between languages. Bialystok et al. (2005) compared 3 groups of bilingual 1st-graders having English as balanced L2 and Cantonese, Hebrew, and Spanish, respectively, as their family languages, with one group of English monolinguals, so that every bilingual group had a different combination of language and writing system. They were tested in decoding and phonological awareness tasks, with bilinguals completing all tasks in both languages. After controlling for random factors, a general increment in reading for all

bilinguals but a larger advantage for children learning two alphabetic systems was found. Moreover, bilinguals transferred literacy skills across languages only when both languages had the same writing system.

Jalali-Moghadan (2015) found better non-word reading as well as better rhyme recognition skills in Iranian bilingual children compared to monolingual ones. The bilingual advantage was more apparent in the first school grades (Kormi-Nouri et al., 2015).

Borragan et al. (2020) showed that bilingual children can be more flexible to acquire a new orthography, provided that their languages are sufficiently distinct from a phono-orthotactical perspective, as in the case of Spanish and Basque, thus suggesting that word learning is not facilitated by bilingualism per se but by the specific combination of languages spoken by the learners.

The amount of same-language input is crucial, as shown by Thordardottir (2015) for the development of morphosyntax in English-French bilingual children. In her study, five-year-old children with low exposure to English displayed optional-infinitive patterns, which are typical clinical markers of developmental language disorders. This might expose them to the risk of misdiagnoses. In these cases, significantly lower scores than monolinguals in both languages might be a reliable sign of possible language impairment.

Bilingual children with a wide range of previous exposure, however, showed a different performance than children with developmental language disorders on non-word repetition and sentence imitation. In particular, they achieved high scores in non-word repetition and were unaffected by non-word length (Thordardottir & Brandeker, 2013).

Interventions on French vocabulary and syntax in monolingual and migrant bilingual children (mean age = 57.7 months; *SD* = 7.8) diagnosed with developmental language disorders in Montreal (Canada) showed efficacy on the targeted skills but short-term gains. Moreover, 3 out of 12 bilingual children obtained scores that were in the non-impaired range in every test

session (including pre-intervention), thus showing the possibility of misdiagnoses or a change in their status (Thordardottir & Rioux, 2019).

A bilingual treatment developed through collaboration with parents, in turn, was deemed ineffective to determine a sufficiently bilingual context to be significantly different from the monolingual treatment (Thordardottir et al., 2015).

Marinelli et al. (2020) found that primary-school children matched for age and school grade are better learners of written words presented in association with pictures when their language (English) has an orthographic inconsistent orthography compared to a consistent one (Italian). This suggests that learning to read in a highly inconsistent orthography enables storing whole-word representations in memory and processing larger orthographic units, whereas learning to read in a consistent orthography promotes grapheme-phoneme conversion rules and processing smaller orthographic units. To support this interpretation, the authors observed that, while spatial-visual memory and phonological awareness explained lexical learning variance in both English and Italian, phonological short-term memory explained further variance in Italian, whereas visuo-attentional skills explained further variance in English.

This partial review shows that bilinguality is a complex variable and its possible advantages (see, e.g., Bialystok et al., 2004) or disadvantages (see, e.g., Michael & Gollan, 2005) can depend on a series of variables including language types, degree of balance between them (Bialystok & Herman, 1999), and biliteracy (Bialystok et al., 2005). Further investigation might reveal whether the (putative) bilingual advantage has linguistic rather than cognitive nature (see, for instance, Borragan et al., 2020), thus revealing some of the “very specific and undetermined circumstances” to which “bilingual advantages in executive functioning [...] are restricted” (Paap et al., 2015).

2. The research project: General Method

The present research is constituted by four studies run from 2018 to 2020.

Study 1 explores the longitudinal development of preliteracy and early literacy skills in a sample of Italian typically developing monolingual and mono-literate bilingual children.

Study 2 analyzes the longitudinal effects of different preliteracy skills on the development of early literacy in the same sample.

Study 3 investigates the effects of metaphonological, RAN, and graphic training on the development of preliteracy and early literacy in a core sample of monolingual and bilingual children with typical development.

Study 4 longitudinally explores the individual, linguistic, and metalinguistic predictors of literacy deficits and possible effects of metaphonological training in a subsample of children who developed some reading or writing deficit during the course of the project.

The general aims are to investigate how diverse preliteracy skills contribute to the development of the first steps of literacy in a language with shallow orthography like Italian; whether this development can be enhanced through a specific training of some preliteracy skills, namely phonological awareness and RAN; and if such training has some effects in case of atypical development. Possible differences between monolingual and bilingual children in the development of these patterns are explored.

This section describes the full sample, the materials used in the studies, and general procedures. Specific methods are described in the dedicated sections of each study.

2.1. Participants

A total number of 172 children took part in the project. They attended a Comprehensive Institute¹⁴ in a neighborhood of Rome, which has a heterogeneous population of more than 56'000 inhabitants (population density = 19'000 per km²), with the predominance of a medium-low socio-economic status¹⁵. Ninety-six were Italian monolingual children, and seventy-six were bilingual children coming from migrant families with different mother languages (Figure 6), who had attended at least two years of kindergarten in the hosting Institute.

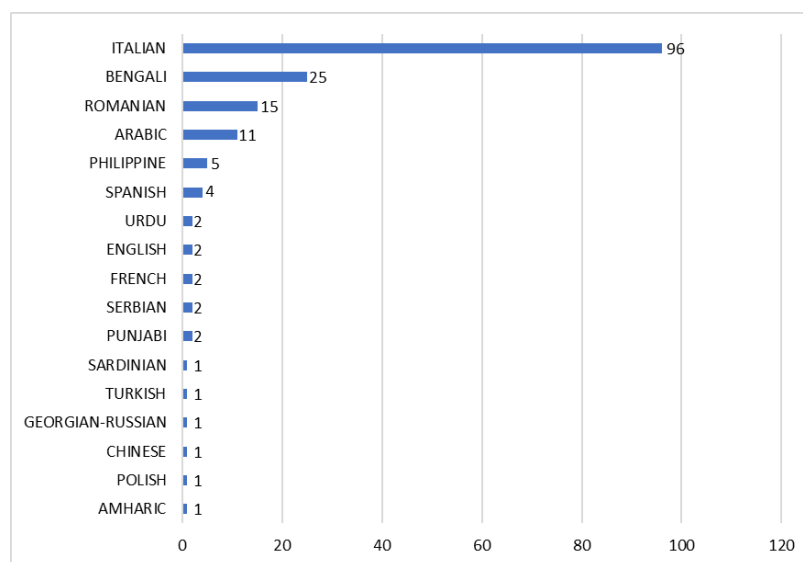


Figure 6. Mother languages of the full sample

2.2. Materials and general procedures

Forty children were longitudinally monitored from the last year of kindergarten to the 2nd grade of primary school; 50 children from the last year of kindergarten to the 1st grade of primary school; 9 children from the 1st to the 3rd grade of primary school. The remaining 73 children were used as controls for collective tests of dictation and reading comprehension.

¹⁴ The Institute includes kindergarten (3 years), primary school (5 years), and lower secondary school (3 years).

¹⁵ Data from the Registry Office of Roma Capitale (2014).

The project was implemented in the framework of experimental monitoring of children's preliteracy and literacy skills promoted by the hosting Institute to early detect possible learning difficulties. Informed consent to the monitoring of children's learning skills is signed by parents upon school enrollment. Moreover, parents provided informed parental consent for children's participation in the present research project. Procedures followed the ethical principles of the Declaration of Helsinki (October 2013).

Children's willingness to take part in the program was ascertained through dedicated meetings. Regular consultations with families, teachers, and school principals took place in the presence of the school counselor. General information on participants was collected through interviews with teachers and parents. The families of monitored children filled in a questionnaire, which consists of 21 questions exploring children's social environment as well as possible familiarity with language disorders or learning disorders¹⁶. Moreover, kindergarten teachers filled in the sections of language, metacognition, pre-literacy, and "other abilities" of the observational questionnaire for the early identification of learning difficulties (Terreni et al., 2011). These sections consist of 29 questions exploring children's behavior during school activities through a frequency Likert scale (never / sometimes / often / always).

Children were tested in the middle and at the end of each school year. Tests were administered by the author and some selected collaborators¹⁷, who received specific training, in silent and adequately lit rooms of school facilities. Tests included language (phonology, receptive and productive vocabulary, syntactic comprehension and production); metalanguage (syllabic and phonological awareness); rapid automatized naming (RAN) of

¹⁶ The questionnaire collects personal information about the child (anagraphic information, health, attended kindergarten) and the family context (components , spoken language(s), parents' occupation and education , familiarity with learning disabilities or developmental disorders), to control for possible risk factors for language or learning disorders

¹⁷ Collaborators were selected and trained by the author after being introduced by the school. They were young graduated fellows (psychologists or speech therapists) running an internship or a collaboration program in the hosting Institute.

pictures, colors, and digits; short-term memory (phonological and verbal); and visual skills (symbol reproduction and capital-letter visual discrimination).

2.3. Language skills

Phonological skills were tested through a non-word repetition test, which was extracted from the PRCR-2 battery (Cornoldi et al., 2009), in the middle and at the end of the last year of kindergarten and in the middle of the 1st grade of primary school. This test consists of a series of 5 monosyllables of increasing complexity, a series of 5 disyllables of 5 letters, a series of 5 disyllables of 6 letters, a series of 5 trisyllables (3 paroxytones + 2 proparoxytones), and a series of 5 quadrisyllables (2 proparoxytones + 3 paroxytones).

Children should repeat the non-word uttered by the administrator. The strings meet the phonotactic rules of Italian, so they are pseudowords (Gathercole, 2006; Santos et al., 2006). Their repetition taps a series of competencies including phonological decoding of the auditory string, phonological short-term memory to retain the auditory material, phonological awareness to match the string to the phonological system of the respective language, phonological encoding, and articulation (see, for instance, Farukh & Vulchanova, 2014; Gathercole, 2006). For this reason, the task has been deemed as a language measure of phonological competence rather than mere phonological short-term memory.

The receptive vocabulary was tested through the Italian version of the PPVT (Peabody Picture Vocabulary Test) (Dunn et al., 2007) once per year, in the last year of kindergarten, in the 1st, 2nd, and 3rd grade of primary school. The test is a word-picture matching task, where the child must choose one option out of four that represents the word uttered by the administrator. Accuracy raw scores were analyzed.

The productive vocabulary was tested in the 1st, 2nd, and 3rd grades of primary school through the Italian version of the BNT (Boston Naming Test) (Kaplan & Goodglass, 1983).

Children should name a series of colored pictures displayed on a computer screen. Accuracy raw scores (correct items) were analyzed.

Syntactic comprehension was tested once per year in the last year of kindergarten, in the 1st, 2nd, and 3rd grade of primary school, through the Italian version of the TROG-2 (Bishop, 2009). It is a sentence-picture matching task that investigates the production of 20 different syntactic structures. Children must choose one picture out of four that represents the sentence uttered by the administrator. Accuracy raw scores were analyzed.

Syntactic production was tested twice in the last year of kindergarten, and once per year in the 1st, 2nd, and 3rd grade of primary school through an elicitation test of Direct Object clitic pronouns (DOcl). In this test, the administrator shows some pictures to children, describes the situation and then asks a question that should force the production of a sentence containing a clitic pronoun instead of a full DP. The original test (Arosio et al., 2014) elicited 12 singular DOcl (6 masculine + 6 feminine) under the condition of gender matching with the sentential subject. The difficulty of this test was increased across progressive test sessions by adding conditions of gender and number mismatch between the DOcl and the sentential subject. A version of intermediate difficulty (Casani & Cardinaletti, 2019) that elicited 15 singular DOcls (8 masculine + 7 feminine) and included 7 conditions of gender mismatch (5 MF + 2 FM) with the sentential subject was employed at the end of kindergarten and, with a small number of children, in the 1st grade of primary school. The last version (Casani et al., in preparation) elicits 48 DOcls under balanced conditions of gender and number mismatch (6 MS-MS; 6 MS-FS; 6 FS-FS; 6 FS-MS; 6 MP-MP; 6 MP-FP; 6 MP-MS; 6 MP-FS). A pilot version of this test was employed with the majority of children in the 1st grade, and with all children in the 2nd and 3rd grade of primary school. Proportional accuracy scores were analyzed.

Tests that required the use of pictures were administered on a 16-inch laptop screen.

2.4. Metaphonological skills

Metaphonological skills were tested twice in kindergarten at the global level (syllable awareness), and five times in primary school at the analytical level (phonological awareness) through the CMF battery (Marotta et al., 2008). Global tests included syllable blending and segmentation, recognition of word and non-word minimal pairs, and recognition of the initial syllable. Analytical tests included phoneme blending and segmentation, deletion of the initial and final syllable, and phonemic verbal fluency. The recognition of rhymes was tested 7 times, twice per year, both in kindergarten and in primary school. Starting from the 2nd grade, spoonerism was also tested. This is a manipulation test requiring to mentally switch corresponding syllables between two words to create a different couple of meaningful words. Every test of this battery consists of 15 items. Raw scores were computed.

2.5. Short-term memory

Short-term/working memory was tested four times, once per year, through digit span forward, digit span backward (Bisiacchi et al., 2005), and word span (Spinnler & Tognoni, 1987). Children should repeat increasing series of digits forward and backward and of disyllabic words. Furthermore, the non-word repetition test (see above) gave a measure of phonological short-term memory.

Short-term/working memory was tested once per year. Raw scores were analyzed.

2.6.. RAN

RAN of pictures was tested in the middle and at the end of the last year of kindergarten through the RAN test of the IPDA materials (Terreni et al., 2002). This consists of 50 black-

and-white printed pictures (5 objects that are randomly displayed 10 times), which must be named as fast and accurately as possible from left to right and from top to bottom as if reading. This test is standardized on 5-year-old children attending the last year of kindergarten and includes only pictures.

In the middle and at the end of the 1st and 2nd grade and in the middle of the 3rd grade, a more comprehensive test (De Luca et al., 2005) was used, which is standardized on primary-school children and includes the RAN of pictures, colors, and numbers. This consists of double lists of items ($N = 50 + 50$) that are presented with the same procedure. RAN speed (object per second) and error percentages were analyzed.

2.7. Visual skills

Visual skills (Cornoldi et al., 2009) were tested through the graphic reproduction of series of one / two / three semicircles (10 items for each series) after brief observation, and through a multiple-choice test consisting of the visual discrimination of the position of 12 printed capital letters. Tests of reproduction of one and two semicircles and visual discrimination of letter position were administered in the middle and at the end of the last year of kindergarten, in the middle and at the end of the 1st grade, and in the middle of the 2nd grade of primary school. The test reproduction of three semicircles was administered in the middle of the 1st grade and in the middle of the 2nd grade of primary school. Raw error scores were analyzed.

2.8. Word decoding

Word decoding was tested through the TILTAN-I battery (Traficante et al., 2017). The word list consists of 125 words (51 disyllabic + 59 trisyllabic + 15 tetra-syllabic words) that are read in a column as fast as possible; a list of 40 non-paroxytone words that are read in a

column to explore lexical decoding; and a list of 25 word pairs (50 words) that are read by row to explore decoding attentional errors.

Decoding speed (syll/sec) and decoding error proportions were analyzed.

Word decoding was tested in the middle of the 2nd grade, at the end of the 2nd grade, and in the middle of the 3rd grade of primary school.

2.9. Sublexical decoding

Sublexical decoding was tested in the middle of the 1st grade of primary school, through a list of alphabetic strings (Calgaro et al., 2018) constituting sublexical components of Italian words (44 monosyllabic + 6 bisyllabic items, among which 26 consisting of 2 letters, 23 of 3 letters, and 1 of 4 letters. Twenty-three are existent function strings (prepositions, articles, and pronouns), 16 of which with high-frequency¹⁸. The items alphabetic structure is CV (N = 25), CVC (N = 11), CCV (N = 6), CVV (3 item), VCV (3 item), VC (1 item)¹⁹, and VCCV (N = 1)

The other sub-lexical tests were administered in the middle and at the end of the 2nd grade, and in the middle of the 3rd grade of primary school through the non-word list of the TILTAN-I battery (Traficante et al., 2017). This is a list of 40 pseudowords²⁰ for Italian (25 disyllabic + 15 trisyllabic strings) conceived to diagnose phonological dyslexia. Decoding speed (syll. / sec.) and error proportions were analyzed.

¹⁸ Data by the CoLFIS data bank, Scuola Normale Superiore of Pisa.

¹⁹ Calgaro et al. (2018) erroneously report the structure CCV and VC to correspond, respectively, to 5 and 2 items.

²⁰ See 2.3 for more information about nonwords / pseudowords.

2.10. Text decoding

Text decoding was tested four times, at the end of the 1st grade, in the middle and at the end of the 2nd grade, and in the middle of the 3rd grade of primary school, through the MT-3 clinical tests (Cornoldi & Carretti, 2016). Length and difficulty are calibrated on the grade and moment of the year (beginning, end). Decoding speed (syll/sec) and error proportion were analyzed.

2.11. Reading comprehension

In the full sample, reading comprehension was tested at the end of the 1st grade, in the middle of the 2nd grade, and in the middle of the 3rd grade through the MT-2 battery (Cornoldi & Colpo, 2011). Children were asked to read a narrative text silently and answer 10 three-choice questions. Texts' length and complexity are calibrated on the grade and moment of the year (beginning / intermediate / end). This test was administered collectively in children's classrooms. A brief familiarization test was administered before the experimental test in 1st and 2nd grade. Children were given all the time needed to complete the task. The proportion of correct responses was analyzed.

An additional diagnostic reading comprehension test (Cornoldi & Carretti, 2016) was administered to the experimental sample. Administrations took place at the end of the 1st grade, in the middle and at the end of the 2nd grade, and in the middle of the 3rd grade. Children were requested to read silently two texts calibrated on the grade and moment of the year and answer 10/12 four-choice questions for each text. The proportion of correct responses was analyzed. These tests were administered individually, as the protocol requires. In any case, no time limit was imposed. The MT-3 clinical tests allow qualitative analysis of responses to questions, which can be specific or central. Specific questions have a literal relationship with the text, where the answers can be directly retrieved, whereas

central questions require inferential operations, with connections with other information and, consequentially, a heavier load of memory. Specific, central, and global (specific + central) reading comprehension were separately analyzed.

2.12. Word writing

Word writing was tested in the middle and at the end of the 1st grade through word dictation (Stella & Apolito, 2004). In this test, children should write 16 words (4 simple disyllables + 4 complex disyllables + 4 simple trisyllables + 4 complex trisyllables) in a limited time (7 seconds per word). Words were uttered once by the administrator. This test was administered collectively.

In the middle and at the end of the 2nd grade and in the middle of the 3rd grade, word writing was tested through the DDE dictation (Sartori et al., 2007). This is part of a diagnostical battery for dyslexia and dysorthography and consists of 48 words (9 simple disyllables + 18 complex disyllables + 19 complex trisyllables + 1 simple tetrasyllable + 1 complex tetrasyllable) that are dictated individually without time limits. Before writing, the child is requested to repeat the word correctly. Error proportions were analyzed.

2.13. Non-word writing

Non-word writing was tested in the middle and at the end of the 2nd grade and in the middle of the 3rd grade of primary school through the non-word dictation of the DDE battery (Sartori et al., 2007). This consists of 24 pseudowords²¹ for Italian (2 simple disyllables + 10 complex

²¹ See 2.3 for more information about nonwords / pseudowords.

disyllables + 3 simple trisyllables + 5 complex disyllables + 1 simple tetrasyllable + 3 complex tetrasyllables). Error proportions were analyzed.

2.14. Text writing

Text writing was tested in the middle of the 2nd and 3rd grade through the BVSCO-2 (Tressoldi et al., 2013) dictation test. A text calibrated on children's grades was dictated collectively. The test was administered once per grade so that longitudinal analyses were not possible. Error proportions were computed (errors/words), and error analysis was conducted.

Error coding includes

- phonological errors, namely sound change (e.g., fiume → piume); syllable (e.g., matita → tita) or sound omission (e.g., carta → cata); methateses (e.g. cinema → cimena); epentheses (e.g. melone → melolone);
- non-phonological errors, namely errors in non-homograph homophones (e.g. cuore → quore); illegal blending (e.g., l'ago → lago); and illegal segmentation (e.g., insomma → in somma);
- stress/geminate errors (e.g., perché → perche; papa → papà; carro → caro; nano → nanno;
- minor errors, namely capital letter errors (e.g. Roma → roma), and punctuation errors;
- word omissions;
- inappropriate words (e.g. viscere → strisce)

3. Study 1

Development of pre-literacy and literacy in typically developing monolingual and mono-literate bilingual children in an Italian school setting

3.1. Aims 1

This study explores the longitudinal development of preliteracy and early literacy skills in a sample of typically developing monolingual children compared to mono-literate bilingual children with minority languages in an Italian school setting.

We concentrate on the differences between monolinguals and bilinguals to investigate whether language, metalanguage, and working memory pre-requisites as well as the first steps of literacy are equally acquired or follow different patterns.

3.2. Method 1

3.2.1. Participants

A total number of 145 children (75 monolinguals + 70 bilinguals) participated in this study. Fourteen participants (6 monolinguals + 8 bilinguals) were excluded from the analyses because they were diagnosed with Language Impairment (henceforth LI²²). Two participants

²² In this study, the acronym LI (Language Impairment), which is not updated (Bishop et al., 2016; 2017), is preferred for Language Disorders to distinguish it from the acronym LD, which is here employed for Learning Disorders (Hammill, 1990; Tzouriadou, 2020). This choice is coherent with the terminology used in the diagnoses, which were issued by the Italian public health system (ASL), and reported the Italian equivalent for Specific Language Impairment, or Language Impairment associated to other conditions of atypical development.

(one monolingual and one bilingual) with LI also had an intellectual disability. One child was excluded for being diagnosed with Autism Spectrum Disorder (ASD)²³.

The final sample included 130 typically developing children (68 monolinguals + 62 bilinguals) who were followed from the last year of kindergarten to the 2nd grade of primary school. Table 1 displays the number of monolingual and bilingual participants in each testing session.

3.2.2. Materials

The materials consisted of two observational questionnaires to be filled out by the teachers and one questionnaire for families:

- the observational questionnaire for the early identification of learning difficulties (IPDA, Terreni et al., 2011), limited to the sections of language, metacognition, pre-literacy, and “other abilities”;
- a brief questionnaire on the types of activities usually made in class.

In addition, children were assessed on different tasks to examine their linguistic and non-linguistic abilities.

The families filled in a questionnaire about the family environment and possible familiarity with LI or LDs (Casani, n/a).

Children were tested in the following abilities:

- forward/backward digit span (Bisiacchi et al., 2005);
- word repetition (Spinnler & Tognoni, 1987);
- non-word repetition (Cornoldi et al., 2019);
- metaphonology (Marotta et al., 2008);

²³ These children followed the scheduled procedures, as possible, but their cases were excluded from the analyses.

- RAN of pictures (De Luca et al., 2005; Terreni et al., 2002), colors, and digits (De Luca et al., 2005) ;
- graphic reproduction of series of one/two semicircles oriented in different fashions (Cornoldi et al., 2019);
- visuo-spatial recognition of the orientation of capital letters (Cornoldi et al., 2019);
- syntactic comprehension (Bishop, 2009);
- syntactic production of clitic pronouns (Arosio et al., 2014; Casani & Cardinaletti, 2019; Casani et al., in preparation);
- sublexical decoding (Calgaro et al., 2018; Traficante et al., 2017);
- word and non-word decoding (Traficante et al., 2017);
- text decoding (Cornoldi & Carretti, 2017);
- reading comprehension (Cornoldi & Carretti, 2016; Cornoldi & Colpo, 2011);
- word dictation (Stella & Apolito, 2004; Sartori et al. 2017);
- text dictation (Tressoldi et al., 2013).

For detailed information about the tests, see sections 2.2 – 2.14.

3.2.3. Procedures

Testing took place over five sessions, in the middle (henceforth, KG-1) and at the end (henceforth, KG-2) of kindergarten, in the middle (henceforth, 1-1) and at the end (henceforth, 1-2) of the 1st grade, and in the middle (henceforth, 2-1) of the 2nd grade of primary school.

A core sample of 29 children (16 monolinguals + 13 bilinguals)²⁴ participated in the first screening in the middle of the last year of kindergarten (KG-1) and was monitored until the middle of the 2nd grade of primary school (2-1). Besides digit span forward, digit span backward, and word span, language measures of receptive vocabulary (raw score), syntactic comprehension (raw score), syntactic production of direct object clitic pronouns (percent correct), and non-word repetition (number of correct syllables) were tested. As for metaphonological skills (raw scores), syllable blending, recognition of word and non-word minimal pairs, recognition of rhymes, recognition of the 1st syllable, and syllabic segmentation were tested. Furthermore, speed (obj./sec.) and error proportional scores in the RAN of pictures (since KG-1), colors, and numbers (since primary school), graphic reproduction of one and two semicircles oriented in different positions (error score), and visual recognition of single capital letters (error score) were tested.

An additional group of 50 children (24 monolinguals + 26 bilinguals) was added in sessions KG-2 and 1-1. Their memory span, metaphonological skills, and RAN were tested, and the same measures were retested in other participants. In 1-1, these abilities plus receptive vocabulary, productive vocabulary, and syntactic production were tested again. Moreover, sublexical decoding (in 1-1 and 2-1), word decoding (in 1-2 and 2-1), text decoding (in 1-2 and 2-1), word writing under dictation (in 1-1, 1-2, and 2-1), non-word writing under dictation (in 2-1), text dictation (in 2-1), and reading comprehension (in 1-2 and 2-1) were tested.

Tests of word-writing in 1-1 and 1-2, and text writing in 2-1 included a further group (henceforth, CG) of 52 children (28 monolinguals + 24 bilinguals). Table 1 displays the number of monolingual and bilingual participants and the abilities tested in each test session. Scoring procedures are described in detail in sections 2.2 – 2.14.

²⁴ The original core sample of 37 children was reduced to 29 after defections and after excluding children with atypical development.

	KG-1	KG-2	1-1	1-2	2-1
forward/backward digit span (Bisiacchi et al., 2005)		72 (39 MONO + 33 BI)		73 (40 MONO + 33 BI)	28 (15 MONO + 13 BI)
word span (Spinnler & Tognoni, 1987)		76 (40 MONO + 36 BI)		73 (40 MONO + 33 BI)	28 (15 MONO + 13 BI)
non-word repetition (Cornoldi et al., 2019)	29 (16 MONO + 13 BI)	76 (40 MONO + 36 BI)		75 (40 MONO + 35 BI)	
metaphonology (syllable awareness) (Marotta et al., 2008)	29 (16 MONO + 13 BI)	75 (40 MONO + 35 BI)			
metaphonology (phonemic awareness) (Marotta et al., 2008)			73 (40 MONO + 33 BI)	27 (16 MONO + 11 BI)	28 (15 MONO + 13 BI)
RAN of pictures (Terreni et al., 2002)	28 (16 MONO + 12 BI)	73 (38 MONO + 35 BI)			
RAN of colors, pictures, and digits (De Luca et al., 2005)			74 (40 MONO + 34 BI)	28 (17 MONO + 11 BI)	27 (15 MONO + 12 BI)
graphic reproduction of series of 1 and 2 semicircles oriented in different fashions (Cornoldi et al., 2019)	29 (16 MONO + 13 BI)	31 (17 MONO + 14 BI)	31 (17 MONO + 14 BI)		
graphic reproduction of series of 3 semicircles oriented in different fashions (Cornoldi et al., 2019)			31 (17 MONO + 14 BI)		
visuo-spatial recognition of capital letters (Cornoldi et al., 2019)	29 (16 MONO + 13 BI)	31 (17 MONO + 14 BI)			
syntactic comprehension (Bishop, 2009)	32 (17 MONO + 15 BI)				27 (15 MONO + 13 BI)
syntactic production of direct object clitic pronouns (Arosio et al., 2014)	30 (16 MONO + 14 BI)				
syntactic production of direct object clitic pronouns (Casani & Cardinaletti, 2019)		31 (17 MONO + 14 BI)			
syntactic production of direct object clitic pronouns (Casani et al., in preparation)					27 (15 MONO + 13 BI)
receptive vocabulary (Dunn et al., 2000)	31 (17 MONO + 14 BI)		47 (23 MONO + 24 BI)		27 (15 MONO + 13 BI)
productive vocabulary (Kaplan et al., 1983)			47 (23 MONO + 24 BI)		27 (15 MONO + 13 BI)
sublexical decoding (Calgaro et al., 2018);			72 (40 MONO + 32 BI)		
sublexical (non-word) decoding (Traficante et al., 2017);					27 (15 MONO + 12 BI)
word decoding (Traficante et al., 2017)					27 (15 MONO + 12 BI)
text decoding (Cornoldi & Carretti, 2017)				28 (17 MONO + 11 BI)	27 (15 MONO + 12 BI)
reading comprehension (Cornoldi & Carretti, 2016)				28 (17 MONO + 11 BI)	28 (15 MONO + 13 BI)
reading comprehension (Cornoldi & Colpo, 2011)				58 (31 MONO + 27 BI)	61 (32 MONO + 29 BI)
word dictation (Stella & Apolito, 2004)			121 (66 MONO + 55 BI)	59 (33 MONO + 26 BI)	
word/non-word dictation (Sartori et al. 2017)					27 (15 MONO + 12 BI)
text dictation (Tressoldi et al., 2013)					61 (32 MONO + 29 BI)

Table 1. Number of monolingual and bilingual subjects and abilities tested in each test session.

3.2.4. Analyses

Since the tests are standardized only on the monolingual population, longitudinal raw scores were analyzed for the abilities tested once (productive vocabulary) and retested through the same tests (digit span, word repetition, non-word repetition, graphic-visual skills, receptive vocabulary, syntactic comprehension, and metaphonological skills). Percentage scores were analyzed for the abilities retested through different tests (syntactic production of direct object clitic pronouns, RAN, reading, and writing).

A series of Generalized (Linear) Mixed Models for Repeated Measures (GMMRM) for longitudinal analyses and Generalized (Linear) Mixed Models (GMM) for cross-sectional analyses were performed in IBM-SPSS, with test scores as dependent variables, bilingualism and administration time (for repeated measures) as fixed effects. Since the core sample that attended the first screening underwent a training program between test sessions KG-1 and KG-2 (for the effects, see Study 3) and between test sessions 1-1 and 1-2, the variable with/without intervention was entered as a random effect. Satterthwaite approximation was applied.

3.3. Results 1

The following section reports the results with a focus on the differences between monolingual and bilingual children. Line graphs display the estimated means (95% *Conf. Int.*) of the two groups for longitudinal administrations, whereas boxplots show the score distribution for cross-sectional tests. Statistics of predictive effects are reported in the following tables. Complete tables including main effects are in Appendix A.

3.3.1. Language skills

Figure 7 shows longitudinal scores obtained by monolingual and bilingual children in the non-word repetition test, receptive vocabulary, syntactic comprehension, and syntactic production of direct object clitics (Arosio et al., 2014, in administration KG-1; Casani & Cardinaletti, 2019, in administration KG-2; and Casani et al., in preparation, in administration 2-1)²⁵.

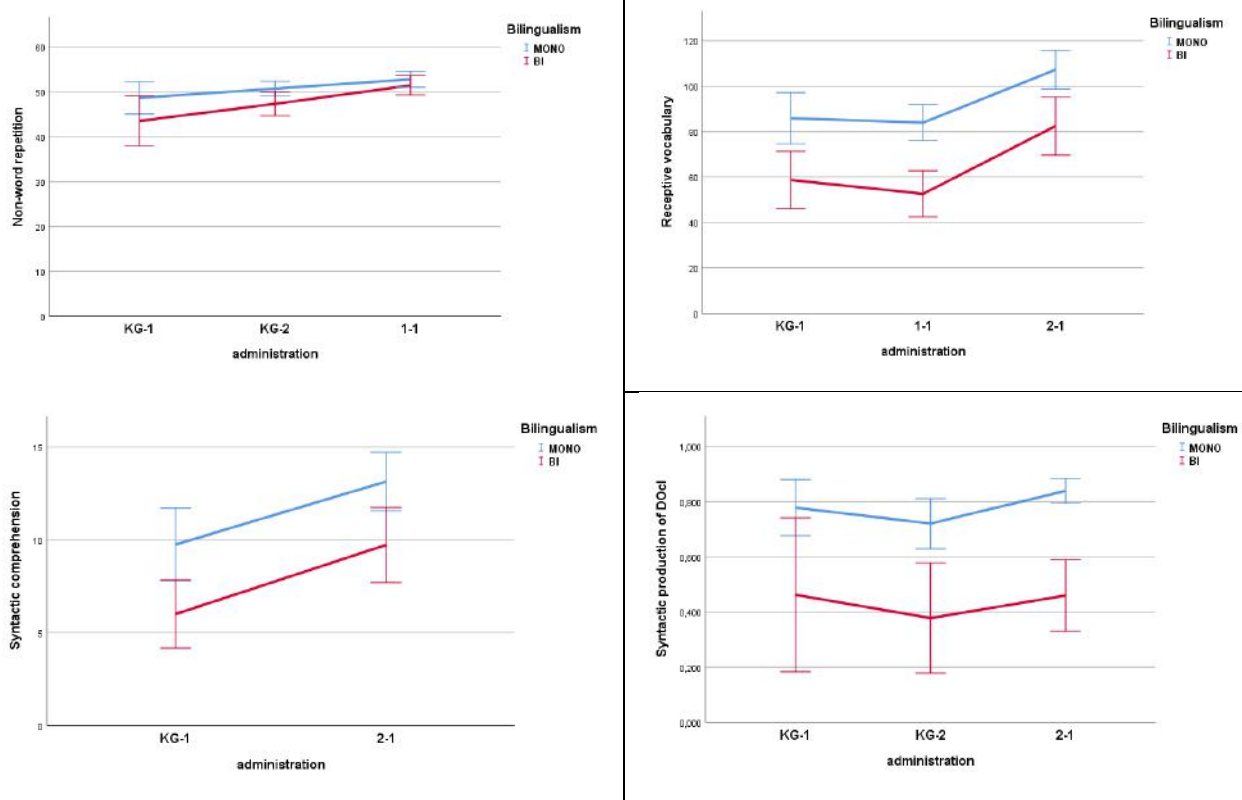


Figure 7. Longitudinal means of monolingual and bilingual children in non-word repetition, receptive vocabulary, syntactic-comprehension, and syntactic-production.

A series of GMMRM with test scores as dependent variables, and bilingualism and administration time as fixed effects correctly predict non-word repetition ($F(3, 62) = 11.363$,

²⁵ In longitudinal tests of syntactic production of direct object clitic pronouns, 3 different versions of increasing difficulty of the clitic test were used (for details, see 2.3).

$p = 0.000$), receptive vocabulary ($F(2, 49) = 19.553, p = 0.000$), syntactic comprehension ($F(2, 52) = 15.220, p = 0.000$), and syntactic production ($F(3, 37) = 4.573, p = 0.008$). We found main effects of bilingualism on every ability except non-word repetition (where the effect was marginally significant), and of administration time on every ability except syntactic production. There are positive effects of monolingualism on every ability (except for the marginally non-significant effect on non-word repetition), negative effects of administrations KG-1 and KG-2 on non-word repetition, and of KG-1 on receptive vocabulary and syntactic comprehension.

Post-hoc pairwise comparison of the effects of administration time showed a significant improvement of non-word repetition between KG-2 and 1-1 ($F(2, 72) = 16.627, p = 0.000$), whereas no significant difference emerged between the two administrations in KG, and a significant increase of receptive vocabulary between KG-1 and 2-1. Significant contrasts are in Table 2.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p (seq. Bonferroni)	Lower	Upper
Non-word repetition	KG-1	1-1	- 5.811	1.576	- 3.687	40	0.001	- 9.840	- 2.142
	KG-2	1-1	- 3.032	0.614	- 4.396	167	0.000	- 4.157	- 1.546
Receptive vocabulary	KG-1	2-1	- 22.689	5.377	4.219	55	0.000	11.912	33.465

Table 2. Significant post hoc comparison between the effects of administration time on non-word repetition and receptive vocabulary.

The combination between bilingualism and administration time as a fixed effect predicts significant differences between monolinguals and bilinguals in receptive vocabulary and syntactic comprehension in every administration, as shown in Table 3

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Receptive vocabulary	MONO*KG-1	BI*KG-1	27.161	7.833	3.468	28	0.002	11.116	43.205
	MONO*1-1	BI*1-1	31.373	6.255	5.015	45	0.000	18.774	43.972
	MONO*2-1	BI*2-1	24.805	6.919	3.585	26	0.001	10.583	32.027
Syntactic comprehension	MONO*KG-1	BI*KG-1	4.083	1.295	3.154	29	0.004	1.435	6.731
	MONO*2-1	BI*2-1	3.364	1.170	2.874	26	0.008	0.958	5.770

Table 3. Significant differences predicted by the combination of bilingualism and administration time in receptive vocabulary and syntactic comprehension.

Figure 8 displays cross-sectional scores of productive vocabulary and syntactic production of direct object clitics²⁶ in the 1st and 2nd grades.

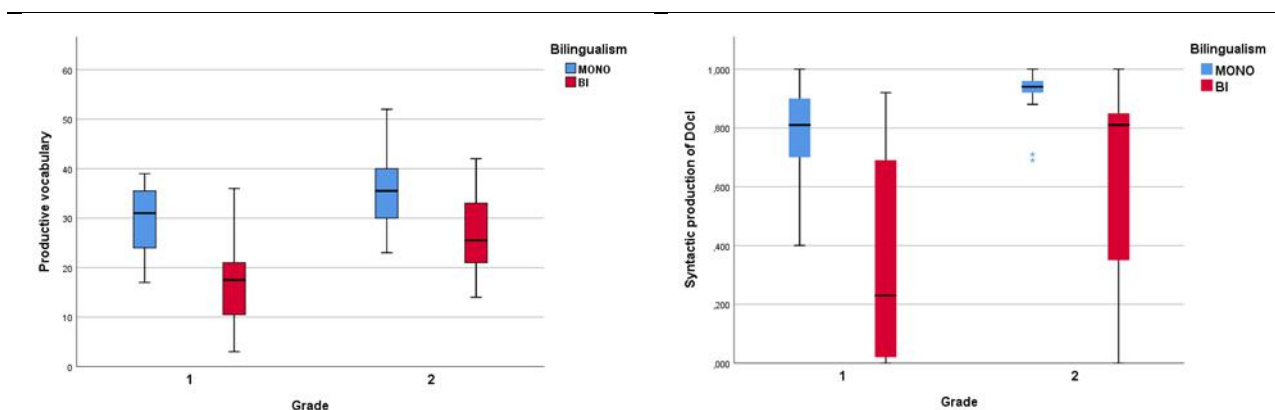


Figure 8. Distribution of the cross-sectional scores of productive vocabulary and syntactic production of direct object clitic pronouns (Casani et al., in preparation) in 1st and 2nd grade.

A GMM with productive vocabulary and syntactic production scores as dependent variables, and bilingualism and grade as fixed effects correctly predicts productive vocabulary ($F(2, 74) = 31.174$, $p = 0.000$) and syntactic production of direct object clitic pronouns ($F(2, 63) = 7.188$, $p = 0.002$). Main effects of bilingualism and grade were found. Productive vocabulary of bilingual children ($M = 21.172$, $SE = 1.302$, $18.577 < CI < 23.766$) is significantly smaller than that of monolinguals ($M = 32.625$, $SE = 1.274$, $30.087 < CI$

²⁶ In cross-sectional tests of syntactic production, the last version of the clitic test was used (Casani et al., in preparation) including balanced (morpho)syntactic elicitation conditions (see 2.3).

<35.163), and that of 1st graders ($M = 22.845$, $SE = 1.150$, $20.553 < CI < 25.138$) is significantly smaller than that of 2nd graders ($M = 30.952$, $SE = 1.441$, $28.080 < CI < 33.823$). Monolingual children produced significantly more target clitics than bilinguals ($M = 0.862$, $SE = 0.083$, $0.711 < CI < 1.045$, vs. $M = 0.523$, $SE = 0.057$, $0.420 < CI < 0.651$). The difference between 1st and 2nd graders in clitic production, tested with the most complex version of the test (Casani et al., 2020), was not significant.

Estimates of significant differences are in Table 4.

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Syntactic production of DOcl	MONO	BI	0.339	0.101	3.372	63	0.001	0.138	0.540
Productive vocabulary	MONO	BI	11.453	1.799	6.366	74	0.000	7.868	15.038
	1	2	- 8.106	1.845	- 4.395	74	0.000	- 11.782	- 4.431

Table 4. Significant differences between mono- and bilinguals in syntactic production of direct object clitics and productive vocabulary, and between 1st and 2nd grade in productive vocabulary.

In sum, language scores of monolingual children were significantly better than those of bilinguals in every ability except phonology. Children's phonological skills improved significantly between the end of kindergarten and the middle of the 1st grade, whereas receptive vocabulary only in 2nd grade. Syntactic production of direct object clitic pronouns including conditions of gender and number mismatch with the sentential subject did not improve significantly from one year to the other one.

3.3.2. Metaphonological skills

Global metaphonological skills

Figure 9 displays longitudinal scores obtained by monolinguals and bilinguals in the two kindergarten administrations of the syllable-awareness tests, and in the five administrations of the rhyme-recognition test.

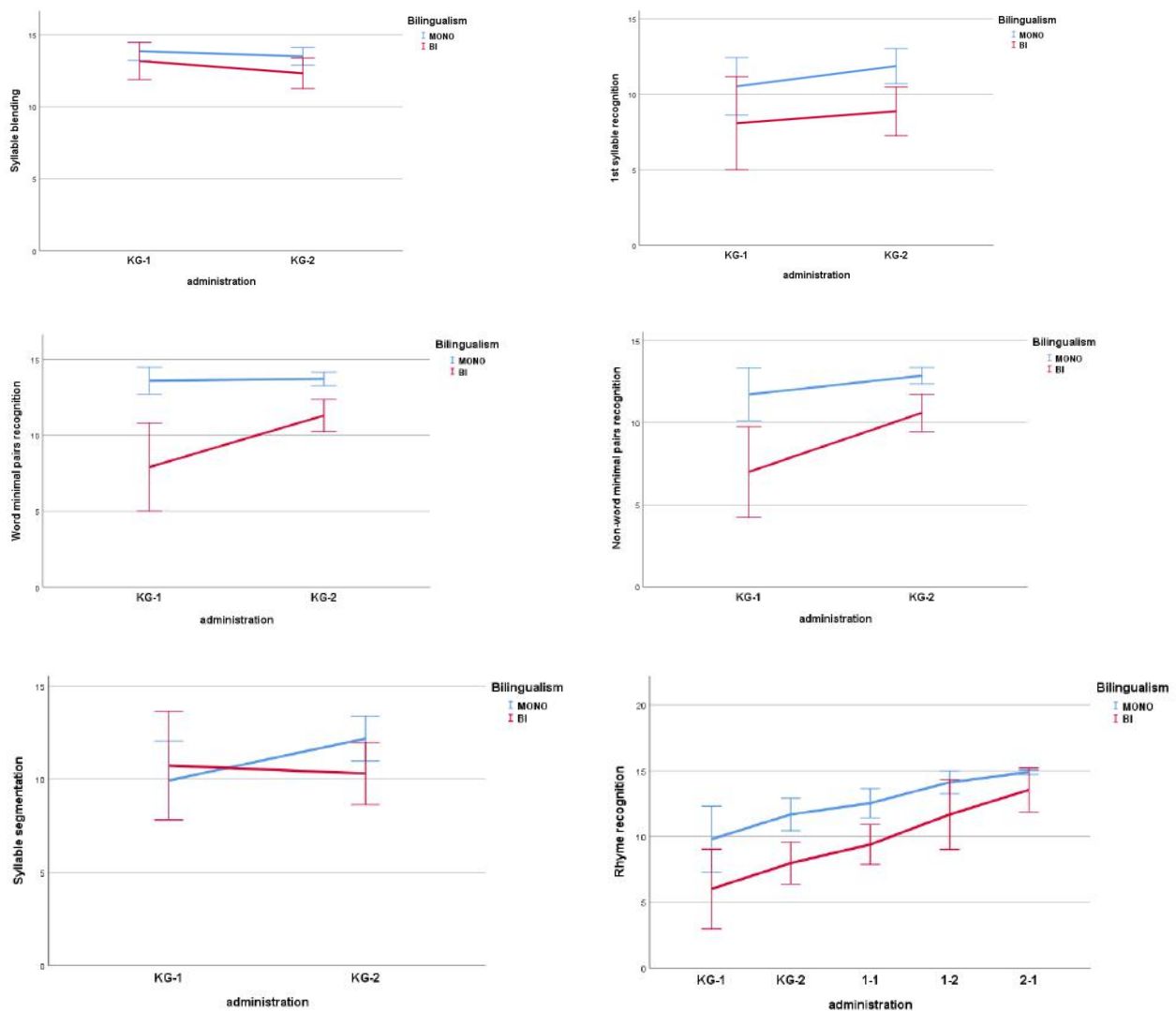


Figure 9. Longitudinal means of monolingual and bilingual children in the two kindergarten administrations of the syllable-awareness tests, and in the five administrations of the rhyme-recognition test.

A series of GMMRM with bilingualism and administration time as fixed effects correctly predict syllable blending ($F(2, 73) = 3.222, p = 0.046$), recognition of the initial syllable ($F(2, 56) = 4.229, p = 0.019$), of word minimal pairs ($F(2, 45) = 15.774, p = 0.000$), of non-word minimal pairs ($F(2, 43) = 10.610, p = 0.000$), and of rhymes ($F(5, 66) = 18.400, p =$

0.000). Main effects of bilingualism on each one of these abilities, and of administration time on the recognition of non-word minimal pairs and rhymes were found. There were positive effects of monolingualism on every ability, a negative effect of administration KG-1 on the recognition of non-word minimal pairs, and negative effects of administrations KG-1, KG-2, and 1-1 on the recognition of rhymes.

The post hoc pairwise comparison of effects of administration time on rhyme recognition revealed a significant improvement between KG-2 and 1-2, and between 1-1 and 2-1. Significant contrasts are in Table 5.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p (seq. Bonferroni)	Lower	Upper
Rhyme recognition	KG-1	1-2	- 4.345	1.037	- 4.191	42	0.001	- 7.278	- 1.411
	KG-1	2-1	- 5.631	0.941	- 5.985	32	0.000	- 8.384	- 2.877
	KG-2	1-2	- 2.532	0.739	- 3.428	64	0.006	- 4.543	- 0.521
	KG-2	2-1	- 3.818	0.597	- 6.398	90	0.000	- 5.535	- 2.100
	1-1	2-1	- 3.209	0.598	- 5.363	88	0.000	- 4.911	- 1.508

Table 5. Significant post hoc comparisons between the effects of administration time on the recognition of rhymes.

The combination between bilingualism and administration time as a fixed effect predicts significant differences between monolinguals and bilinguals in the second administration of syllable blending (marginally significant) and initial-syllable-recognition, in both administrations of word and non-word minimal pair recognition, and in the first three administrations of rhyme recognition, as shown in Table 6.

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Syllable blending	MONO*KG-2	BI*KG-2	1.148	0.567	2.024	75	0.046	0.018	2.277
Word minimal pair recognition	MONO*KG-1	BI*KG-1	4.145	1.091	3.800	24	0.001	1.894	6.397
	MONO*KG-2	BI*KG-2	2.071	0.451	4.589	71	0.000	1.171	2.971

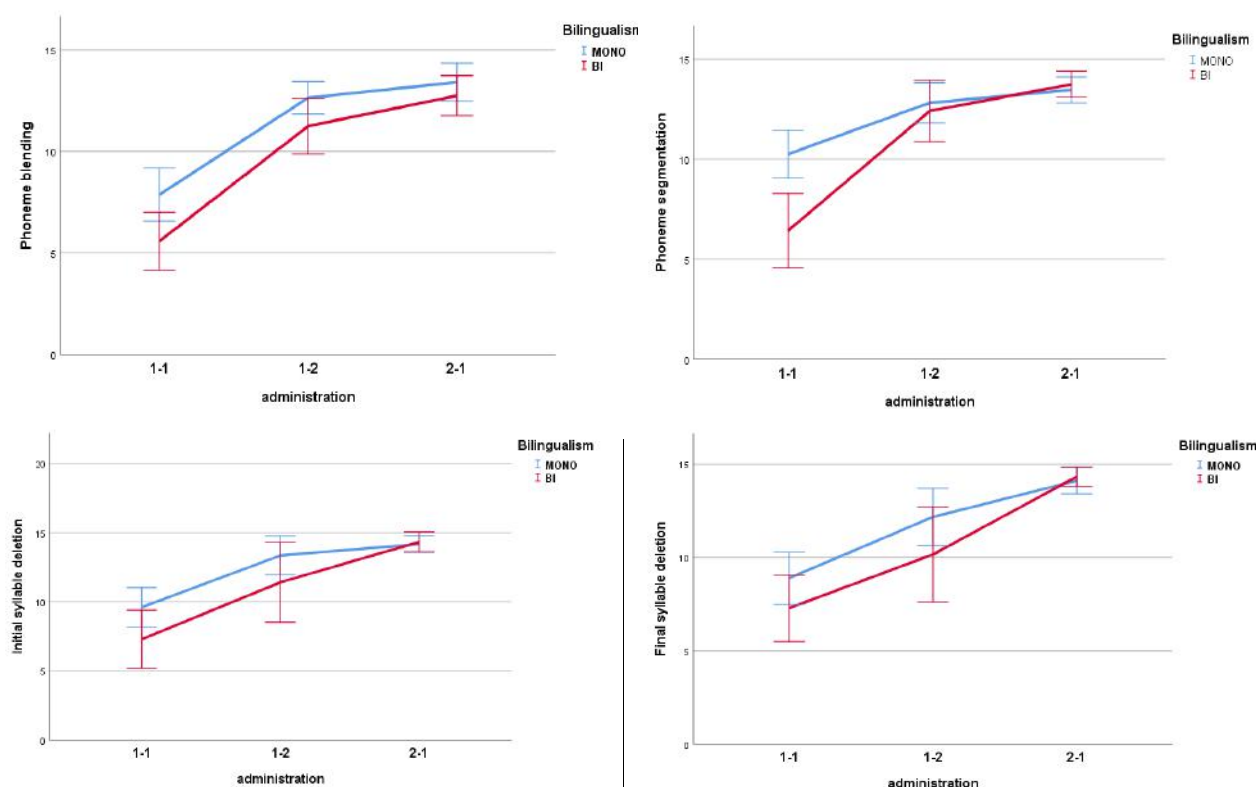
Non-word minimal pair recognition	MONO*KG-1	BI*KG-1	3.006	1.235	2.435	24	0.023	0.458	5.554
	MONO*KG-2	BI*KG-2	1.931	0.537	3.592	71	0.001	0.859	3.002
Initial syllable recognition	MONO*KG-2	BI*KG-2	2.407	0.857	2.808	73	0.006	0.698	4.116
Rhyme recognition	MONO*KG-1	BI*KG-1	4.990	1.831	2.725	18	0.014	1.143	8.837
	MONO*KG-2	BI*KG-2	3.757	0.972	3.864	63	0.000	1.814	5.700
	MONO*1-1	BI*1-1	3.397	0.929	3.657	66	0.001	1.542	5.251

Table 6. Significant differences predicted by the combined effect of bilingualism and administration time on global metaphonological skills.

There were no significant effects of either bilingualism or administration time on syllable segmentation ($p \geq 0.055$).

Analytical metaphonological skills

Phoneme blending and segmentation, first and final syllable deletion and phonemic verbal fluency were tested three times, whereas spoonerism only once in the middle of the 2nd grade of primary school. Figure 10 shows the results of the three longitudinal administrations of phonemic-awareness tests in the 1st and 2nd grades.



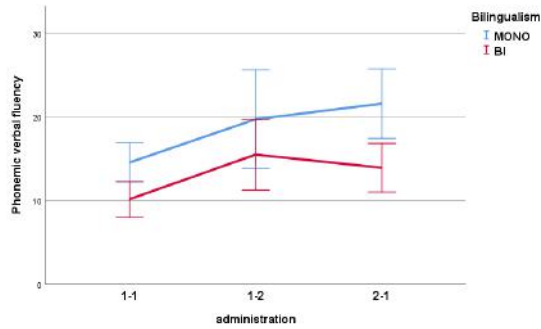


Figure 10. Longitudinal raw scores obtained by bilingual and monolingual children in phoneme blending and segmentation, initial and final syllable deletion, and phonemic verbal fluency.

A GMMRM with bilingualism and administration time as fixed effects correctly predicts phoneme blending ($F(3, 45) = 31.111, p = 0.000$) and segmentation ($F(3, 62) = 15.400, p = 0.000$), initial syllable deletion ($F(3, 58) = 16.000, p = 0.000$), final syllable deletion ($F(3, 46) = 24.069, p = 0.000$), and phonemic verbal fluency ($F(3, 59) = 14.541, p = 0.000$). A main effect of bilingualism on phonemic verbal fluency, and effects of administration time on every ability emerged.

The post hoc pairwise comparison of effects of administration time on phoneme blending ($F(2, 120) = 76.072, p = 0.000$), phoneme segmentation ($F(2, 68) = 32.812, p = 0.000$), deletion of the initial syllable ($F(2, 73) = 35.052, p = 0.000$), deletion of the final syllable ($F(2, 86) = 56.827, p = 0.000$), and phonemic verbal fluency ($F(2, 96) = 17.557, p = 0.000$) predict significant differences that are in Table 7.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p (seq. Bonferroni)	Lower	Upper
Phoneme blending	1-1	1-2	- 4.769	0.447	- 10.657	120	0.000	- 5.855	- 3.682
	1-1	2-1	- 5.846	0.480	- 12.192	120	0.000	- 6.395	- 4.758
	1-2	2-1	- 1.078	0.311	- 3.469	120	0.001	- 1.693	- 0.463
Phoneme segmentation	1-1	1-2	- 3.394	0.585	- 5.802	118	0.000	- 4.721	- 2.066

	1-1	2-1	- 4.325	0.536	- 8.065	92	0.000	- 5.633	- 3.018
	1-2	2-1	- 0.932	0.473	- 1.971	47	0.055	- 1.833	0.019
Initial syllable deletion	1-1	1-2	- 3.010	0.698	- 4.312	117	0.000	- 4.595	- 1.425
	1-1	2-1	- 4.702	0.563	- 8.357	85	0.000	- 6.076	- 3.328
	1-2	2-1	- 1.692	0.664	- 2.548	34	0.016	- 3.041	- 0.342
Final syllable deletion	1-1	1-2	- 2.662	0.685	- 3.885	114	0.000	- 4.020	- 1.305
	1-1	2-1	- 5.525	0.524	- 10.544	95	0.000	- 6.802	- 4.248
	1-2	2-1	- 2.862	0.622	- 4.603	36	0.000	- 4.317	- 1.408
Phonemic verbal fluency	1-1	1-2	- 4.788	1.236	- 3.875	46	0.001	- 7.652	- 1.925
	1-1	2-1	- 5.755	0.987	- 5.832	103	0.000	- 8.157	- 3.353

Table 7. Significant post hoc comparisons between effects of administration time on phoneme blending and segmentation, deletion of the initial and final syllable, and phonemic verbal fluency.

The combination between bilingualism and administration time as a fixed effect predicts a significant difference between monolinguals and bilinguals in administrations 1-1 and 2-1 of the phonemic-verbal-fluency test, as shown in Table 8.

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Phonemic verbal fluency	MONO*1-1	BI*1-1	4.080	1.555	2.625	68	0.011	0.978	7.182
	MONO*2-1	BI*2-1	7.677	2.353	3.262	26	0.003	2.839	12.514

Table 8. Significant differences predicted by the combined effect of bilingualism and administration time on analytical metaphonological skills.

Moreover, there is a significant negative effect of the association between bilingualism and administration 1-1 on phoneme segmentation ($Coeff = - 0.576$, $SE = 0.098$, $t = - 5.893$, $p = 0.000$, $- 0.770 < CI < - 0.381$).

Figure 11 displays cross-sectional results of the spoonerism test in the 2nd grade of primary school.

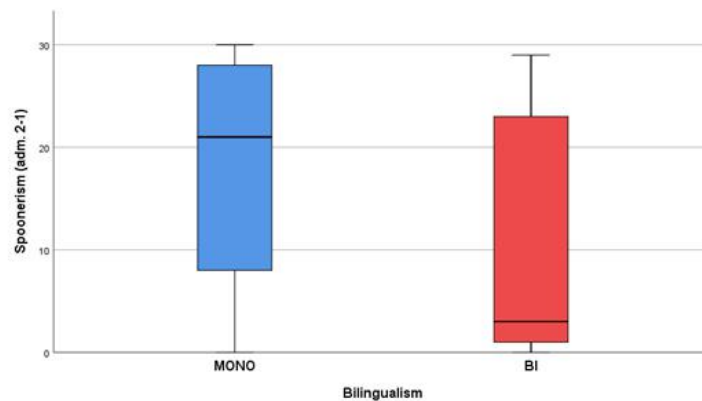


Figure 11. Distribution of cross-sectional raw scores obtained by monolingual and bilingual children in the spoonerism test in the middle of the 2nd grade.

A GMM found no effect of bilingualism on spoonerism ($p = 0.069$).

To sum up the outcomes of metaphonological tests, monolingual kindergarteners showed significantly higher syllable awareness than bilinguals in every ability except syllable segmentation, which was similar between groups. The rhyme-recognition performance was significantly different between groups until the middle of the 1st grade but it turned to be similar from the end of the 1st grade.

As for phonemic awareness in primary school, bilingual children showed an initial disadvantage in phoneme segmentation, but they reached comparable outcomes as monolinguals at the end of the 1st grade in every ability except phonemic verbal fluency, which was still significantly different in the middle of the 2nd grade.

3.3.3. RAN

RAN of pictures was tested five times from KG to the 2nd grade, whereas RAN of colors and numbers three times in primary school. Speed and error longitudinal scores are in Figure 12.

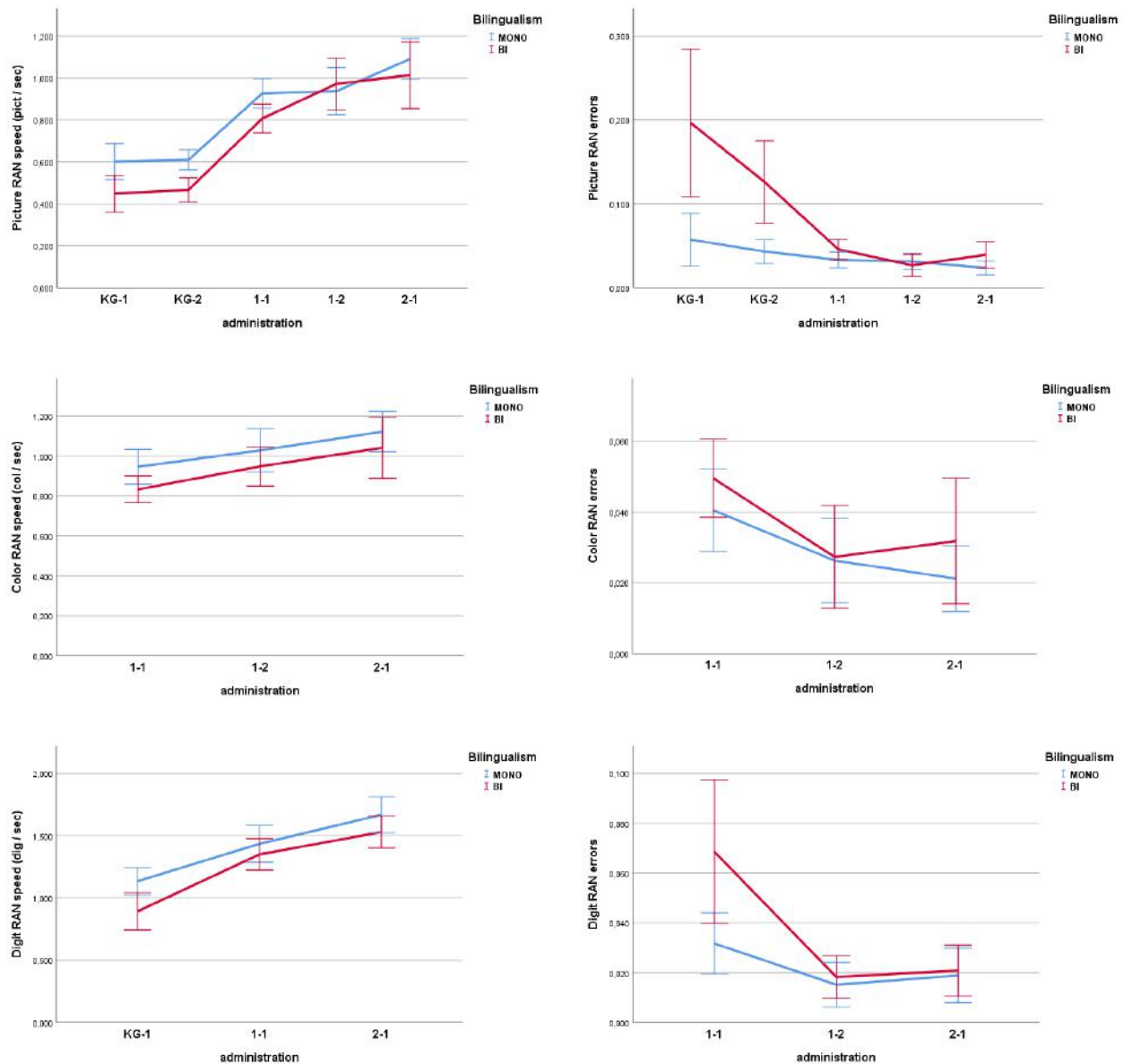


Figure 12. Speed and error scores obtained in the five longitudinal administrations of picture RAN, and in the three longitudinal administrations of color and digit RAN.

A series of GMMRM with bilingualism and administration time as fixed effects correctly predicts picture speed ($F(5, 73) = 55.783, p = 0.000$) and errors ($F(5, 76) = 17.834, p =$

0.000), color speed ($F(3, 80) = 9.622, p = 0.000$) and errors ($F(3, 64) = 4.284, p = 0.008$), and digit speed ($F(3, 70) = 29.654, p = 0.000$) and errors ($F(3, 59) = 11.159, p = 0.000$).

Main effects of bilingualism on speed and errors in the RAN of pictures and on errors in the RAN of digits, as well as main effects of administration time on every RAN measure emerged.

The post hoc pairwise comparison of effects of administration time on RAN found:

- for pictures, a significant increase of speed between administration KG-2 and 1-1, 1-1 and 2-1, and a significant decrease of errors between KG-2 and 1-1;
- for colors, a significant increase of speed between administration 1-1 and 1-2, 1-2 and 2-1, and a significant decrease of errors between administration 1-1 and 2-1;
- for digits, a significant increase of speed between 1-1 and 1-2, between 1-2 and 2-1, and a significant decrease of errors between 1-1 and 1-2, and between 1-1 and 2-1.

Significant contrasts are in Table 9.

CI									
Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p	Lower	Upper
(seq. Bonferroni)									
Picture RAN speed	KG-1	1-1	- 0.339	0.037	- 9.129	59	0.000	- 0.443	- 0.236
	KG-1	1-2	- 0.418	0.050	- 8.394	48	0.000	- 0.551	- 0.284
	KG-1	2-1	- 0.522	0.050	- 10.350	49	0.000	- 0.666	- 0.378
	KG-2	1-1	- 0.330	0.031	- 10.803	130	0.000	- 0.417	- 0.243
	KG-2	1-2	- 0.408	0.045	- 9.062	39	0.000	- 0.534	- 0.283
	KG-2	2-1	- 0.513	0.046	- 11.193	41	0.000	- 0.647	- 0.379
	1-1	2-1	- 0.183	0.048	- 3.802	49	0.002	- 0.308	- 0.058
Picture RAN errors	KG-1	1-1	0.077	0.020	3.855	22	0.004	0.021	0.134
	KG-1	1-2	0.086	0.020	4.255	23	0.002	0.025	0.146
	KG-1	2-1	0.083	0.020	4.161	22	0.003	0.024	0.143
	KG-2	1-1	0.045	0.012	3.665	73	0.003	0.012	0.079
	KG-2	1-2	0.053	0.012	4.306	73	0.001	0.017	0.089
	KG-2	2-1	0.051	0.012	4.189	69	0.001	0.016	0.086
Color RAN speed	1-1	1-2	- 0.096	0.038	- 2.566	101	0.014	- 0.173	- 0.019
	1-1	2-1	- 0.190	0.044	- 4.330	69	0.000	- 0.297	- 0.082
	1-2	2-1	- 0.094	0.034	- 2.752	123	0.014	- 0.171	- 0.016

Color RAN errors	1-1	2-1	0.017	0.006	2.995	75	0.011	0.003	0.031
Digit RAN speed	1-1	1-2	- 0.367	0.053	- 6.890	121	0.000	- 0.488	- 0.246
	1-1	2-1	- 0.578	0.055	- 10.436	122	0.000	- 0.713	- 0.444
	1-2	2-1	- 0.212	0.042	- 4.992	122	0.000	- 0.296	- 0.128
Digit RAN errors	1-1	1-2	0.034	0.009	3.896	63	0.001	0.012	0.055
	1-1	2-1	0.027	0.009	3.164	71	0.005	0.008	0.047

Table 9. Significant post hoc pairwise comparisons between effects of administration times on RAN speed and errors for pictures, colors, and digits.

The combination of bilingualism and administration time as a fixed effect predicts a significant difference between monolinguals and bilinguals in picture speed in the first three administrations (until the middle of the 1st grade); a significant difference in picture errors in the first two administrations (until the end of kindergarten); and a significant difference in digit errors in the first administration (1-1) (Table 10).

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Picture RAN speed	MONO*KG-1	BI*KG-1	0.132	0.054	2.423	22	0.024	0.019	0.245
	MONO*KG-2	BI*KG-2	0.139	0.039	3.616	67	0.001	0.062	0.216
	MONO*1-1	BI*1-1	0.120	0.048	2.483	68	0.016	0.024	0.216
Picture RAN errors	MONO*KG-1	BI*KG-1	- 0.143	0.054	- 2.636	20	0.016	- 0.256	- 0.030
	MONO*KG-2	BI*KG-2	- 0.072	0.023	- 3.078	59	0.003	- 0.119	- 0.025
Digit RAN errors	MONO*1-1	BI*1-1	- 0.037	0.017	- 2.196	55	0.032	- 0.070	- 0.003

Table 10. Significant differences predicted by the combined effect of bilingualism and administration time on RAN speed and errors.

Moreover, there is a significant negative effect of the combination between bilingualism and administration 1-1 on speed in the RAN of digits (*Coeff* = - 0.594, *SE* = 0.102, *t* = - 5.819, *p* = 0.000, - 0.798 < *CI* < - 0.390).

In sum, monolingual children were significantly faster than bilinguals in the RAN of pictures until the middle of the 1st grade of primary school, and significantly more accurate during the last year of kindergarten.

In primary school, monolinguals and bilinguals revealed comparable performances in the RAN of colors and RAN speed of digits. Bilinguals made more errors in digit RAN in the middle of the 1st grade. From the end of the 1st grade, the groups' performances were similar in every RAN category.

3.3.4. Short-term memory

Figure 13 shows the results of the three longitudinal administrations of digit span forward, digit span backward, and word span tests.

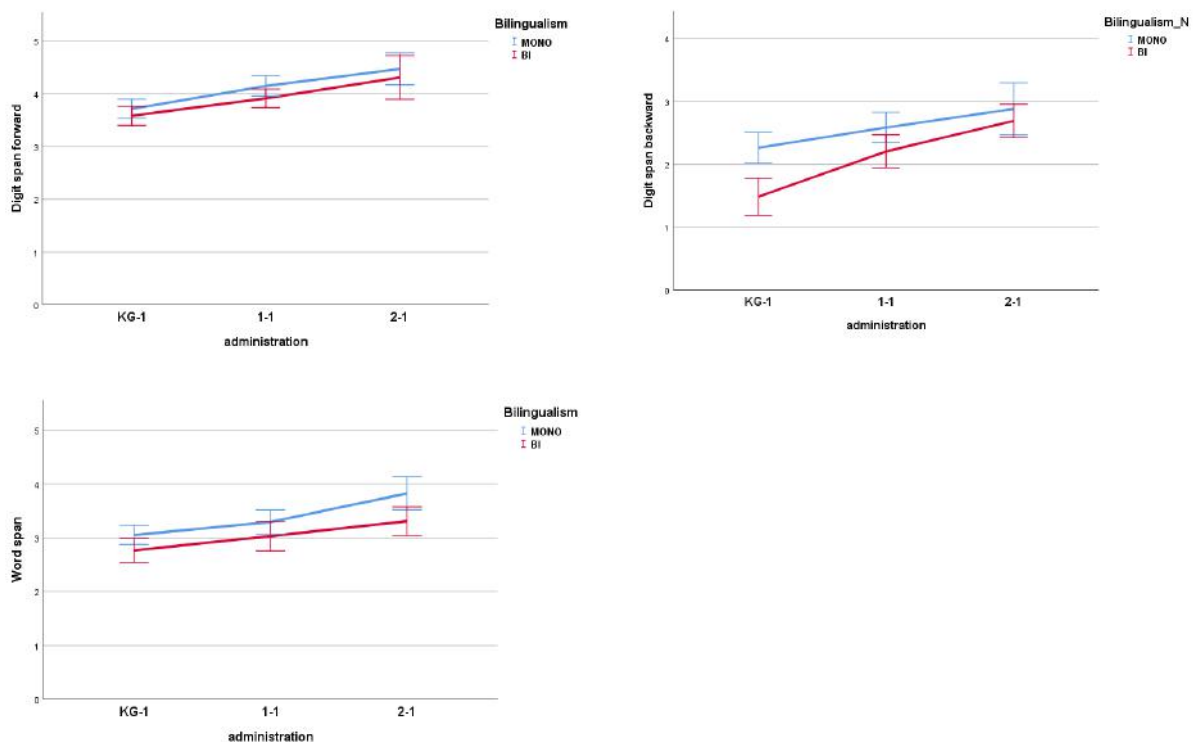


Figure 13. Longitudinal results obtained by monolingual and bilingual children in the tests of digit span forward, digit span backward, and word span.

A series of GMMRM with bilingualism and administration time as fixed effects correctly predicts digit span forward ($F(3, 90) = 24.294, p = 0.000$), digit span backward ($F(3, 101) = 14.307, p = 0.000$), and word span ($F(3, 104) = 14.075, p = 0.000$).

There are main effects of bilingualism on digit span backward and word span, and of administration time on digit span forward and backward and word span, with positive effects of monolingualism on digit span backward and word span, and negative effects of administration KG-1 and 1-1 on every ability. The post hoc pairwise comparison of effects of administration time showed significant differences between each pair, with improvements in every administration, as shown in table 11.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p (seq. Bonferroni)	Lower	Upper
Digit span forward	KG-1	1-1	- 0.387	0.066	- 5.289	170	0.000	- 0.547	- 0.226
	KG-1	2-1	- 0.743	0.120	- 6.185	56	0.000	- 1.019	- 0.466
	1-1	2-1	- 0.356	0.134	- 2.564	51	0.011	- 0.625	- 0.087
Digit span backward	KG-1	1-1	- 0.401	0.102	- 3.918	166	0.000	- 0.633	- 0.170
	KG-1	2-1	- 0.785	0.139	- 5.651	66	0.000	- 1.126	- 0.444
	1-1	2-1	- 0.384	0.154	- 2.488	51	0.016	- 0.693	- 0.074
Word span	KG-1	1-1	- 0.251	0.090	- 2.802	174	0.009	- 0.445	- 0.058
	KG-1	2-1	- 0.662	0.117	- 5.666	67	0.000	- 0.949	- 0.375
	1-1	2-1	- 0.411	0.139	- 2.960	63	0.009	- 0.730	- 0.092

Table 11. Significant post hoc pairwise comparisons between effects of administration times on digit span forward and backward, and word span.

The combination between bilingualism and administration time predicted significant differences between monolinguals and bilinguals in the first two administrations of digit span backward, and in the first and third administration of the word-span test, as shown in Table 12.

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Digit span backward	MONO*KG-1	BI*KG-1	0.559	0.170	3.296	63	0.002	0.220	0.899
	MONO*1-1	BI*1-1	0.403	0.185	2.182	70	0.032	0.035	0.772
Word span	MONO*KG-1	BI*KG-1	0.318	0.137	2.320	68	0.023	0.044	0.591
	MONO*2-1	BI*2.1	0.661	0.176	3.749	28	0.001	0.300	1.022

Table 12. Significant differences predicted by the combined effect of bilingualism and administration time on short-term memory.

In sum, bilingual children showed comparable digit span forward as monolinguals but smaller digit span backward in kindergarten and 1st grade, and smaller word span in kindergarten and 2nd grade.

3.3.5. Visual memory and discrimination

Figure 14 shows the results of the three longitudinal administrations of the visual tests of graphic reproduction of semicircles and discrimination of capital letters.

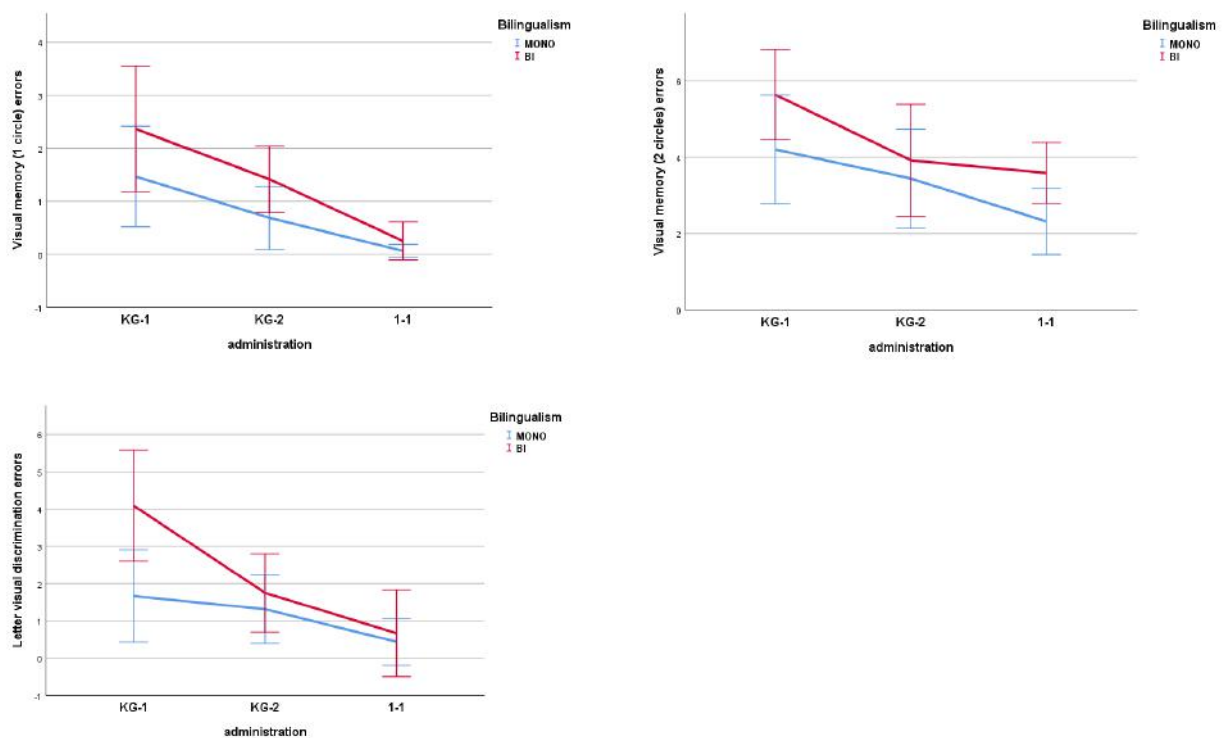


Figure 14. Error scores obtained by monolingual and bilingual children in the three longitudinal administrations of graphic reproduction of one and two semicircles, and in the visual discrimination of capital letters with different orientations.

A GMMRM with bilingualism and administration time as fixed effects correctly predicts the visual reproduction of one ($F(3, 14) = 4.372, p = 0.022$) and two ($F(3, 45) = 8.934, p = 0.000$) semicircles, whereas there is no significant effect on the discrimination of capital letters ($p = 0.100$).

There were significant positive effects of administration KG-1 on errors in the reproduction of one circle, and of administration KG-1 and KG-2 on errors in the reproduction of two semicircles.

The post hoc pairwise comparison of effects of administration time on visual errors showed a significant improvement in the reproduction of one semicircle between KG-1 and KG-2, and of two semicircles between KG-1 and KG-2 and between KG-2 and 1-1. An improvement in the discrimination of capital letters was also found between KG-1 and KG-2, with a marginally significant effect of administration time. Significant contrasts are in Table 13.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p (seq. Bonferroni)	Lower	Upper
Visual memory errors (one-circle reproduction)	KG-1	KG-2	1.077	3.078	2.851	32	0.023	0.122	2.032
	KG-1	1-1	1.501	0.495	3.030	16	0.023	0.188	2.813
Visual memory errors (two-circle reproduction)	KG-1	KG-2	1.646	0.466	3.534	58	0.002	0.498	2.794
	KG-2	1-1	1.303	0.383	3.401	70	0.002	0.412	2.194
Letter discrimination errors	KG-1	KG-2	1.565	0.565	2.769	32	0.028	0.137	2.992

Table 13. Significant post hoc pairwise comparisons between effects of administration time on errors in the graphic reproduction of one and two semicircles oriented differently, and in the visual discrimination of capital letters.

An analysis of the combined effect of bilingualism and administration time as a fixed effect found a marginally significant positive effect on letter-recognition errors in administration KG-1 ($Coeff = 0.539, SE = 0.255, t = 2.118, p = 0.045, 0.014 < CI < 1.064$).

In sum, monolingual and bilingual children showed comparable performances of reproduction of one and two semicircles. There was a significant improvement in the reproduction of single semicircles between the two kindergarten administrations, whereas the reproduction of two semicircles improved significantly at every administration. The ability to visually discriminate the position of capital letters oriented in different ways showed a marginally significant positive effect of bilinguals' errors in the first administration (KG-1) but the difference between groups was not significant. This ability improved significantly between the first two administrations in kindergarten,

3.3.6. Literacy skills

Sublexical decoding

Figure 15 shows speed and error scores obtained by monolingual and bilingual children in the two longitudinal administrations of sublexical decoding tests²⁷ in the middle of the 1st and 2nd grades of primary school.

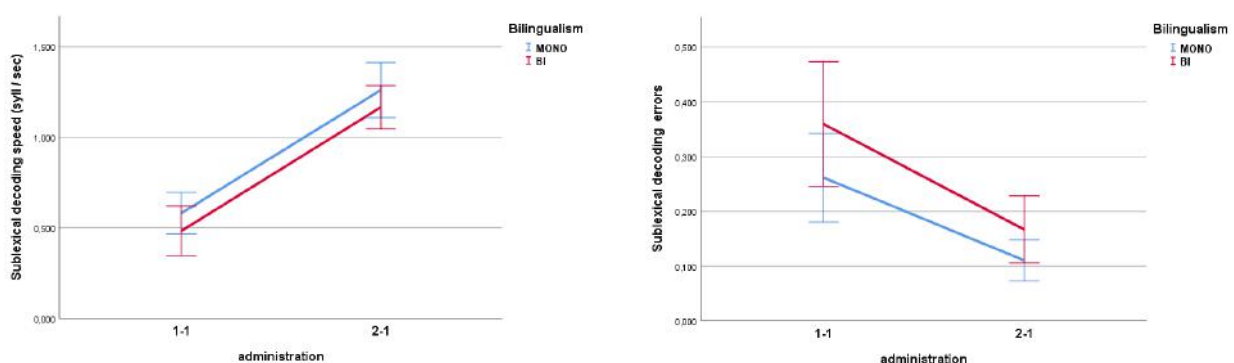


Figure 15. Speed and error scores obtained by bilingual and monolingual children in the two longitudinal administrations of the sublexical decoding test.

²⁷ Sublexical decoding was tested through a list of alphabetic strings constituting sublexical components of Italian words (Calgaro et al., 2018) in the middle of the 1st grade (1-1), and through a pseudoword decoding test (Traficante et al., 2017) in the middle of the 2nd grade of primary school (2-1). For more details about the tests, see 2.9.

A GMMRM with bilingualism and administration time as fixed effects correctly predicts sublexical decoding speed ($F(2, 57) = 40.909, p = 0.000$) and errors ($F(2, 74) = 14.877, p = 0.000$).

Main effects of bilingualism on sublexical decoding errors, and of administration time on sublexical decoding speed and errors were found. Monolingualism had a significant negative effect on sublexical decoding errors, and administration 1-1 a negative effect on sublexical decoding speed and a positive effect on sublexical decoding errors.

The combination between bilingualism and administration time predicts a negative effect of monolingualism in the second administration, and a positive effect of bilingualism in the first administration on sublexical decoding errors, as shown in Table 14, but no significant difference between groups.

Dependent variable	Predictor	Coeff.	StdErr	t	p	Lower	Upper
Sublexical decoding errors	MONO*2-1	- 0.507	0.237	- 2.143	0.043	- 0.999	- 0.016
	BI*1-1	0.753	0.248	3.034	0.004	0.257	1.250

Table 14. Significant combined effects of bilingualism and administration time on sublexical decoding.

Word and non-word decoding

Figure 16 shows speed and error scores obtained in the middle of the 2nd grade by bilingual and monolingual children in decoding generic lists of words and non-words (first and second row); stress errors in decoding lists of non-paroxitone words; and attentional errors in decoding word pairs.

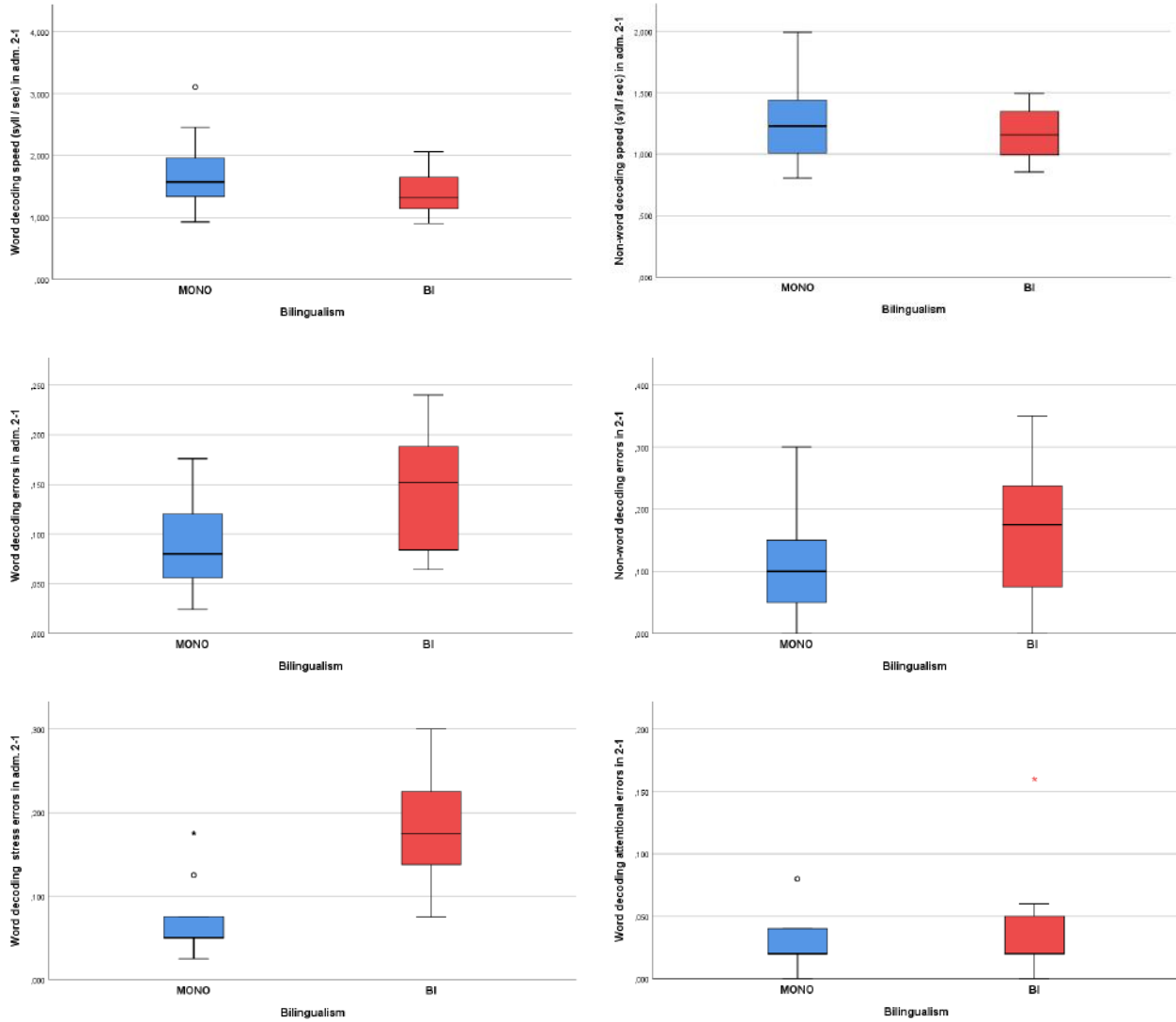


Figure 16. Speed and errors in word and non-word decoding by monolingual and bilingual children in the middle of the 2nd grade of primary school.

A GMM with bilingualism as fixed effect correctly predicts errors in decoding generic word lists ($F(1, 24) = 5.783, p = 0.024$) and words with atypical stress ($F(1, 25) = 27.132, p = 0.000$). Significant positive effects of bilingualism on errors in decoding generic word lists and words with atypical stress emerged.

Bilingual 2nd graders make significantly more generic errors in word decoding ($M = 0.137, SE = 0.015, 0.105 < CI < 0.168$) than monolinguals ($M = 0.089, SE = 0.013, 0.062 < CI < 0.115$), and significantly more stress errors ($M = 0.177, SE = 0.037, 0.004 < CI < 0.290$) than

monolinguals ($M = 0.067$, $SE = 0.014$, $0.000 < CI < 49.408$). Significant contrasts are in Table 15.

Dependent variable	Contrast		Est.	StdErr	t	df	p	CI	
								Lower	Upper
Word decoding errors	MONO	BI	- 0.048	0.020	- 2.405	24	0.024	- 0.090	- 0.007
Stress errors	MONO	BI	- 0.110	0.031	- 3.559	3	0.041	- 0.212	- 0.009

Table 15. Estimates of significant differences between monolingual and bilingual children in word decoding generic errors and stress errors.

Text decoding

Figure 17 displays the results of the two longitudinal sessions of text decoding at the end of the 1st grade and in the middle of the 2nd grade of primary school.

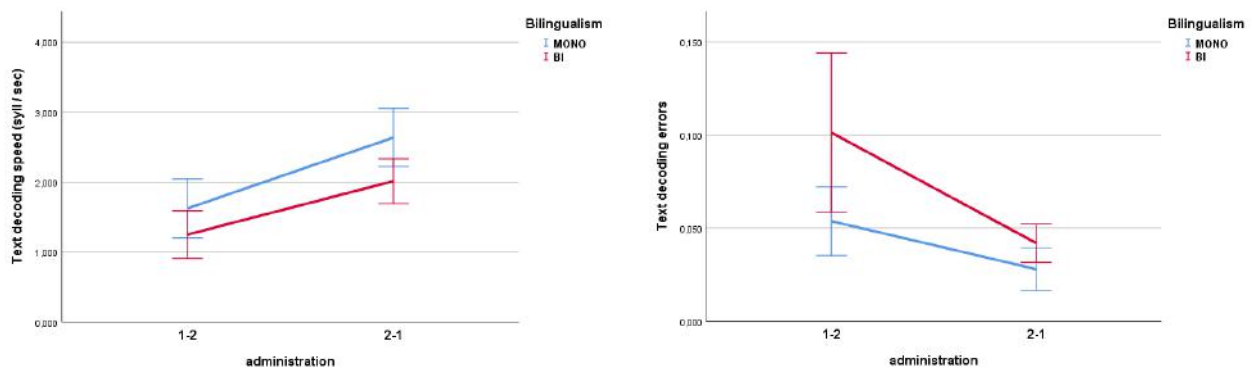


Figure 17. Speed and error score by monolingual and bilingual children in the two longitudinal tests of text decoding.

A GMMRM with bilingualism and administration time as fixed effects correctly predicts text decoding speed ($F(2, 44) = 13.289$, $p = 0.000$) and errors ($F(2, 52) = 12.559$, $p = 0.000$). Main effects of bilingualism and administration time on text decoding speed and errors were found. Monolingualism predicts faster decoding and fewer errors, whereas administration 1- 2 predicts slower decoding and more errors.

The combination between bilingualism and administration time predicts a significant difference between monolinguals and bilinguals in text decoding speed in the second administration (2-1), as shown in Table 16.

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Texr decoding speed	MONO*2-1	BI*2-1	0.740	0.269	2.745	24	0.011	0.184	1.296

Table 16. Significant difference predicted by the combined effect of bilingualism and administration time on text decoding speed.

The bilingual group predicts a significant positive effect on text decoding errors in the first administration at the end of the 1st grade (*Coeff.* = 0.930, *SE* = 0.303, *t* = 3.073, *p* = 0.004, $0.321 < CI < 1.539$). The difference between monolinguals and bilinguals is marginally non-significant (*p* = 0.059).

Reading comprehension

Reading comprehension was tested at the end of the 1st grade and in the middle of the 2nd grade. The full sample was tested through the MT-2 battery, which requires the global comprehension of one narrative text, whereas the core sample was tested through the MT-3 clinical tests, which require a more thorough comprehension of two narrative texts in the 1st grade, and one narrative plus one informative text in the 2nd grade. For the latter test, an analysis of specific and central comprehension was performed (see 2.11) besides reporting total comprehension scores (specific + central).

Longitudinal results are in Figure 18.

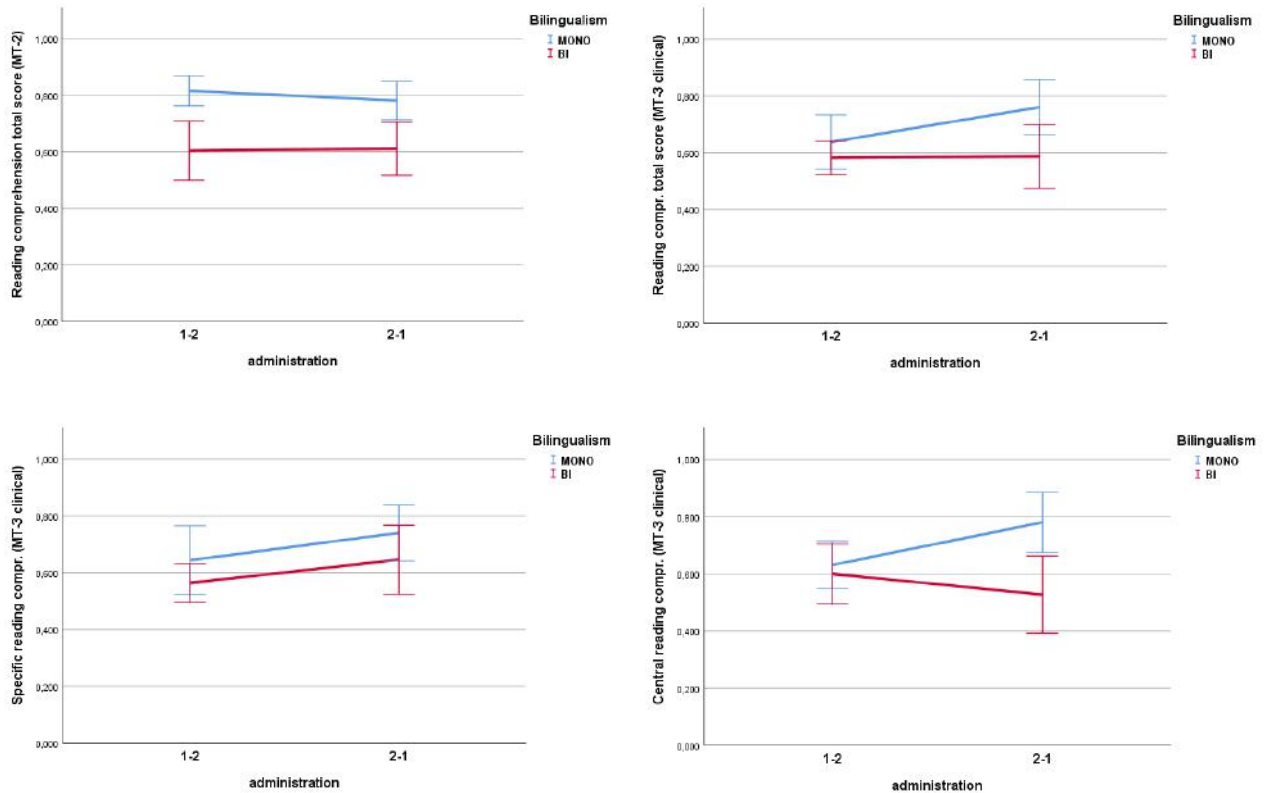


Figure 18. Raw scores obtained by bilingual and monolingual children in the two longitudinal tests of reading comprehension.

A GMMRM with bilingualism and administration time as fixed effects correctly predicts reading comprehension total scores (for MT-2: ($F(2, 110) = 9.784, p = 0.000$); for MT-3: ($F(2, 50) = 3.447, p = 0.040$)), and comprehension central scores ($F(1^{28}, 50) = 9829.478, p = 0.000$), whereas no significant effect on comprehension specific scores emerged ($p = 0.084$).

The combination between bilingualism and administration time predicts a significant difference between groups' comprehension total scores in both administrations for the MT-2 test, and in the second administration for the MT-3 tests, and a significant difference in comprehension central scores in the second administration, as shown in Table 17.

²⁸ Robust model

Dependent variable	Contrast		Est.	StdErr	t	df	p	Lower	Upper
Reading comprehension (MT-2)	MONO*1-2	BI*1-2	0.212	0.060	3.557	53	0.001	0.092	0.331
Reading comprehension (MT-2)	MONO*2-1	BI*2-1	0.147	0.055	2.682	56	0.010	0.037	0.256
Reading comprehension (MT-3) (total score)	MONO*2-1	BI*2-1	0.174	0.075	2.329	24	0.029	0.020	0.327
Reading comprehension (MT-3) (central score)	MONO*2-1	BI*2-1	0.253	0.084	2.993	24	0.006	0.078	0.427

Table 17. Significant differences predicted by the combined effect of bilingualism and administration time on reading comprehension.

Word and non-word writing

Word writing was tested three times between the 1st and 2nd grade, whereas non-word writing was tested once in the middle of the 2nd grade. Figure 19 shows the results of longitudinal tests of word-dictation, and of the cross-sectional test of non-word dictation in the middle of the 2nd grade.

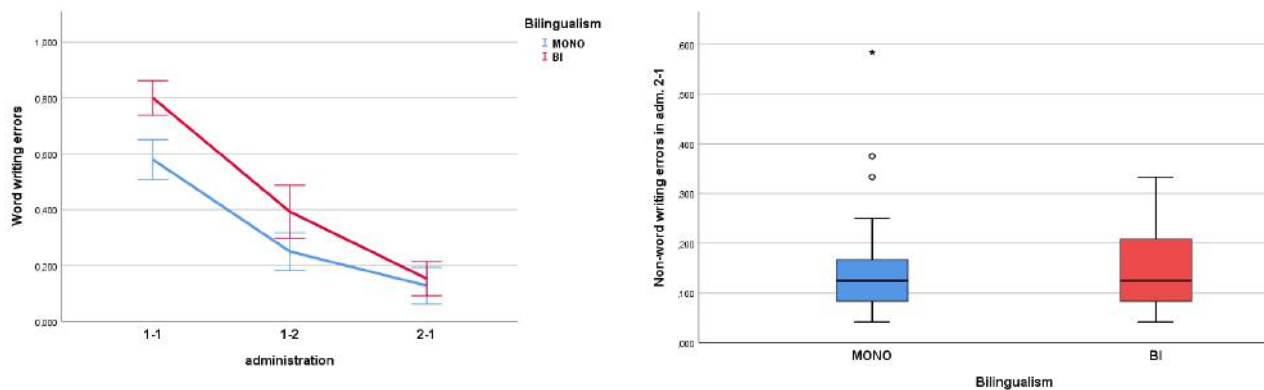


Figure 19. Error scores obtained by monolingual and bilingual children in the three longitudinal tests of word dictation, and distribution of error scores in the cross-sectional test of non-word dictation (2-1).

A GMMRM with bilingualism and administration time as fixed effects correctly predicts word writing ($F(3, 64) = 61.902, p = 0.000$). Main effects of bilingualism and administration time on word writing errors were found. Monolingualism has a significant negative effect on word writing errors, whereas administrations 1-1 and 1-2 have a significant positive effect on word writing errors.

A post hoc pairwise comparison of the effects of administration time revealed a significant difference between each administration, with significant improvement in the second and third administrations, as shown in Table 18.

Dependent variable	Contrast		Est.	StdErr	t	df	Adj. p	Lower	Upper
	(seq. Bonferroni)								
Word writing errors	1-1	1-2	0.376	0.034	11.208	173	0.000	0.295	0.457
	1-1	2-1	0.551	0.030	18.188	145	0.000	0.483	0.620
	1-2	2-1	0.175	0.036	4.858	96	0.000	0.104	0.247

Table 18. Significant post hoc pairwise comparisons between effects of administration time on word dictation errors.

The combination between bilingualism and administration time predicts a significant error difference between monolinguals and bilinguals in the 1st grade, in both the first and second administration (Table 19).

Dependent variable	Contrast		Est.	StdErr	t	df	p	CI	
								Lower	Upper
Word writing errors	MONO*1-1	BI*1-1	- 0.203	0.054	- 3.765	107	0.000	- 0.309	- 0.096
	MONO*1-2	BI*1-2	- 0.175	0.060	- 2.916	51	0.005	- 0.295	- 0.054

Table 19. Significant differences predicted by the combined effect of bilingualism and administration time on word writing errors

A GMM revealed no effect of bilingualism on non-word writing errors ($p = 0.194$).

Text writing

Text writing under dictation was tested once in the middle of the 2nd grade. The proportion of writing errors is in Figure 20.

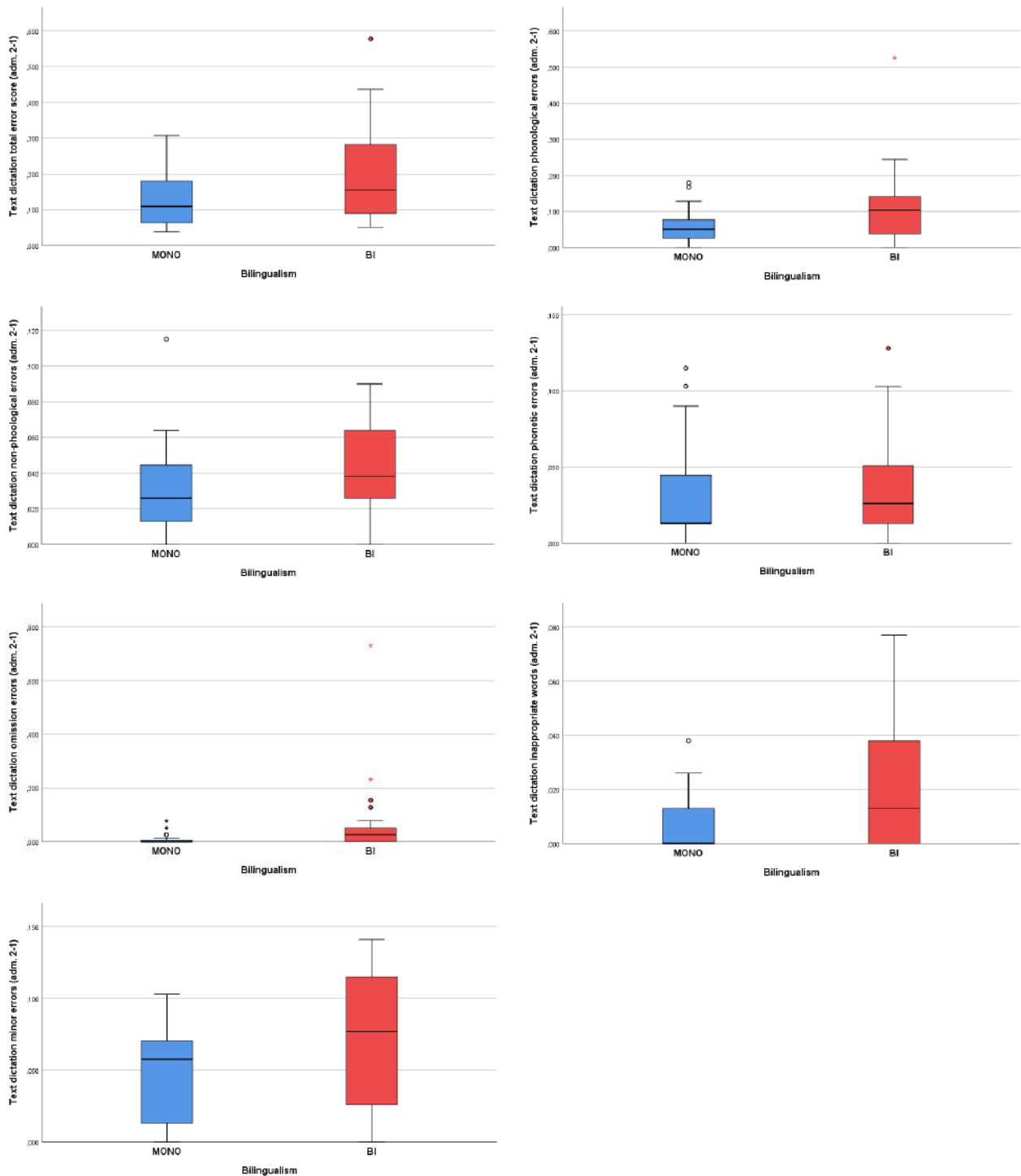


Figure 20. Distribution of writing errors by monolingual and bilingual children in the cross-sectional test of text dictation in the middle of the 2nd grade.

A GMM with bilingualism as a fixed effect correctly predicts text writing total errors ($F(1, 57) = 6.661, p = 0.012$); phonological errors ($F(1, 53) = 6.736, p = 0.012$); non-phonological errors ($F(1, 55) = 5.313, p = 0.025$); and minor errors ($F(1, 50) = 5.954, p = 0.018$).

Significant positive effects of bilingualism on total error score, phonological errors, non-phonological errors, and minor errors emerged.

Total error score of bilingual children is significantly higher ($M = 0.180$, $SE = 0.019$, $0.145 < CI < 0.223$) than that of monolinguals ($M = 0.123$, $SE = 0.012$, $0.101 < CI < 0.150$).

In particular, bilinguals make significantly

- more phonological errors ($M = 0.093$, $SE = 0.011$, $0.073 < CI < 0.119$) than monolinguals ($M = 0.061$, $SE = 0.007$, $0.049 < CI < 0.076$);
- more non-phonological errors ($M = 0.044$, $SE = 0.005$, $0.036 < CI < 0.054$) than monolinguals ($M = 0.032$, $SE = 0.003$, $0.026 < CI < 0.039$);
- more minor errors ($M = 0.076$, $SE = 0.007$, $0.063 < CI < 0.092$) than monolinguals ($M = 0.055$, $SE = 0.005$, $0.046 < CI < 0.067$).

Significant contrasts are in Table 20.

Dependent variable	Contrast		Est.	StdErr	t	df	p	CI	
								Lower	Upper
Total error score	MONO	BI	- 0.056	0.023	- 2.470	57	0.017	- 0.102	- 0.011
Phonological errors	MONO	BI	- 0.032	0.013	- 2.442	53	0.018	- 0.059	- 0.006
Non-phonological errors	MONO	BI	- 0.013	0.006	- 2.235	55	0.030	- 0.024	- 0.001
Minor errors	MONO	BI	- 0.021	0.009	- 2.389	50	0.021	- 0.039	- 0.003

Table 20. Significant error differences predicted between monolingual and bilingual 2nd graders in text dictation.

To sum up the outcomes of literacy tests, the longitudinal analysis of sublexical decoding (administrations 1-1 and 2-1) revealed comparable sublexical decoding speed by monolingual and bilingual children. Bilingualism showed a positive effect on errors in the 1st grade, whereas monolingualism a negative effect in the 2nd grade but the difference between groups was not significant.

Cross-sectional analyses of word/non-word decoding in the middle of the 2nd grade confirmed comparable decoding speed and non-word decoding accuracy but significantly lower accuracy of bilinguals in decoding generic word lists and non-paroxytone words.

The longitudinal analysis of text decoding skills (administrations 1-2 and 2-1) revealed a significant advantage of monolingual children in both speed and accuracy.

The longitudinal analysis of reading comprehension (administrations 1-2 and 2-1) showed a significant advantage of monolinguals in central comprehension, but no significant difference between groups in specific comprehension.

The longitudinal analysis of word writing under dictation (administrations 1-1, 1-2, and 2-1) revealed significantly less accuracy of bilinguals in the 1st grade but a comparable performance as monolinguals in the middle of the 2nd grade. The cross-sectional analysis of non-word writing in the middle of the 2nd grade confirmed no significant difference between monolinguals and bilinguals.

The cross-sectional analysis of text writing under dictation in the middle of the 2nd grade showed significantly less accuracy by bilingual children. In particular, bilinguals made significantly more phonological errors, non-phonological errors, and minor errors than monolinguals.

3.4. Discussion 1

3.4.1. *Language skills*

The longitudinal analysis revealed that phonological skills of bilingual children do not differ significantly from those of monolinguals (the difference is marginally non-significant ($p = 0.055$)). Phonological skills of both groups followed a parallel developmental pattern and improved significantly in the 1st grade.

Bilinguals' receptive vocabulary, instead, as well as their syntactic abilities (both receptive and productive), were found to be lower than those of monolinguals, from the last year of kindergarten to the 2nd grade of primary school. Cross-sectional tests showed that bilinguals'

productive vocabulary was also significantly smaller than that of monolinguals until the 2nd grade. A significant enlargement of their receptive vocabulary was not present between kindergarten and 1st grade but only between kindergarten and 2nd grade, thus showing that the instructional input they found in primary education was crucial for the development of their vocabulary. There is a body of research aiming at proving the effects of vocabulary on school achievements (for a review, see Elleman et al., 2009; for an intervention on 1232 primary school children, see Vadasy et al., 2015) but, to the best of our knowledge, no longitudinal study showing the effects of instructional input on 1st and 2nd graders. What we know is that preschoolers' vocabulary size is mainly determined by parents' vocabulary (for studies on American English children, see Hart & Risley, 1993; 1999; 2003; Wells, 1985). Given the basic observation that "one cannot acquire words that are not encountered"²⁹, it has been suggested that much of the difference among children in vocabulary acquisition is "a matter of opportunity rather than ability" because "children with small vocabularies do acquire new words during instruction about as well as children with larger vocabularies" (Biemiller, 2005: 6). We also know that, for the majority of children in this age coming from minority-language families, school is the only opportunity to encounter Italian input³⁰, and that every child of our sample had attended 2 or 3 years of Italian kindergarten. Since a significant improvement in children's vocabulary was found between the middle of the 1st grade and the middle of the 2nd grade but not between the middle of the last year of kindergarten and the middle of the 1st grade, without any difference between the patterns of monolinguals and bilinguals (see Figure 7), we can say that the instructional input found in primary school, which also entails a certain metalinguistic component, must have played a role in the development of children's receptive vocabulary.

²⁹ This is true for root word meaning but not for derivations and compound words.

³⁰ In our parents' questionnaire, 7 families declared to use Italian in association with their mother tongue at home.

Syntactic comprehension abilities significantly increased between kindergarten and 2nd grade, but we do not have the intermediate measure of 1st grade. We can observe, however, that syntactic production of direct object clitic pronouns does not improve significantly between kindergarten and the 2nd grade if we gradually increase the test difficulty. Direct object clitic pronouns are a complex domain because their use requires a range of phonological, morphosyntactic, syntactic, and pragmatic skills. A cross-linguistic research on 16 languages (Varlokosta et al., 2016) shows that clitics should be already acquired in typically developing monolingual children aged 5 years. After this age, high omission rates might be due to language atypical development (see also Tuller et al., 2011). Italian direct object clitics, in particular, are candidate clinical markers of specific LI (Arosio et al., 2010; 2014; Bortolini et al. 2002, 2006). Recent studies show, however, that the specificity of this marker may have been overrated (Casani, 2020a, Tuller et al., 2011). Errors in clitic production appear in several other atypical developmental conditions (for LI with borderline intellectual functioning, see Casani, 2020a; for developmental dyslexia, see Cardinaletti & Casani, 2019; Guasti, 2013; and Zachou et al., 2013; for the deaf population, see Tuller, 2011 and Tuller & Jakubowicz, 2004; for Down syndrome, see Rondal & Comblain 1996 and Sorianello, 2012; for high functioning autism, see Meir & Novogrodsky, 2019). Moreover, testing clitics is highly sensitive to elicitation conditions. Casani & Cardinaletti (2019) showed that both monolingual and bilingual children in the last year of kindergarten and primary school produce significantly more target clitics under the condition of gender matching between the object, namely the clitic antecedent, and the sentential subject. In the same study, a small sample of primary school children (age $M = 9;10$; $SD = 0;10$) with developmental language disorders ($N = 3$), developmental dyslexia ($N = 3$), and typical development ($N = 2$) produced significantly more target clitics when the stimulus question contained a lexical subject compared to a null subject. For this reason, the original test was modified (Casani, Cardinaletti, & Volpato, in preparation) to present balanced conditions of

gender and number matching and mismatch. The longitudinal analysis of the present sample, besides confirming a particular difficulty of bilingual children in clitic production, confirmed clitic agreement to be more difficult under the condition of gender mismatch with the sentential subject, in particular with a feminine antecedent³¹. The gradual complexification of the test (presenting gender and number matching conditions in the first administration (KG-1); 47% conditions of gender mismatch with 100% of number matching in the second administration (KG-2); and balanced conditions of gender and number matching and mismatch in the third administration (2-1)) might explain the reason why no significant improvement of either monolingual and bilingual children was observed between the last year of kindergarten and the 2nd grade of primary school in the present study.

Cross-sectional analyses found that monolinguals produce significantly more correct clitic pronouns than bilinguals, but no significant improvement of either monolinguals or bilinguals was observed between the 1st and 2nd grade by using the most recent test.

A fine-grained analysis of cross-sectional responses by children tested with the most recent version of the test in 1st (N = 49), 2nd (N = 36), and 3rd (N = 9) grade, however, revealed that a significant increase of target clitics and a significant decrease of clitic errors are associated to each school grade (Cramer's $V = 0.208$, $p < 0.001$)³². In the 1st grade, in particular, 14% of clitic errors and 12% of clitic omissions were still present, whereas the proportion of target clitics by typically developing monolingual children was only 60%. This is far from considering direct object clitics to be fully acquired at age 4 (see, for instance, Arosio et al. 2014; Belletti & Guasti, 2015; or Hyams & Schaeffer, 2008, after Schaeffer, 2000, who found

³¹ Data on the new clitic-elicitation test were preliminarily presented by Casani in a pilot joint study with Anna Cardinaletti in the online Conference BeMeLAcq 2020 (Behavioural Measures of Language Acquisition and language processing in typical and atypical populations), Ca' Foscari University of Venice (Venice Accessibility Lab) and NTNU Trondheim (Language Acquisition and Language Processing Lab). The presentation is available at the link <https://www.youtube.com/watch?v=EPNgOZaNoBs>

³² The above-mentioned pilot study demonstrated that such a fine-grained analysis should include among target answers only clitics that are pragmatically adequate, beyond being grammatically correct.

100% of target clitics in the elicited production of 77 4-year-old and 72 5-year-old children). This difference, besides suggesting the evaluation of diverse contexts, supports the opportunity of including balanced syntactic and morphosyntactic conditions in testing clitics, to create the most appropriate elicitation context.

Bilingual children confirmed significant problems in clitic production compared to monolinguals (see also Chondrogianni et al., 2015; Karpava, 2017). The cross-sectional response analysis showed a mean of 31% target clitics, with a peak of 71% omissions by bilingual typically developing children. This is strictly dependent on the family language, as children with a Romance mother language showed 77% and 76% accuracy in the most accessible conditions, namely masculine singular (MS) clitic with masculine singular and masculine plural (MP) subject, respectively, versus 42% (MS – MS) and 39% (MP – MS) of children with a non-romance mother language. Casani (2020a) found that, in addition to resorting to full DPs (12%), bilingual children in the last year of kindergarten replaced or omitted clitics, thus producing significant rates of ungrammatical sentences including omissions (20%), infinitive verbs instead of main sentences (17%), and proto-morphemes (4%). Age-matched monolingual children, instead, used more refined strategies, such as clitic elision (11%), which allow them for the production of grammatically correct sentences that are not completely adequate from a pragmatic point of view, as a change of the verb tense from present continuous to present perfect is required (see Example 1).

Example 1

Question: *C'è un poliziotto che sta arrestando un ladro. Guarda, cosa sta facendo il poliziotto al ladro?*

(There is a policeman that is arresting a thief. Look, what is the policeman doing to the thief?)

Expected answer: **Lo**_{MS_dir_obj_cl} *sta arrestando*. ([Ø] is arresting **him**_{MS_dir_obj_cl})

Elided clitic: **L'**_{MS/FS_dir_obj_cl} *ha arrestato*. ([Ø] has arrested **h[im]**_{MS/FS_dir_obj_cl})

Although the answer does not completely adhere to the elicitation context, it is more appropriate than one containing a full DP, because the clitic pronoun is present but without overt morphological agreement. This might be considered as an avoiding strategy, which was also reported in Italian monolingual children with developmental dyslexia and with developmental language disorder (age 6;7 – 9;6) (Cardinaletti & Casani, 2019), and in typically developing French monolingual children (age 4;7 – 8;5) (Délage et al., 2016). In the present sample, this strategy was adopted by monolingual children with typical development (35%) and with language disorders³³ (31%), and by bilingual typically developing children (35%), whereas it was absent in 2 bilingual 2nd-graders with language disorders, one of whom with mild intellectual disability. This might indicate that bilingualism matched with a language deficit and non-brilliant intellectual functioning might hinder access to this kind of strategy. When children employ it, however, the past-participle agreement is generally correct, thus showing local agreement to be in place, and the problem to lie in the distance of the antecedent. An analysis of bilingual children with minority languages in the last year of kindergarten (Casani, 2020a) also showed that a Romance native language, as well as the use of Italian along with their mother tongue in their family, was significantly associated with better performance (significantly more target clitics) and with the use of more refined strategies (grammatically correct even though pragmatically inadequate). This suggests that the clitic production deficit among bilingual children is not due to a delay in their language development but to their reduced exposure to the Italian input (see also Karpava, 2017). These data show that, while bilinguals' receptive and productive vocabulary, as well as their syntactic skills between the last year of kindergarten and the 2nd grade, are kept lower compared to those of age-matched monolinguals, their phonological skills between the last year of kindergarten and the 1st grade of primary school can be comparable, thus

³³ Children with language disorders were excluded from the present study and are analyzed in Study 4.

suggesting the possibility to use the non-word repetition test to detect possible phonological deficits in bilingual children (Coady & Evans, 2008). It is worth noticing that there might be marginal differences with monolingual children, probably due to a greater familiarity of monolinguals with the structural properties of the Italian lexicon, which could not be the dominant one for bilingual children (see Summers et al., 2010). These differences might emerge as pseudowords are tested that are modeled on phonotactic rules of the non-dominant language and might be affected by the frequency of strings used.

In sum, the development of phonology, vocabulary, and (morpho)syntax proceeds in a parallel manner among monolinguals and bilinguals, but this might not be evident in the production of complex structures like clitic pronouns. In this case, a fine-grained response analysis is necessary to detect children's strategies, which can determine ungrammatical responses or different grades of grammatical but pragmatically inadequate productions, according to their extent of development.

3.4.2. Metaphonological skills

Global metaphonological skills

Global metaphonology, namely syllable awareness, was tested twice in the last year of kindergarten, except for rhyme recognition, which was tested 5 times (twice per year). Monolingual children showed an advantage in every ability except syllable segmentation, where no difference from bilinguals emerged. The strongest differences were in the first administration (KG-1) of rhyme-recognition, word, and non-word minimal pairs, meaning that monolinguals get a head start particularly in these abilities. The strongest positive effects of monolingualism were on rhyme recognition, followed by word minimal pairs. This was probably due to the lexical advantage of monolinguals, as shown by the vocabulary tests and by the strong correlation of rhymes ($\rho_s = 0.671$, $p = 0.000$) and the moderate correlation

of word minimal pairs ($\rho_s = 0.460$, $p = 0.014$), respectively, with receptive vocabulary. The quite strong positive effect of monolingualism in the first administration of non-word minimal pairs (which do not correlate with vocabulary measures) might depend on a greater familiarity of monolinguals with the structural properties of the Italian lexicon (see Summers et al., 2010), which might also be responsible for the marginally non-significant difference shown by the groups in the non-word repetition test. It is plausible that the marginally significant difference between monolinguals and bilinguals in syllable blending reported at the end of kindergarten, as opposed to the equivalent performance reported, instead, in syllable segmentation, also depends on a sort of monolingual phonotactic advantage, which makes blending syllables into words more accessible for monolinguals. In this perspective, the syllable segmentation task presents a higher degree of abstractness, albeit not as high as at the phonemic level (Guedens, 2006), which might make this task less advantageous for monolingual children compared to syllable blending. Figure 9 shows that syllable segmentation scores were, in general, a bit lower than syllable blending scores, and monolingual children had a numerically lower performance than bilinguals in the first test session (KG-1), whereas, in every other ability, monolinguals' scores were numerically higher than those of bilinguals. The comparable difficulty of syllable segmentation for the two groups might make it a possible tool for testing metaphonology in monolingual and bilingual kindergarteners without linguistic bias.

Another strong positive effect of monolingualism was found on the recognition of the initial syllable, where the difference between groups got significant in the second administration. According to the *Model of Phonological and Phonetic Encoding* (Levelt, 1989, 1992, 1993; Levelt & Wheeldon, 1994), syllable-sized *chunks* or *schemas* (for the difference, see Seevald et al., 1995) of the language are stored in a *mental syllabary*, the access being regulated by frequency principles, as well as the access to the lexicon, which is, however, independently stored. Hence, the more a syllable is used the more it is accessible. In our

case, monolingual children take advantage of their use frequency of Italian, which constitutes 100% of their input and output, thus having easier access to the mental syllabary. Rhyme-recognition also showed a positive effect of monolingualism, which determined a significant difference in the first three administrations, namely until the middle of the 1st grade, compared to bilinguals. A general significant improvement in rhyme recognition was observed in the middle of the 1st grade and of the 2nd grade. Rhyme sensitivity is deemed as a crucial ability for the acquisition of metaphonology because it can develop naturally in prereaders and illiterate adults, irrespective of specific training and letter knowledge (Guedens, 2006; Morais et al., 1979), whereas the development of explicit phonemic awareness requires, at least, a rudimental print experience or specific training (Guedens, 2006; Johnston et al., 1996). Our data showed that the recognition of rhymes develops parallelly in monolingual and bilingual children between the last year of kindergarten and the 2nd grade of primary school, but bilinguals' proficiency is at a lower level until the middle of the 1st grade. Input processing requires the activation of multiple competing words including onset competitors, which share their beginning sound (e.g., **cane**, **casa**), and rhyme competitors, which share their end (e.g., **naso**, **vaso**) (for the *onset-rime* syllable structure theory, see Fudge, 1987; and Treiman & Kessler, 1995). Through an eye-tracking experiment on a group of adult migrants who were proficient in both Dutch and English, Bruggeman & Cutler (2019) found that bilinguals were prone to discard rhyme competitors in favor of onset competitors, probably because their working memory was overloaded by the activation of onset competitors (which appear earlier in speech) from both the L1 and the L2 (McDonald et al., 1992; Service et al., 2002). If this happens in proficient adult L2 speakers, it might be true for preliteracy children, too. A qualitative analysis of bilingual children's responses in the rhyme-recognition task might reveal whether they tend to choose onset competitors instead of rhyme competitors, or if they adopt different strategies (for instance, semantic strategies). The difference between monolingual and bilingual children

in rhyme recognition, however, is not significant anymore at the end of the 1st grade, probably for the combined effect of increased language skills and executive functions. The development of children's language skills thanks to the instructional input (see 3.4.1), in particular, might allow children for the allocation of an effective load of working memory in the inhibition of onsets during the rhyme recognition task (see also 3.4.3).

Analytical metaphonological skills

After the 3 administrations of analytical metaphonology tests (in 1st and 2nd grade), monolinguals revealed higher proficiency in the phonemic verbal fluency task. The reason can lie in that this ability is strictly related to the indicators of vocabulary size and lexical-access speed, as confirmed by the correlations with receptive vocabulary ($\rho_s = 0.509$, $p = 0.008$) and with RAN speed (digits: $\rho_s = 0.536$, $p = 0.000$; colors: $\rho_s = 0.493$, $p = 0.000$; pictures: $\rho_s = 0.358$, $p = 0.000$), as well as to inhibitory control, thus showing a hybrid nature that suggests a cautious use of its measures (Shao et al., 2014). Phonemic verbal fluency underwent a significant improvement between the middle and end of the 1st grade for both monolinguals and bilinguals, so that their performance was comparable, but in the middle of the 2nd grade, monolinguals outperformed bilinguals again and more markedly.

The two groups did not reveal any difference in phoneme blending, thus showing that bilinguals' disadvantage in (syllable) sound blending was completely overcome at the beginning of the 1st grade. The development of this skill proceeded parallelly in the two groups, showing a significant improvement at every test session. The most important developmental step was in the first grade, as the positive effect between the middle and the end of the 1st grade was 4.4 times greater than that shown between the end of the 1st grade and the middle of the 2nd grade. This shows that analytical blending develops in parallel with early literacy learning. As for the other analytical metaphonological skills, monolinguals were advantaged in the first administration of the phoneme-segmentation test. Italian

monolinguals, who did not show any advantage in syllable segmentation, seemed to acquire phoneme segmentation earlier than their bilingual fellows, but this difference disappeared at the end of the 1st grade (see Figure 10 and Table A.4). The initial disparity is probably due to reduced exposure to the Italian input. As mentioned before, analytical phonological awareness develops along with the exposure to the written language (Guedens, 2006; Johnston et al., 1996) or after specific training aiming at developing children's ability to "hear individual sounds in the continuous airflow". This takes place as "they have acquired the symbols which help them to abstract over the highly variable acoustic events" (Guedens, 2006: 37). Bilinguals' reduced exposure to the Italian input makes this process somewhat slower in the very early steps of literacy acquisition, but, according to our data, comparable outcomes are achieved at the end of the 1st grade. Analytical-segmentation abilities developed rapidly and parallelly in the two groups, with a powerful increase in the second half of the 1st grade, which was 3.6 times greater than that shown between the end of the 1st grade and the middle of the 2nd grade.

A combined positive effect of monolingualism and second administration (see Table A.4) at the end of the 1st grade on the deletion of the initial syllable indicates that the development of this ability proceeds a bit faster in monolinguals, even though this did not determine any significant difference between groups. This slightly different pace was not observed in final syllable deletion. This might confirm that in the first steps of literacy acquisition, bilinguals are more prone to discard *rhyme competitors* than *onset competitors* (see above), because the latter, occurring earlier in speech, might be more prominent, especially under conditions of working memory overload. As their (meta)linguistic competence and executive functions interactively develop, bilinguals rapidly get comparable sound-manipulation skills as their monolingual fellows. This is confirmed by the cross-sectional results of a highly complex manipulation task like spoonerism, in which monolingual and bilingual 2nd-graders did not show any significant difference.

These data suggest that, except for slight differences in the acquisition pace of phoneme segmentation and initial-syllable deletion, children's analytical metaphonology in the literacy language develops parallelly in monolinguals and bilinguals, thus achieving comparable levels at the end of the 1st grade. An exception is constituted by phonemic verbal fluency, which is related to vocabulary size and lexical retrieval (as shown by strong relations with receptive ($\rho_s = 0.527$, $p = 0.000$) and productive ($\rho_s = 0.627$, $p = 0.000$) vocabulary, and by the more moderate relations with RAN speed ($0.367 < \rho_s < 0.423$, $p \leq 0.001$) in 1st grade³⁴), thus suggesting that intervention aiming at vocabulary development in the literacy language starting from kindergarten might help improve phonemic verbal fluency in bilingual children.

3.4.3. RAN

Longitudinal scores of the five test administrations of picture RAN (twice per year, from kindergarten to the 2nd grade) and of the three test administrations of color and digit RAN (twice in the 1st grade and once in the 2nd grade) revealed that monolingual children are significantly faster in the RAN of pictures during the last year of kindergarten, until the middle of the 1st grade of primary school, and significantly more accurate than bilinguals during the last year of kindergarten. A significant increase in picture speed was observed in the middle of the 1st grade and of the 2nd grade, whereas picture errors decreased significantly in the middle of the 1st grade of primary school. This means that RAN speed of pictures develops quite slowly and gradually, whereas accuracy should be in place in the 1st grade (Table 9). Concerning RAN abilities that were tested starting from primary school, monolinguals and bilinguals revealed comparable performances in the RAN of colors. Color speed significantly increased in every administration, namely at the end of the 1st grade and in the middle of

³⁴ Interestingly, the relation of phonemic verbal fluency in 2nd grade persists only with receptive vocabulary ($\rho_s = 0.509$, $p = 0.008$), thus suggesting a possible automatization of this ability.

the 2nd grade, whereas a significant decrease of color errors was observed in the 2nd grade, with a mild effect.

Bilinguals were as fast as monolinguals in the RAN of digits, whereas they made significantly more errors in the first administration in the middle of the 1st grade. From the end of the 1st grade, the groups' performances were comparable. Digit speed increased significantly in every administration, whereas digit errors decreased between the middle and the end of the 1st grade, as long as children acquired numeracy thus consolidating their familiarity with digits. The negative effect of bilingualism and first administration (1-1) on speed in the RAN of digits indicates some initial difficulties of bilinguals in this measure, which are, however, early overcome, as no significant difference between groups' performances in this measure of RAN emerged.

These data suggest that colors constitute the most accessible RAN measure for bilingual children, which might be used as a possible testing tool since the last year of kindergarten, whereas the RAN of pictures and digits could determine some differences that do not depend on a RAN deficit in bilinguals but on linguistic reasons. Starting from the end of the 1st grade, however, all RAN abilities of typically developing monolingual and bilingual children should be at a comparable level. The present outcomes are not completely in line with the literature, which describes significant deficits in picture naming tasks by bilingual children (Poarch et al., 2012), even having comparable receptive vocabulary as monolingual controls³⁵ (Poulin-Dubois et al., 2013), and by bilingual adults (Gollan et al., 2005; 2008; 2011; Ivanova & Costa, 2008). These deficits have been attributed to language frequency, namely to the fact that bilinguals use words from each language less frequently than their monolingual counterpart (for the *frequency lag account*, see Gollan et al., 2005; 2008; 2011), or to cognitive reasons, namely to the contemporary activation of cross-language conflicting

³⁵ For vocabulary measures of the present sample, see section 3.3.1.

competitors determining slower naming latencies (for the *competition account*, see Abutalebi & Green, 2007; Green, 1998; Kroll & Gollan, 2014), or to a combination of both linguistic and cognitive reasons (Sullivan et al., 2018). None of the mentioned studies, however, analyzed RAN longitudinally. The present data might demonstrate that bilingual children with lower receptive and productive vocabulary than monolingual controls matched for age, social, and educational environment, can match their peers in color RAN since kindergarten and, with a slightly different pace, can get comparable results in other RAN measures within the end of the 1st grade of primary school. Further experiments might be designed by controlling for cognitive, psycholinguistic (e.g., verbal IQ and reaction times), and language variables (e.g., mother tongue, proficiency level in the two languages, knowledge of more languages). Controlling for language variables, in particular, might reveal differences due to the presence of word cognates (Costa et al., 2005), which might help cross-language activation in children that are proficient in more than two languages (Poarch et al., 2012).

3.4.4. Short-term memory

Bilingual children showed comparable digit span forward as monolinguals but smaller digit span backward in kindergarten and 1st grade, and smaller word span in kindergarten and 2nd grade, thus revealing reduced verbal working memory until the 1st grade, and a generally reduced word span compared to their monolingual fellows. The word span did not correlate with vocabulary in kindergarten but did moderately correlate with receptive vocabulary in 1st grade ($\rho_s = 0.454$, $p = 0.001$), and robustly in 2nd grade ($\rho_s = 0.605$, $p = 0.001$); at the same time, it correlated robustly with digit span forward in 1st grade ($\rho_s = 0.506$, $p = 0.000$) but weaker in 2nd grade ($\rho_s = 0.388$, $p = 0.034$). So, the word span tended to lose its relationship with digit span forward, which is also a measure of verbal short-term memory, and to

increase its relationship with the vocabulary size as long as vocabulary increases. This might suggest that younger children use a heavier load of short-term memory to accomplish the word-span task, whereas they progressively tend to rely on lexical strategies as they grow up. Bilinguals' smaller word span might then be due to their reduced vocabulary (see 3.3.1 and 3.4.1), thus supporting the idea that timely lexical training might have positive fallout on bilinguals' skills.

These data might constitute a further argument (see 1.8) versus the idea that bilingual children outperform monolinguals in a range of executive functions (for a review arguing no systematical evidence for the *bilingual executive advantage* in adults, see Lehtonen et al., 2018), at least for what concerns working memory (for a review arguing in favor of the *bilingual executive advantage* in the domain of working memory, see Adesope et al., 2010; Grundy & Timmer, 2017). In this vein, recent studies found that working memory span is affected by language abilities, in particular by sentence comprehension (Service et al., 2002) and lexical decision (Bruggeman & Cutler, 2019). Bruggeman & Cutler's (2019) hypothesis based on the competition account has been assumed as a possible explanation of the initial disadvantage of bilingual children in the rhyme-recognition test (see 3.3.2 and 3.4.2). Notice that bilinguals' scores in non-word repetition, a task involving a load of phonological working memory (see 2.1.3), were kept numerically lower than those of bilinguals and this difference was marginally non-significant ($p = 0.055$).

Other aspects of working memory should be taken into consideration to get an exhaustive profile of children's abilities. In the next section, we analyze the test of graphic reproduction, which puts into play visual working memory beyond graphic skills.

3.4.5. Visual memory and discrimination

Visual tests were administered twice in kindergarten and once in the 1st grade of primary school. Bilingual and monolingual children showed comparable visual skills in the test of reproduction of one and two semicircles, which sets in place a prevalent component of visual memory and a moderate component of graphic abilities. A significant improvement in the reproduction of single circles was observed between the two kindergarten administrations, whereas the more demanding task of reproduction of two semicircles improved significantly at every administration. The ability to visually discriminate the position of capital letters oriented in different ways showed a marginally significant effect of bilinguals' errors in the first administration (KG-1). This was not sufficient, however, to determine a significant difference between groups, thus revealing only a possible lack of familiarity of minority-language children with the Latin alphabet compared to Italian monolinguals. A significant improvement of this ability was observed only in kindergarten, thus showing it to be completely acquired by both monolingual and bilingual typically developing children when entering primary school.

These data suggest no bilingual advantage in a graphic task involving visual working memory (see (3.4.4), as bilinguals' outcomes were not significantly different and were rather maintained numerically lower than those of monolinguals. Furthermore, we confirm that non-verbal aspects of short-term/working memory, like those involved in the graphic reproduction of simple pictures after brief observation, are at the same level in monolinguals as in bilinguals, thus confirming verbal working memory to be affected by language skills (see 3.3.4).

3.4.6. Literacy skills

Sublexical decoding

Sublexical decoding was tested in the middle of the 1st grade and in the middle of the 2nd grade. The longitudinal analysis revealed comparable sublexical decoding speed by the two groups but some accuracy differences. Bilingualism showed a positive effect on errors in the 1st grade, whereas monolingualism a negative effect in the 2nd grade. This means that both groups improve their accuracy between the first and second administration, but bilinguals start with a disadvantage, and monolinguals tend to be more accurate in 2nd grade.

Talli & Emmanouil (2020) compared the cases of 2 Greek-Italian bilingual male adults aged 22 – 26 with 2 Greek monolingual adults with and without developmental dyslexia in word and non-word decoding and found the typically developing monolingual to outperform the typically developing bilingual. Interestingly, the bilingual dyslexic subject outperformed the monolingual dyslexic subject. The authors also refer both bilingual subjects to outperform monolinguals in a non-word repetition task, differently from our sample, where we found no differences in non-word repetition (see 3.3.1). Furthermore, they observed the Greek language dominance to affect Italian decoding but not non-word repetition, where children performed better in Italian than in Greek. The authors attribute the reason to the phonotactic complexity of the Greek orthography, thus supposing possible independence of phonological skills from language dominance and orthographic transparency. This would imply a sort of moderating power of bilingualism on decoding deficits in dyslexia but no facilitative effect in typical development, a hypothesis that needs further exploration in wider samples. It might be possible that their bilingual sample started with better phonological skills, or that bilingualism offers a phonological advantage only in the long term when both languages are well acquired. Meanwhile, similar results as in our study have been described, for instance, by Bonifacci et al. (2017) in a mixed context. They explored literacy skills in English as a foreign language in a sample of 19 monolingual children with developmental dyslexia, compared to 19 typically developing bilingual language-minority children, and 76 Italian monolingual controls in 4th and 5th grade of primary school. As in our

case, they found typically developing bilingual children to have comparable decoding speed but lower accuracy than monolinguals in word and non-word decoding, and word writing. Interestingly, the two samples did not differ in either speed or accuracy in English tasks, thus showing a sort of bilingual benefit to be evident not in their second language, but in learning a third language, even though this language has a deep orthography.

Word and non-word decoding

Cross-sectional results in the middle of the 2nd grade revealed that the decoding speed of bilingual children was comparable to that of monolinguals for both words and non-words. Moreover, non-word decoding accuracy was also in place, thus confirming that bilinguals' difficulties with sublexical decoding are overcome (see above). Bilinguals, however, were significantly less accurate in decoding generic word lists and words with atypical stress, namely non-paroxytones. These problems are not to be attributed to any deficit in the bilinguals' lexical route of reading but, probably, to their smaller vocabulary, as shown by the lower performance in vocabulary tests (see 3.3.1). The decoding of generic word lists, however, does not correlate with vocabulary tests. This might indicate that bilinguals' problems do not lie in children's vocabulary size but in their readiness in lexical retrieval and familiarity with structural properties of the Italian lexicon, as shown by the strong correlations of word decoding with the phonemic-verbal-fluency task ($\rho_s = -0.579$, $p = 0.001$) and with spoonerisms ($\rho_s = -0.512$, $p = 0.000$), and by the more moderate correlations with (morpho)syntactic tests (comprehension: $\rho_s = -0.472$, $p = 0.015$; production: $\rho_s = -0.384$, $p = 0.040$). The phonemic verbal fluency task, as a measure of the word-retrieval ability after phonemic hint, and spoonerism, as a measure of the ability to manipulate word chunks to create new words, require acquired familiarity with the Italian lexicon and its structural properties, in particular with *onset-rimes* (see 3.4.2) and with syllable manipulation. Furthermore, the moderate relation of word decoding errors with syntactic comprehension

(generic errors: $\rho_s = -0.472$, $p = 0.015$; stress errors: $\rho_s = -0.557$, $p = 0.003$) and production of direct object clitic pronouns (generic errors: $\rho_s = -0.384$, $p = 0.040$; stress errors: $\rho_s = -0.556$, $p = 0.002$) might confirm a role of (morpho)syntactic abilities not only in reading comprehension but also in written decoding. The effect of morphological information on decoding accuracy and speed has been suggested by different studies (for Italian, see Angelelli et al., 2014; Burani, 2010; Burani et al., 2002). Casani (in preparation) analyzed text decoding in a sample of Italian children with and without developmental dyslexia in primary school (from the 1st to the 5th grade) and found that a considerable amount of variance in text decoding accuracy was explained by syntactic-comprehension skills, followed by syntactic production of direct object clitics, together with 3 decoding error categories that have to do with morphology to different extents, namely morphological errors, semantic errors, and phonological-output-buffer³⁶ errors (see 1.3.1; for a comprehensive description of decoding error categories associated to different types of developmental dyslexia, see Friedmann & Coltheart, 2018). In Study 2, an analysis of the effects of children's prerequisites on literacy will help disentangle whether the mentioned related abilities confirm an actual role in children's decoding.

Bilingual children did not show any significant difference compared to monolinguals in attentional errors. This shows that the peripheral stage of the reading model, which is responsible for processes of visual-orthographic analysis (see Friedmann & Coltheart, 2018), is not affected by bilingualism, and confirms, more in general, that visual skills involved in decoding are equivalent in bilinguals and monolinguals (see 3.3.5 and 3.4.5).

³⁶ Preliminary results of the mentioned study were presented at the 19th International Morphology Meeting "Morphology in Contact" (Vienna, 6 – 8 February 2020), and at the National Congress AIRIPA (Associazione Italiana per la Ricerca e l'Intervento nella Psicopatologia dell'Apprendimento) (online, 25 – 26 September 2020). In these studies, multiple errors in long and morphological complex words were coded as phonological -output-buffer errors. Further analyses are in progress to ascertain whether these errors actually depended on a phonological -output-buffer phonological dyslexia (see Friedmann & Coltheart, 2018).

Text decoding

The longitudinal analysis of text decoding skills, tested at the end of the 1st grade and in the middle of the 2nd grade, revealed an advantage of monolingual children in both speed and accuracy. Text decoding seems to be a more demanding activity than word decoding, where bilingual children were less accurate but not less fast. Similar reasons as for word decoding might be responsible for problems in text decoding accuracy (see above). Moreover, the context offered by the text might play a facilitating role for monolingual children, who have a larger vocabulary and better syntactic skills than bilinguals (see 3.3.1 and 3.4.1) and can easier recur to orthographic and lexical strategies that allow them for faster decoding.

Reading comprehension

A longitudinal analysis of the two administrations of the reading comprehension test revealed a disadvantage of bilinguals. Total scores, as well as central comprehension scores, differed significantly, whereas no difference was found for specific comprehension scores. This could mean that smaller vocabulary, weaker syntactic comprehension skills (see 3.3.1), as well as weaker decoding skills (see above), and reduced working memory capacity (see 3.4.4) of bilingual children are sufficient to get a partial comprehension of text literal aspects, at least in the first grades, but can have a deeper impact on the inferential operations required by more demanding texts. If so, unless an improvement of bilinguals' language and decoding skills, a subsequent increase of bilinguals' problems in reading comprehension is expectable, due to the increasing difficulty of texts proposed in the school. This may be suggested by the fact that only monolingual children showed a significant improvement in their central comprehension at the second test session (Figure 18).

Word and non-word writing

Monolingual and bilingual children showed similar results in non-word writing under dictation, whereas bilinguals had a significantly lower performance than monolinguals in word writing after the two test sessions in 1st grade. As we see in Figure 19, bilinguals made a higher rate of errors in the first administration (1-1), which decreased significantly (see Table 17) at the end of the 1st grade, to get a comparable performance as monolinguals in the middle of the 2nd grade. It is worth mentioning that 1st-grade tests (Stella & Apolito, 2004) were administered collectively, and dictation took place without repetition and with standard pace (7 seconds per word), according to the standard procedures (see 2.12). The 2nd-grade dictation of words and non-words (Sartori et al., 2007), which is longer and more complex, was administered individually, with the possibility to listen to words as many times as needed for a correct acoustic analysis. The test standard procedure includes the repetition of the word/non-word before writing it. This procedural variation might influence the performance, thus affecting the longitudinal analysis³⁷. The between-group analysis, however, tells us that in the 1st grade, under the condition of collective word dictation, bilingual children have a lower word-writing performance than monolinguals. This performance improves at the end of the 1st grade but is still at a lower level. In the middle of the 2nd grade, on the contrary, the word and non-word writing performance of the two groups, under the condition of individual dictation and after ascertaining the word/non-word correct acoustic analysis, is at the same level.

The 2nd-grade text dictation, which was administered with a similar procedure as the 1st-grade word dictation, might help disentangle whether the good performance of bilingual 2nd-graders might be related to the administration conditions.

³⁷ This issue might be disentangled by administering word, non-word, and text dictation with both procedures (collective without repetition and individual with repetition).

Text writing

The performance of bilingual children in text dictation confirmed to be significantly less accurate than that of monolinguals. Bilinguals made significantly more phonological errors, non-phonological errors, and minor errors than monolinguals. Phonological errors were not expected because bilinguals had similar phonological skills (see 3.3.1. and 3.4.1) and showed a similar non-word writing performance (see 3.3.6 and 3.4.6) as their monolingual fellows, thus revealing a proper functioning of the sublexical route of writing. The problem might then lie in their reduced vocabulary. In particular, phonological errors correlate with productive vocabulary ($\rho_s = -0.448, p = 0.022$) and minor errors with receptive vocabulary ($\rho_s = -0.436, p = 0.026$). Moreover, phonological ($\rho_s = -0.503, p = 0.022$) and minor errors ($\rho_s = -0.407, p = 0.039$) correlate with (morpho)syntactic comprehension. This does not mean that these related skills are the direct cause of all these errors, but the reduced familiarity with structural properties of lexicon and morphology (see 3.3.1) might hinder the passage of the acoustically analyzed strings to the lexical route through the phonological input lexicon, thus making writing a higher demanding task for bilinguals, which turns to be inaccurate particularly with meaningful words and text³⁸. The working memory overload (see 3.3.4 and 3.4.4) involved in these processes, in particular under the condition of time restriction and without repetition as in the collective dictation, might explain the higher rate of minor errors (punctuation and capital letter errors)³⁹, which have not to do with language skills⁴⁰ but are more numerous in bilinguals, anyway.

³⁸ Similar reasons might underlie inaccuracy in word decoding and slowness and inaccuracy in text decoding by bilingual children (see above).

³⁹ This category, according to the standard procedure, was not included in the general error score, which is composed of the 3 main error categories (phonological, non-phonological, stress/geminate errors).

⁴⁰ It might be interesting to explore whether these error types can be related to children's familiarity with books, newspapers, and other written materials in the instructional language.

3.5. Interim summary 1

We described the longitudinal development of preliteracy and literacy skills of typically developing monolingual and mono-literate bilingual children with minority languages from the last year of kindergarten to the 2nd grade of primary school.

We found equivalent phonological skills but significant differences in vocabulary and (morpho)syntactic skills. The instructional input seems to have a decisive role in the development of children's vocabulary size. Language development of monolinguals and bilinguals proceeds in a parallel manner, but this might not be evident in some complex areas. For direct object clitic pronouns, in particular, a qualitative analysis of productions is necessary to distinguish the extent of development of different language skills involved in their use. We highlighted that these structures are particularly sensitive to the production context and suggested a test with balanced morphosyntactic conditions for their elicitation (Casani, Cardinaletti, & Volpato, in preparation).

Syllable segmentation might be an affordable testing tool of global metaphonology since kindergarten because it is not biased from children's language skills, thus allowing for the comparison of monolingual and bilingual performances. The majority of other metaphonological skills, except for slight differences in the acquisition pace, achieve comparable levels between the middle and the end of the 1st grade.

We showed that verbal working memory is affected by language abilities and suggested that lexical training could be useful since kindergarten to lighten the burden of children's working memory engaged in activities that require complex or automatized language processes, such as phonemic verbal fluency. Bilinguals' visual and phonological memory were equivalent to those of monolinguals, and we did not find any trace of a bilingual advantage in activities involving the use of visual working memory and high-level executive functions, such as graphic reproduction of semicircles after brief observation and RAN.

In the 2nd grade, bilinguals reached a comparable level of sublexical decoding skills as monolinguals but kept accuracy problems related to their smaller vocabulary, such as in decoding words with atypical stress. Their language skills were sufficient to reach a literal comprehension but were not enough for a deeper comprehension of written texts.

Non-word writing skills were equivalent, but word writing (in the 1st-grade collective dictations) and text writing (in 2nd-grade collective dictation) were less accurate in bilinguals. Pressing conditions required by the collective dictation might affect children's writing skills, as shown by the higher rate of errors that have not to do with language. Alternative procedures could be advisable for writing assessments during school activities.

4. Study 2

Effects of short-term memory, language, metaphonology, and RAN on the development of literacy in typically developing monolingual and mono-literate bilingual children in an Italian school setting

4.1. Aims 2

This study explored the longitudinal effects of different preliteracy skills on the development of early literacy in a sample of Italian monolingual children compared to mono-literate bilingual children belonging to migrant families.

We also aimed to investigate whether RAN accuracy in addition to RAN speed has effects on the decoding speed of typically developing children (Casani, 2020b) in the early stages of Italian literacy acquisition, whether this might influence decoding accuracy (Talli et al., 2015; Vaessen et al., 2009), and what the relationship is with reading comprehension.

We examined whether different levels of phonological awareness (e.g., syllabic/phonemic) and of different semantic categories of RAN (e.g., pictures/colors/numbers) (Denckla & Rudel, 1974) differentially contribute to reading development in a shallow orthography.

We investigated possible effects of (morpho)syntactic competence (comprehension and production) on written decoding (Angelelli et al., 2014; Burani, 2010; Burani et al., 2002; Casani, in preparation), and written comprehension (Casani, 2020b; 2020c; Huggins, 1977; Morvay, 2012; Siu & Ho, 2020).

Finally, we examined the effects of phonological awareness and RAN on writing (Zoccolotti et al., 2005) and investigated possible differences between monolingual and bilingual children (Hsu et al., 2016).

Predictions

We expected to find positive effects of metaphonological skills on decoding accuracy (Ball & Blachmann, 1988; Bradley, 1988; Frith, 1985; Hulme & Snowling, 1988; Lundeberg et al., 1988) and of RAN speed on decoding speed (Wolf & Obregon, 1992; Ziegler et al., 2010), as well as relations of RAN with the early indicators of phonological awareness, and, possibly, with grapheme recognition (Decker et al., 2013).

4.2. Method 2

4.2.1. Participants, materials, and procedures

Participants (see 3.2.1), materials (see 3.2.2), and procedures (see 3.2.3) are the same as in Study 1 and are summarized in Table 1.

Detailed information about the materials is in sections 2.2 – 2.14.

4.2.2. Analyses

We conducted a series of correlation analyses between each literacy skill and measures of language, metaphonology, RAN, and short-term memory analyzed longitudinally in Study 1.

We performed Generalized (Linear) Mixed Models (GMM) for cross-sectional measures and Generalized (Linear) Mixed Models for Repeated Measures (GMMRM) for longitudinal measures. Literacy measures were assumed as dependent variables, correlating preliteracy skills as fixed effects, and bilingualism as a random effect. Different test sessions (KG-1; KG-2; 1-1; 1-2; and 2-1) were entered as a repeated measure, as in Study 1. Satterthwaite approximation was used.

4.3. Results 2

The analyses of relations between children's pre-literacy skills at the end of kindergarten (KG-2) and early literacy skills in the middle of the 1st grade (1-1) gave the following outcomes (for complete statistics, see Table B.1 in Appendix B).

Sublexical decoding speed correlates

- strongly ($\rho > 0.50$) with syllable blending; initial syllable recognition; visual discrimination of capital letters;
- moderately ($0.30 < \rho < 0.50$) with digit span backward; syntactic comprehension; syllable segmentation; recognition of word and non-word minimal pairs; of rhymes; graphic reproduction of one semicircle; and speed in the RAN of pictures;
- weakly ($\rho < 0.30$) with word span and non-word repetition.

Sublexical decoding errors correlate

- strongly with syllable blending;
- moderately with digit span backward; non-word repetition; syllable segmentation; recognition of word and non-word minimal pairs, rhymes, and initial syllable; speed in the RAN of pictures;
- weakly with word span and errors in the RAN of pictures.

Word writing under dictation correlates

- strongly with digit span backward; syntactic comprehension; syntactic production of DOcl tested with the most accessible test version including only items with gender and number matching, whereas the intermediate test version including gender mismatch correlates moderately only with grapheme errors but not with word and syllable errors. There are strong correlations also with visual discrimination of capital letters (only with word errors and, moderately, syllable errors, whereas there is no relation with grapheme errors), and speed in the RAN of pictures;
- moderately with non-word repetition (only with word and syllable errors, whereas the relation with grapheme errors is weaker); receptive vocabulary; recognition of non-word minimal pairs, rhymes (only with word errors, whereas the relations with syllable and grapheme errors are weaker), initial syllable, and errors in the RAN of pictures;
- weakly with syllable blending (only word and syllable errors) and recognition of word minimal pairs;

In the middle of the 1st grade (Table B.2), sublexical decoding speed correlates

- strongly with initial and final syllable deletion; phonemic verbal fluency; and speed in the RAN of colors, pictures, and digits;
- moderately with productive vocabulary; phoneme blending; rhyme recognition; and phoneme segmentation; reproduction of two and three semicircles;
- weakly with errors in the RAN of digits.

Sublexical decoding errors correlate

- strongly with phoneme blending; initial and final syllable deletion; phoneme segmentation; speed in the RAN of digits;
- moderately with digit span backward; non-word repetition; receptive and productive vocabulary; recognition of rhymes; phonemic verbal fluency; graphic reproduction of three

semicircles; speed in the RAN of colors and pictures; and errors in the RAN of pictures and digits;

- weakly with errors in the RAN of colors.

Word writing under dictation correlates

- strongly with productive vocabulary; phoneme blending, initial syllable deletion, and phoneme segmentation (word and syllable errors, whereas the relation with grapheme errors is more moderate); speed in the RAN of colors, pictures, and digits;
- moderately with digit span backward (only word and syllable errors); non-word repetition; receptive vocabulary; final syllable deletion; rhyme recognition; phonemic verbal fluency; and errors in the RAN of pictures and digits (only word and syllable errors).

As for reading skills in the middle of the 2nd grade (Table B.3),

text decoding speed correlates

- strongly with spoonerism and errors in the RAN of colors;
- moderately with word span; receptive and productive vocabulary; syntactic comprehension; phoneme blending; phonemic verbal fluency; speed in the RAN of colors and pictures; and errors in the RAN of pictures.

Text decoding errors correlate

- strongly with word span;
- moderately with phonemic verbal fluency, spoonerism, and errors in the RAN of colors.

Word decoding speed correlates

- strongly with spoonerism and errors in the RAN of colors;
- moderately with word span, phoneme blending; rhyme recognition; phonemic verbal fluency; and speed in the RAN of colors and pictures.

Word decoding errors correlate

- strongly with phonemic verbal fluency and spoonerism;

- moderately with word span; syntactic comprehension; speed in the RAN of colors; and errors in the RAN of pictures.

Stress errors in word decoding correlate

- strongly with word span and syntactic comprehension;
- moderately with receptive vocabulary; phonemic verbal fluency; and errors in the RAN of pictures.

Non-word decoding speed correlates

- strongly with phonemic verbal fluency and spoonerism;
- moderately with speed in the RAN of colors, pictures, and digits.

Non-word decoding errors correlate

- strongly with spoonerism;
- moderately with word span, phoneme blending, and phonemic verbal fluency.

Reading comprehension total scores correlate

- very strongly ($\rho > 0.70$) with receptive vocabulary and syntactic comprehension, and strongly with word span, productive vocabulary, and errors in the RAN of pictures;
- moderately with rhyme recognition, phonemic verbal fluency, and speed in the RAN of colors, pictures, and digits.

Reading comprehension central scores correlate

- very strongly with syntactic comprehension; strongly with receptive and productive vocabulary, word span, phonemic verbal fluency, and errors in the RAN of pictures;
- moderately with spoonerism, speed in the RAN of colors and pictures, and errors in the RAN of colors.

Reading comprehension specific scores correlate

- strongly with receptive vocabulary and syntactic comprehension;
- moderately with word span; productive vocabulary; rhyme recognition; speed in the RAN of colors, pictures, and digits; and errors in the RAN of pictures.

As for writing skills in the middle of the 2nd grade (Table B.4),

word dictation errors correlate

- strongly with phoneme blending, phonemic verbal fluency, and spoonerism;
- moderately with word span, syntactic comprehension, and rhyme recognition;

Non-word dictation errors correlate

- moderately with phoneme blending, rhyme recognition, and spoonerism.

Text dictation total errors correlate

- strongly with phoneme blending;
- moderately with syntactic comprehension; phoneme segmentation; final syllable deletion; spoonerism; speed and errors in the RAN of colors.

Text dictation phonological errors correlate

- strongly with syntactic comprehension; phoneme blending and segmentation; final syllable deletion; and speed in the RAN of colors;
- moderately with receptive and productive vocabulary; rhyme recognition; spoonerism; speed in the RAN of pictures; and errors in the RAN of colors.

Text dictation non-phonological errors correlate

- strongly with phoneme blending;
- moderately with digit span backward; phoneme blending and segmentation; spoonerism; and errors in the RAN of colors.

Stress/geminate errors in text dictation correlate moderately with phoneme blending.

Omissions in text dictation correlate moderately with receptive vocabulary and syntactic comprehension;

Minor errors in text dictation correlate

- moderately with receptive vocabulary; syntactic comprehension; and errors in the RAN of pictures.

Inappropriate words in text dictation correlate

- strongly with receptive vocabulary;
- moderately with syntactic comprehension.

Finally, moderate relations emerged in the last year of kindergarten between speed in the RAN of pictures and all global metaphonological skills ($0.337 < \rho_s < 0.477$; $0.000 < p < 0.005$) except syllable segmentation, and moderate to strong relations between visual discrimination of letters ($-0.457 < \rho_s < -0.531$; $0.004 < p < 0.014$) or graphic reproduction of single semicircles ($-0.429 < \rho_s < -0.492$; $0.001 < p < 0.023$) and all global metaphonological skills except syllable blending and word minimal pairs.

These relations are longitudinally present in all administrations until the 2nd grade between all the domains of RAN speed/errors and analytical metaphonological skills.

4.3.1. Effects of preliteracy at the end of kindergarten on literacy in the middle of the 1st grade

A series of Generalized Linear Mixed Models (GLMM) were analyzed with sublexical decoding (speed and errors) and word-dictation error scores (word errors, syllable errors, and grapheme errors) obtained in the middle of the 1st grade as dependent variables; short-term memory skills, RAN of pictures, and visual scores obtained at the end of KG and correlating with the dependent variable⁴¹ as fixed effects; and bilingualism as a random effect.

The models correctly predict sublexical decoding speed ($F(10, 53) = 7.276$, $p = 0.000$), sublexical decoding errors ($F(7, 60) = 6.593$, $p = 0.000$), and word-dictation word errors (F

⁴¹ See 4.3 and Tables B.1 – B.4.

(10, 51) = 2.527, $p = 0.015$), syllable errors ($F(9, 52) = 3.128$, $p = 0.004$), and grapheme errors ($F(8, 53) = 3.796$, $p = 0.001$).

Syllable segmentation and recognition of word minimal pairs show significant positive effects on sublexical decoding speed, whereas syllable blending shows a significant negative effect on sublexical decoding errors. Digit span backward has a significant negative effect on word dictation word errors, syllable errors, and grapheme errors, as shown in Figure 21.

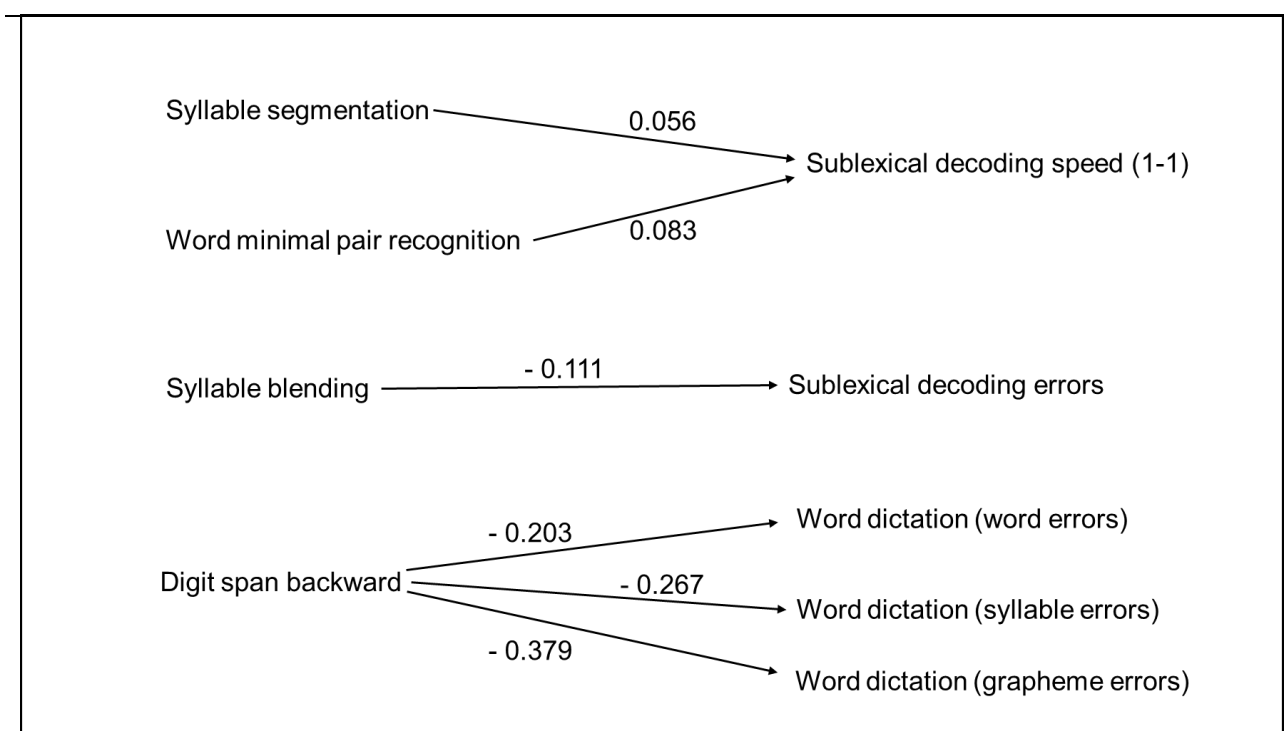


Figure 21. Significant effects of short-term memory, metaphonology, and RAN skills at the end of kindergarten on sublexical decoding and word dictation in the middle of the 1st grade of primary school. Coefficients are indicated on the respective arrows. Complete statistics of effects are in Appendix B (Table B.5).

No effects of bilingualism emerged on both sublexical decoding speed and word dictation ($p > 0.05$).

4.3.2. Longitudinal effects of preliteracy skills on word writing between the middle of the 1st grade and the middle of the 2nd grade

To explore the longitudinal effect of short-term memory, analytical metaphonological skills, and RAN skills on word dictation in the three test sessions in the middle (1-1) and end (1-2) of the 1st grade and in the middle of the 2nd grade (2-1), a series of GMMRM were analyzed with short-term memory measures (digit and word span, non-word repetition, graphic reproduction of one and two semicircles, visual recognition of capital letters), analytical metaphonological skills, and RAN skills (colors, pictures, and digits) correlating with the dependent variable as three distinct blocks of fixed effects, and bilingualism as a random effect.

Word dictation errors are correctly predicted by phonological/visual memory ($F(7, 66) = 5.571, p = 0.000$), analytical metaphonological skills ($F(5, 115) = 24.826, p = 0.000$), and RAN ($F(6, 111) = 10.622, p = 0.000$). Among short-term memory measures, we found main effects of digit span backward and phonological short-term memory (non-word repetition); among metaphonological skills, a main effect of phoneme blending, and among RAN measures, a main effect of digit speed. Digit span backward, non-word repetition, phoneme blending, and digit RAN speed have a significant negative effect on word writing errors, as shown in Figure 22.

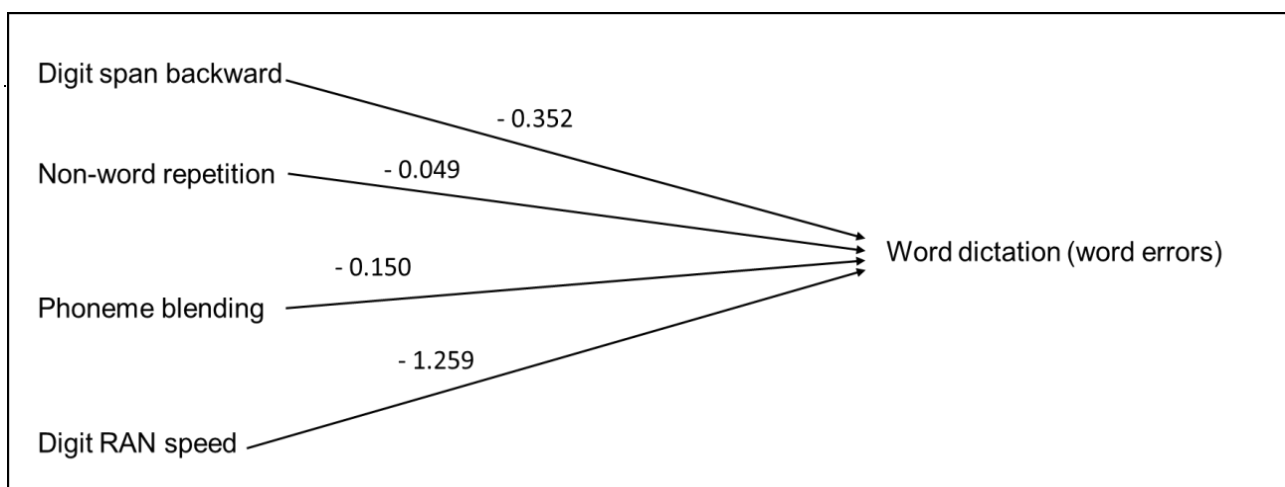


Figure 22. Significant longitudinal effects of short-term memory measures, analytical metaphonological skills, and RAN on word dictation between the middle of the 1st grade and the middle of the 2nd grade of primary school. Coefficients are indicated on the respective arrows. Complete effect statistics are in Appendix B (Table B.6)

No effects of bilingualism on word dictation emerged ($p > 0.05$).

4.3.3. Longitudinal effects of preliteracy skills on text reading and word writing between the end of the 1st grade and the middle of the 2nd grade

To explore the longitudinal effects of the same skills on text decoding, text comprehension, and word writing in the two administrations between the end of the 1st grade (1-2) and the middle of the 2nd grade (2-1), separate GMMRM were analyzed with short-term memory (digit span forward, digit span backward, and word span), language (receptive vocabulary, syntactic comprehension, and syntactic production of DOcls), metaphonological, and RAN scores correlating with the dependent variable as four distinct blocks of fixed effects, and bilingualism as a random effect.

The global GMMRM correctly predict text decoding speed ($F(9, 42) = 9.459, p = 0.000$); text decoding errors ($F(8, 42) = 5.002, p = 0.000$); reading comprehension total scores ($F(9, 42) = 4.975, p = 0.000$), specific scores ($F(7, 44) = 5.719, p = 0.000$), and central scores ($F(7, 44) = 5.459, p = 0.000$); and word dictation errors ($F(7, 42) = 4.872, p = 0.000$).

Text decoding speed is predicted by phoneme blending, speed in the RAN of digits, and final-syllable deletion, whereas text decoding errors only by final-syllable deletion.

Reading comprehension total and central scores are predicted by syntactic comprehension, whereas comprehension specific scores by syntactic comprehension, deletion of the initial syllable, and speed in the RAN of digits, as shown in Figure 23.

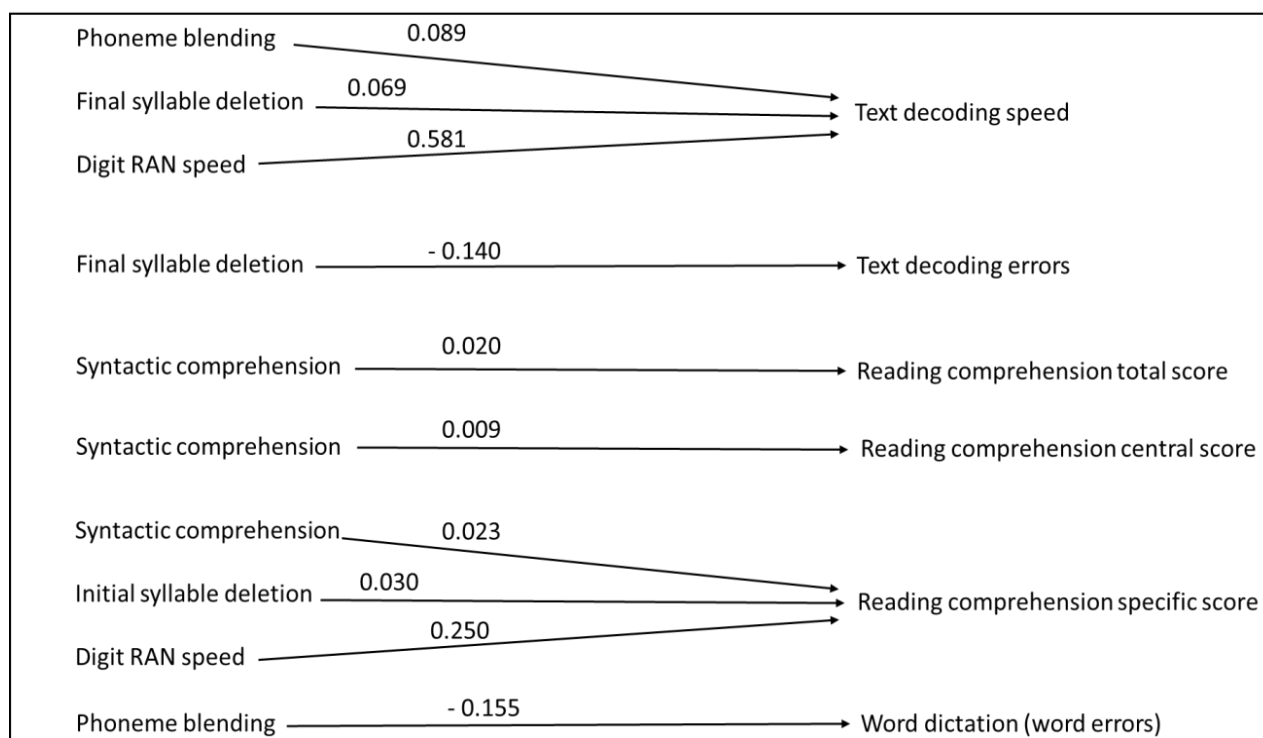


Figure 23. Significant longitudinal effects of memory measures, analytical metalinguistic skills, and RAN on text decoding, text comprehension, and word dictation between the end of the 1st grade and the middle of the 2nd grade of primary school. Coefficients are indicated on the respective arrows. Complete effect statistics are in Appendix B, Table B.7

No effects of bilingualism on either reading or word dictation emerged ($p > 0.05$).

4.3.4. Longitudinal effects of 1st and 2nd-grade preliteracy skills on literacy in the middle of the 2nd grade

A series of GMMRM were analyzed with scores obtained in literacy tests in the middle of the 2nd grade as dependent variables; short-term memory, language, metaphonology, and RAN longitudinal scores obtained in the three administrations of 1st and 2nd grade and correlating with the dependent variable as fixed effects; and bilingualism as a random effect. The global models correctly predict text decoding speed ($F(13, 60) = 3.300, p = 0.001$); text decoding errors ($F(5, 75) = 5.031, p = 0.000$); word decoding speed ($F(11, 62) = 5.222, p$

= 0.000); errors in decoding generic word lists ($F(4, 76) = 4.552, p = 0.002$); stress errors in word decoding ($F(10, 63) = 2.087, p = 0.039$); speed ($F(4, 75) = 8.223, p = 0.000$) and errors ($F(6, 62) = 4.212, p = 0.001$) in decoding non-words; word decoding attentional errors ($F(2, 63) = 18.676, p = 0.000$); reading comprehension total score ($F(12, 61) = 4.774, p = 0.000$), central score ($F(14, 60) = 2.705, p = 0.004$), and specific score ($F(9, 64) = 4.053, p = 0.000$); non-word dictation ($F(2, 81) = 5.331, p = 0.007$); text dictation total errors ($F(13, 61) = 3.320, p = 0.001$), phonological errors ($F(6, 65) = 9.014, p = 0.000$), non-phonological errors ($F(2, 72) = 10.578, p = 0.000$), stress/geminate errors ($F(8, 52) = 2.093, p = 0.053$), omissions ($F(5, 32) = 3.369, p = 0.015$), and minor errors ($F(8, 52) = 4.187, p = 0.001$);.

Figure 24 shows significant predictive effects.

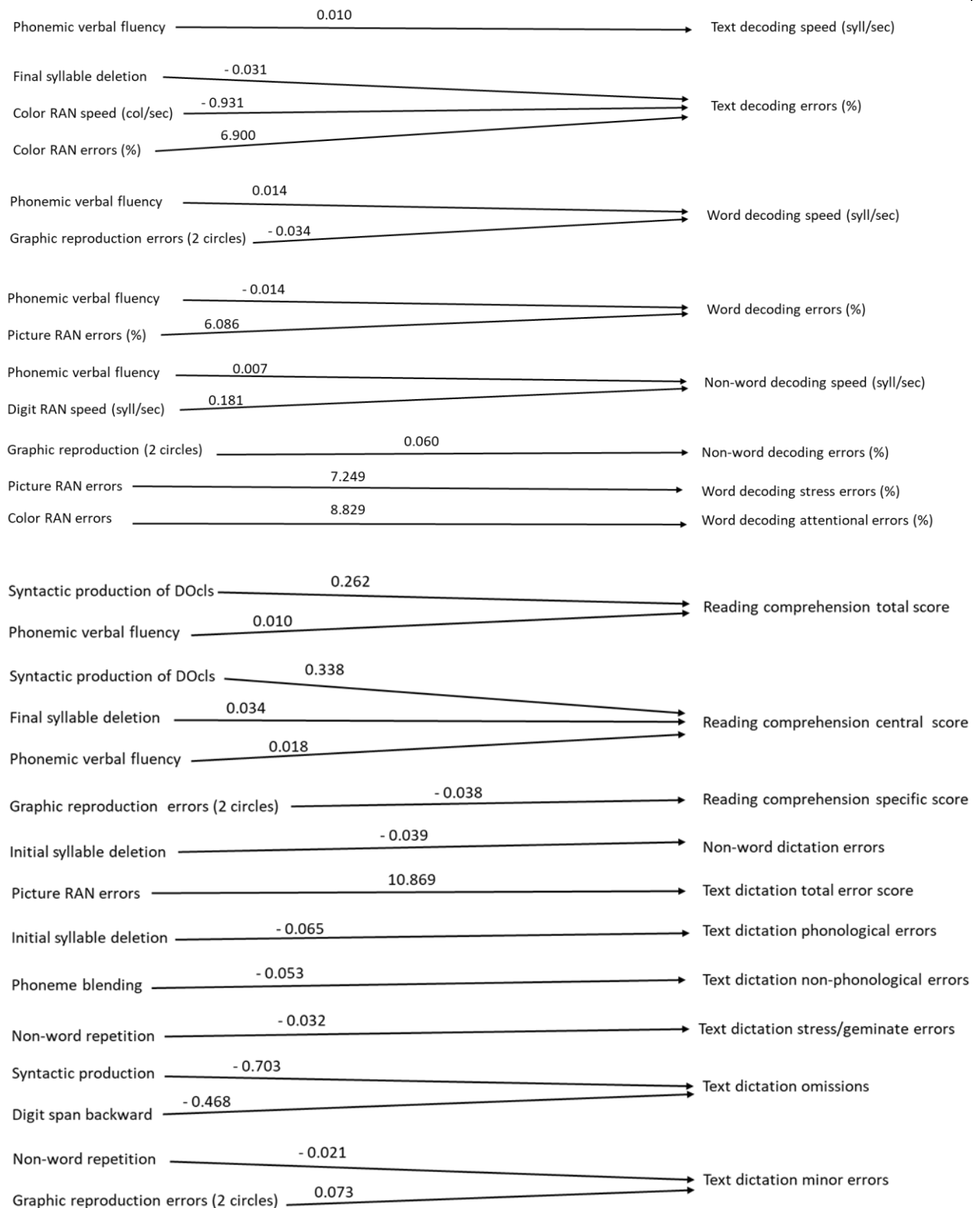


Figure 24. Significant effects of longitudinal scores of short-term memory, metaphonology, and RAN obtained in the three administrations of 1st and 2nd grade, on literacy scores obtained in the middle of the 2nd grade. Coefficients are indicated on the respective arrows. Complete effect statistics are in Appendix B, Table B.8.

No effects of bilingualism emerged on any of the dependent variables ($p > 0.05$).

4.3.5. Effects of language skills and spoonerism on literacy in the middle of the 2nd grade

Since productive vocabulary was tested once per grade and spoonerism was tested only once in the middle of the 2nd grade, a series of Generalized Linear Mixed Models (GLMM) were analyzed to explore the effects of language skills (receptive and productive vocabulary, syntactic comprehension, and syntactic production of DOcls) and spoonerism on different measures of literacy. The models correctly predict text decoding errors ($F(1, 24) = 4.492, p = 0.045$), word decoding speed ($F(3, 22) = 3.627, p = 0.029$), word decoding stress errors ($F(2, 23) = 3.702, p = 0.040$), non-word decoding speed ($F(3, 22) = 3.839, p = 0.024$), non-word decoding errors ($F(3, 22) = 3.392, p = 0.036$), word dictation errors ($F(3, 21) = 3.403, p = 0.037$), text dictation phonological errors ($F(2, 22) = 8.086, p = 0.002$), reading comprehension total score ($F(4, 21) = 10.344, p = 0.000$), central score ($F(5, 20) = 9.304, p = 0.000$), and specific score ($F(4, 21) = 4.321, p = 0.010$). Spoonerism predicts word decoding speed; non-word decoding speed and errors; and errors in word-dictation and text-dictation. Syntactic comprehension predicts word decoding stress errors; text decoding errors; and reading comprehension total, central, and specific scores. Productive vocabulary predicts phonological errors in text dictation, as shown in Figure 25.

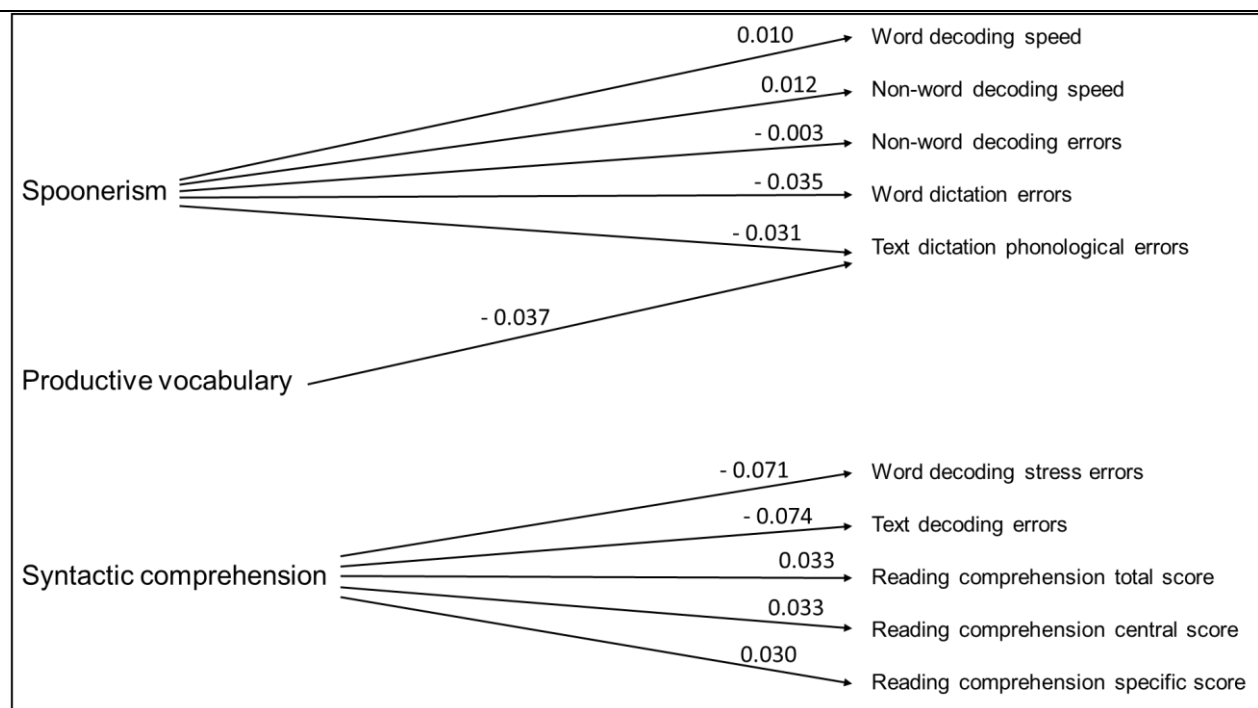


Figure 25. Significant effects of language skills on literacy skills in the middle of the 2nd grade. Coefficients are indicated on the respective arrows. Complete effect statistics are in Appendix B, Table B.9.

No effects of bilingualism on any literacy skills emerged ($p > 0.05$)

4.4. Discussion 2

4.4.1. Effects of preliteracy skills at the end of kindergarten on literacy in the middle of the 1st grade

In preschool-aged children, working memory (initially backward digit span and then word span), metaphonological skills and RAN, and syntactic abilities correlate to different extents with both reading and writing (see 4.2). We explored whether some of these skills have a predictive effect on literacy.

Firstly, we asked whether preliteracy skills at the end of kindergarten would predict reading (sublexical decoding) and writing (word dictation) in the middle of the 1st grade. We found

(see 4.3.1) that children's syllable segmentation abilities and their ability to recognize minimal pairs at the end of kindergarten predicted sublexical decoding speed in the middle of the 1st grade, whereas the ability to blend syllables into words predicted sublexical decoding accuracy. The segmentation of written words into syllables constitutes an intermediate step between phonological recoding and orthographic representations of word forms in the mental lexicon for efficient word recognition (Ehri, 2001, 2005). Heikkilä et al. (2013) analyzed 265 poor readers in 2nd and 3rd grade and found that repeated reading of infrequent syllables increased decoding speed at the word level in a transparent orthography like Finnish. Müller et al. (2017) administered a 16-session intervention aiming at improving the mental representation of word syllabic structure to 75 German poor readers in 4th grade and found a positive effect on fluency in word decoding but not in text decoding. Bernard et al. (2014) found that segmentation of words into syllables through the use of different colors could be useful for young adult readers (age 22 - 26) with a central vision deficit, who are usually poor readers, to reduce the word recognition time in French. The magnitude of the effect was small (10.38% for low-frequency words and smaller print size words) but reliable. This suggests possible patterns for intervention, which might be extended to different functional syllabic units, such as morphemes. In our sample, we found that syllable segmentation is already acquired in the middle of the last year of kindergarten, with no distinction between monolinguals and bilinguals (see 3.3.2). So, the present outcomes suggest that syllable segmentation is a possible predictive tool for decoding speed in both monolingual and bilingual children. It also suggests that syllable segmentation could serve as the target for training aimed at increasing fluency in children who have slow decoding and/or poor metaphonological skills.

Recognition of minimal pairs also predicted decoding speed. These consist of word pairs with an equal number of segments, one phonemic contrast in a reciprocal position, and a single distinctive feature contrast at the phonetic level. For these properties, they were used

by Blache et al. (1981) on children with multiple articulation errors for a clinical intervention aiming at teaching the linguistic significance of distinctive feature properties. The authors suggest that minimal pairs can be manipulated to teach or analyze phonemic as well as morphemic competence. Cilibrasi et al. (2014) found that young adult English monolingual speakers were faster in discriminating words generating lexical minimal pairs (e.g., *back* / *badge*) than morphosyntactic minimal pairs (e.g., *asks* / *asked*), thus arguing that these two categories must be stored differently in memory, namely lexical minimal pairs are stored as units, whereas morphosyntactic minimal pairs are not. To our knowledge, there is no study showing the direct effect of minimal pairs on reading. Roberts & Walter (2016) worked on Kabiye, a language of Togo with a toneless standard orthography, which presents many tonal minimal pairs, namely heterophonic homographs, as long as many homophonic homographs within conjugations. They proposed a morphographic spelling where elided root-final labial consonants were superscripted as silent letters to facilitate the lexeme identification and compared the reading of this spelling with a diacritic tonographic alternative. They found that participants who learned the morphographic spelling read more accurately than those who learned the tonographic spelling.

In our sample, word minimal pair recognition showed to predict sublexical decoding speed. Non-word minimal pairs have the same surface characteristics as word minimal pairs but did not show any effect on reading. Hence, we guess that the effect found has to do with lexical/semantic properties of word minimal pairs. Interestingly, word minimal pair recognition correlates with receptive lexicon ($\rho_s = 0.460$; $p = 0.014$) and with syntactic comprehension ($\rho_s = 0.510$, $p = 0.006$), which, in turn, correlates with sublexical decoding speed ($\rho_s = 0.409$, $p = 0.409$) and word writing errors ($\rho_s = -0.533$, $p = 0.003$).

As for decoding accuracy, we found it to be predicted by syllable blending. Müller et al. (2020), after a 24-session syllable-based intervention on 549 German 2nd graders, found positive effects on phonological recoding, direct word recognition, and text-based reading

comprehension. Poor readers of the experimental group showed greater improvement in word recognition than children with good word-recognition skills. The results of the mentioned studies, albeit conducted on older samples than ours, provide evidence that a syllable-based reading intervention can increase decoding skills, even in struggling readers, and improve their reading comprehension.

Our analyses confirmed a longitudinal effect of syllable segmentation and blending in the last year of kindergarten on reading in the middle of the 1st grade and reported an effect of syllable segmentation on decoding speed, and of syllable blending on decoding accuracy. The most solid kindergarten predictor of word writing accuracy in 1st-grade was digit span backward. The effect was greater on grapheme errors, intermediate on syllable errors, and smaller on word errors, thus showing that working memory is particularly engaged in the phoneme-grapheme conversion in the first steps of literacy acquisition when children begin to use the sublexical route of writing. The highly demanding nature of word writing for working memory was shown by Colombo et al. (2009), who found that an articulatory suppression task affects the selection of the serial grapheme order in word writing even in skilled writers⁴².

Our outcome can be explained through a *capacity* approach to working memory (for the *capacity theory*, see Just & Carpenter, 1992; for the *capacity theory* applied to writing, see McCutchen, 1996). In this view, a central source of working memory is shared between the storage and processing involved in writing, which absorbs it completely, thus impeding the interaction with other processes. This might be the reason why no other ability appeared to predict word writing at this stage. It is plausible that, as long as writing processes are automatized, other abilities can interact with working memory in predicting writing. With this regard, McCutchen (1988) considers skilled writing to be achieved as all processes involved

⁴² They tested Italian University students.

(namely text planning, translation, and revision) are interactively coordinated in working memory.

4.4.2. Longitudinal effects of preliteracy skills on word writing between the middle of the 1st grade and the middle of the 2nd grade

A longitudinal analysis of the predictors of word writing in 1st and 2nd grade (see 4.3.2) showed, as hypothesized in 4.4.1, that word writing errors are predicted by digit span backward in interaction with other abilities, namely non-word repetition, phoneme blending, and speed in the RAN of digits. We found evidence of an effect of RAN on writing accuracy. RAN has been mainly studied in relation to reading (Wolf & Obregon, 1992) rather than writing (Albuquerque, 2017), and to speed rather than accuracy (Ziegler et al., 2010). This might constitute a bias in the evaluation of its role (see Martin et al., 2010; Patel et al., 2004; Talli et al., 2015; Vaessen et al., 2009). Here, we found a robust effect of digit RAN on word writing accuracy, followed by digit span backward, phoneme blending, and non-word repetition. Accuracy in word writing under dictation, then, turns to be the result of executive, working memory, metaphonological, and phonological skills. In the very early steps of literacy acquisition (administration 1-1), working memory span had a prevalent role. As long as children's abilities develop at the analytical level, executive and (meta)phonological skills acquire more importance (see 4.4.1). The prominent role of working memory span and RAN (which, in the case of digits, in particular, mirrors the ability to rapidly retrieve the phonological equivalents of graphemes), might be related to the kind of task used. Dictation, especially under the condition of collective administration, without repetition and with standard times, requires, beyond executive skills, the fast elaboration of phonological strings, fast retrieval of phoneme correspondences, and grapheme serial recoding, so that the role of memory span, phonological memory, fast retrieval, and phoneme blending are

crucial. In this view, it is not surprising that RAN speed shows an effect on accuracy. The large size of this effect reveals that digit RAN has a greater effect than expected on accuracy⁴³, but this might be related more to writing processes⁴⁴ than to writing accuracy per se. It seems more plausible that interacting phonological and phoneme-blending abilities could be responsible for accuracy. It is worth noting that RAN has been deemed as a measure of automaticity, namely of the extent of speed and autonomy in process execution⁴⁵, which allows wasting less cognitive resources (Norton & Wolf, 2012). Furthermore, among all RAN measures, digit RAN has been considered the most automatic and the least demanding in terms of language and cognitive processes involved (Närhi et al., 2005; Wiig et al., 2000). In this vein, the high predictive power of digit RAN on word writing might indicate that children have acquired a good rate of automaticity in writing. The other abilities that interact with working memory span and RAN to the achievement of the word writing task are non-word repetition and phoneme blending. The former, as a measure of phonological working memory and phonological ability, is responsible for the acoustic-phonological elaboration of the phonemic string, the temporary allocation to the phonological buffer, and the phonological – orthographic conversion, whereas the latter contributes to the operations of acoustic – phonological conversion and sequential elaboration of the phoneme-grapheme conversion. The entire route process is regulated by working memory (digit span backward) and automaticity processes (RAN). The different coefficient values (Figure 22) give an idea of the contribution extent of each ability to the achievement of the writing model (Figure 26).

⁴³ Modest relations of RAN with spelling accuracy are reported by Babayiğit & Stainthorp, 2011; Ding et al., 2010; Georgiou et al., 2011; Justi & Roazzi, 2012; Moll et al., 2009; Savage et al., 2005; 2008; Stainthorp, Powell, & Stuart, 2013.

⁴⁴ Albuquerque (2017) found moderate relations of digit RAN with handwriting fluency in 3rd and 4th graders.

⁴⁵ Kail et al. (1999) consider RAN to be a measure of processing speed

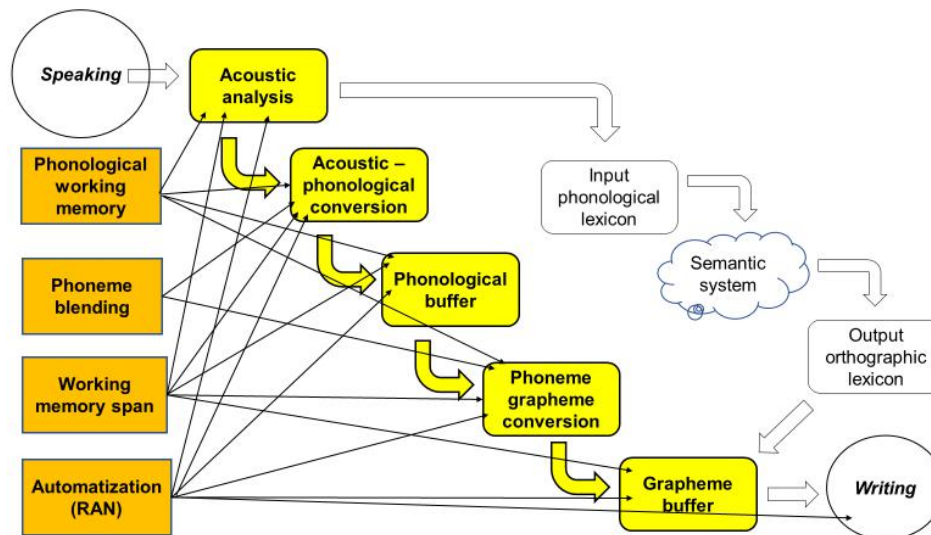


Figure 26. Hypothesis of the contribution of phonological working memory, phoneme blending, working memory span, and automaticity processes to the sublexical route of writing (yellow cases) in the implementation of the word dictation task in the middle of the 1st grade of primary school. White cases represent the lexical route of writing, which is not still used in this stage.

4.4.3. Longitudinal effects of preliteracy skills on text reading and word writing between the end of the 1st grade and the middle of the 2nd grade

Between the end of the 1st grade and the middle of the 2nd grade (see 4.3.3), text decoding speed is predicted by phoneme blending, final-syllable deletion, and speed in the RAN of digits, whereas text decoding accuracy only by final-syllable deletion.

In the very early steps of reading acquisition, decoding speed was predicted by syllable segmentation and a global discrimination ability like word minimal pair recognition (see 4.3.1 and 4.4.1). As long as literacy and metaphonology are refined, phoneme blending and a manipulation ability like final-syllable deletion get responsible for decoding speed. RAN speed, which is known by the literature as a typical predictor of decoding speed (for a review, see Norton & Wolf, 2012), reveals to take this role in particular with digits and to contribute also for a portion of writing accuracy (see 4.3.2 and 4.4.2). Final-syllable deletion, beyond

contributing to text decoding speed, is the main predictor of text decoding accuracy, thus showing that sound-manipulation abilities have a key role in decoding accuracy and might underlie the strong relationship existing between decoding speed and accuracy ($\rho_s = -0.725$, $p = 0.000$).

Reading comprehension central and total (central + specific) scores are predicted by syntactic comprehension, whereas comprehension specific scores, beyond syntactic comprehension, are predicted by deletion of the initial syllable and speed in the RAN of digits. Specific comprehension, having a literal relationship with the text, shows to be directly affected by the abilities that predict decoding. This is in line with the Simple View of Reading (Gough & Tunmer, 1986), according to which reading is the product of written decoding and language comprehension. The Simple View of Reading could then explain the predictive role of RAN in reading comprehension, which might be a secondary effect due to the primary effect on decoding. This seems to be confirmed from the effect size of digit RAN speed, which is more than twice as big on text decoding ($Coeff = 0.581$) as on reading comprehension ($Coeff = 0.250$). Furthermore, the effect is present only on specific comprehension, thus demonstrating that reading comprehension predicted by RAN is text-related.

Interestingly, while decoding accuracy and speed are affected by final syllable deletion, reading comprehension is affected by initial syllable deletion, thus showing partial independence between decoding accuracy and reading comprehension.

Oral syntactic comprehension is known to be a predictor of reading comprehension in both first and second languages, even in non-alphabetic languages (see, for instance, Chik, 2012; Morvay, 2012; Simpson et al., 2020). In a cross-sectional study on a sample of primary-school children with developmental dyslexia, Casani (2020c) found that syntactic comprehension correlated with reading comprehension only in the younger group (age 6;7 – 9;3), whereas this correlation was absent among elder children (9;5 – 11;3), probably

because older children develop different strategies for comprehension, e.g. contextual strategies. Interestingly, among typically developing controls, the correlation between syntactic comprehension and reading comprehension was absent even in the younger group, where reading comprehension correlated with decoding, thus showing that children adopt different comprehension strategies fitting their abilities. In this regard, Hulme & Snowling (2009: 127) note that children with dyslexia use some particular strategies to compensate for their decoding difficulties, such as “getting the gist from a passage”, and adult subjects with developmental dyslexia rely on context to decode words (Bruck, 1990)⁴⁶. Word dictation errors, in the last two administrations (the former (1-2) collective and the latter (2-1) individual), are predicted only by phoneme blending. This might confirm that the effects of digit span backward, non-word repetition, and RAN, which emerged from the analyses of both collective administrations (1-1 and 1-2), might be related to the particular administration conditions of the collective test, namely standard times and no repetition. The next analysis, which assumes as an objective only the individual administration (2-1), will help disentangle this aspect.

4.4.4. Longitudinal effects of 1st and 2nd-grade preliteracy skills on literacy in the middle of the 2nd grade

We analyzed the abilities that, during the course of 1st and 2nd grade, predict literacy achievements in 2nd grade.

The outcomes (Figure 24) show that phonemic verbal fluency is a key ability for reading, as it predicts text decoding speed, word decoding speed and errors, non-word decoding speed,

⁴⁶ These findings are related to connectionist models of reading development (Plaut et al., 1996), according to which phonological and semantic processes interact.

reading comprehension total and central score. Verbal fluency has been considered a cognitive function rather than a metalinguistic ability, which facilitates information retrieval from memory. This requires executive control over cognitive processes, such as selective attention, selective inhibition, mental set shifting, internal response generation, and self-monitoring (Patterson, 2011).

Another ability that affects many measures of reading is RAN. In particular, speed and errors in the RAN of colors predict text decoding errors, and color errors predict attentional errors in word decoding; errors in the RAN of pictures predict word decoding speed, and stress errors in word decoding; speed in the RAN of digits predicts non-word decoding speed. Except for digit speed predicting non-word decoding speed, RAN accuracy more than RAN speed turns to be a predictor of decoding, and more for accuracy than for speed. There are some differences in predictive semantic categories of RAN. Color errors predict text decoding errors and word decoding attentional errors. A fine-grained analysis of text decoding errors might be useful to explore whether color errors in RAN are predictive of attentional errors also for text decoding, thus suggesting a particular relation of color RAN with the orthographic-visual analysis stage of the dual-route model of reading. This might have interesting implications for the early detection of reading attentional deficits in both monolinguals and bilinguals, as colors are the most accessible condition to test RAN in bilingual children, who showed a similar performance as monolinguals only in this RAN category (see Study 1).

Errors in the RAN of pictures predict errors in decoding generic word lists and stress errors in particular. Even in this case, a fine-grained qualitative analysis of errors made in generic word lists and text might help determine whether picture errors in RAN specifically predict lexical errors in reading, another aspect that may open interesting research questions in the field of the early detection and treatment of surface reading deficits.

Digit RAN speed predicts non-word decoding speed. Digits are considered the most automatic and least demanding category to be named, in terms of language and cognitive processes involved, compared to colors and pictures (Närhi et al., 2005; Wiig et al., 2000), provided that children have learned to visually distinguish digits. Furthermore, RAN with alphanumeric stimuli was reported to be associated with more reading accuracy (Manis et al., 2000, Savage et al., 2008), and with more reading accuracy and speed (Cardoso-Martins & Pennington, 2004; Georgiou, Parrila, & Liao, 2008; van den Bos et al., 2002).

Our data tell us that RAN is a good predictor of decoding speed in Italian, which has a shallow orthography, only for non-words and with RAN digit measures. Other RAN measures (colors and pictures) are better predictors of accuracy (in particular, picture RAN predicts lexical errors in word decoding; and color RAN predicts (generic) errors in text decoding and attentional errors in word decoding).

Picture RAN errors revealed to be also good predictors of writing errors (for some effects of RAN on writing fluency and spelling in European Portuguese, see Albuquerque, 2012).

Different predictive patterns emerged for different writing error types. Phonological errors in writing are predicted by a phonological manipulation ability like deletion of the initial syllable. This is valid for both phonological errors in text dictation and errors in non-word dictation. Non-phonological errors in text dictation are predicted by phoneme blending. This is plausible, as these errors include illegal fusion and segmentation (see 2.14). A more fine-grained error coding distinguishing these errors from non-homograph homophones could shed more light on this issue.

Stress/geminate writing errors are predicted by non-word repetition. This might indicate that these errors are phonological in nature but with a different status compared to sound change and migrations, which are predicted by syllable manipulation (deletion of the initial syllable), instead. Further qualitative analyses might explore whether metaphonological manipulation skills predict particularly sound migration errors. In this case, a more fine-grained error

coding distinguishing sound migration and addition from sound change within the phonological-error category might be proposed.

Digit span backward and syntactic production of direct object clitics predict omissions in text writing. Word omissions during dictation might be due to a range of reasons that have to do with working memory, (morpho)syntax, and phonology. A working memory overload due to the complexity of the task and to pressing conditions associated with the collective dictation, which is implemented in standard times and without repetition, might be a cause of omission errors. The predictive role of syntactic production should be searched in the particular nature of the structure investigated, namely direct object clitic pronouns. As shown in 3.4.1, the use of direct object clitic pronouns requires a range of phonological, morphosyntactic, syntactic, and pragmatic abilities. In particular, clitics are small and phonologically weak words, so their production might be overlooked by children, due to scarce phonological salience (Leonard, 1998) or prosodic reasons (Bortolini & Leonard, 1996). At the same time, they convey a heavy load of morphosyntactic information, as they are marked for person, gender, number, and case, they are subject to syntactic movement and can assume marked positions within the clause. Furthermore, they are mandatory in some contexts, either forbidden or optional in others. These characteristics make them particularly sensitive to reveal language difficulties (Casani 2020a; Zachou et al., 2013). In this case, omissions in text dictation might be due to a failure in phonological skills that lead children to overlook portions of the dictated string; or to immature morphosyntactic competence that could induce children to overlook functional strings (Casani, in preparation); or to incomplete pragmatic competence that may provoke the omission of mandatory elements. A qualitative analysis of the omitted text portions might help explore the reasons for the omissions more deeply, thus allowing the planning of ad hoc training tapping children's writing skills.

These data, besides highlighting the complexity of a composite task like text writing under dictation, confirm that morphosyntactic competence has a significant role in literacy, and

reveal that this role extends to writing since the first steps of literacy acquisition. Further research is needed to investigate which aspects of morphosyntactic competence act on reading and writing, how this happens, and if training morphosyntax might indirectly affect decoding and surface aspects of writing.

Minor errors (punctuation, and capital letter errors) in text writing are predicted by the graphic reproduction of two semicircles, and, curiously, by non-word repetition.

In a componential approach to working memory (see Baddeley, 1986; 2000), the phonological loop and the visual sketchpad interact to accomplish the writing task. One possibility is that minor errors might be due to a non-optimal performance of the visual component. The issue, however, might be more complex and could be interpreted in light of integration between a componential and a capacity (Just & Carpenter, 1992; McCutchen, 1996) model of working memory. In this vein, the negative effect of non-word repetition associated with the positive effect of graphic reproduction on minor errors might be due to the condition of cognitive overload determined by a complex activity like dictation. Under this condition, the central executive (Baddeley, 1996; 2007), which governs the two slave systems, might not be able to fulfill the entire load of functions related to the visual component, thus determining the graphic errors. In such a model, the phonological system would assume a competing role compared to the visual system. The existence of competing components of working memory responsible for these outcomes implies the presence of specialized memory registers devoted to managing the visual and phonological information separately, which might indicate a maturation of children's skills and strategies (see Kellogg, 1996).

4.4.5. Effects of language skills and spoonerism on literacy in the middle of the 2nd grade

In the middle of the 2nd grade, spoonerism predicts word and non-word decoding speed, non-word decoding errors, word dictation errors, and text dictation phonological errors.

Syntactic comprehension predicts word decoding stress errors, text decoding errors, reading comprehension total score, central score, and specific score.

Productive vocabulary predicts text dictation phonological errors.

Mastering complex sound manipulation activities like spoonerisms turns out to have similar effects on speed in decoding words and non-words. The weaker effect on decoding accuracy of non-words (but no effect on words) suggests that high-level sound-manipulation abilities might be taken as a measure of processing efficiency in phonological activation, which allows for faster decoding. Perry et al. (2019) and Ziegler et al. (2020), in simulating individual learning trajectories and intervention outcomes based on the five main component skills of the self-teaching-hypothesis model (letters, phonemes, decoding network, phonological lexicon, and orthographic lexicon), took phoneme-deletion abilities as a measure of correct activation of phonemes (see 1.4). In this vein, spoonerisms can be deemed as a further analytical development of sound deletion skills, and these, in turn, as a development of syllable deletion skills, which we have seen to predict different measures of reading and writing between the end of the 1st grade and the middle of the 2nd grade.

Spoonerisms also predict word and text writing, and the strong effects demonstrate that the ability to manipulate phonemes and word chunks is crucial for writing in shallow orthographies. This ability also allows to decode pseudowords (which are inexistent words constituted by existing chunks) more accurately (and fastly), as shown by the effect of spoonerism on non-word decoding accuracy (and speed).

Both word decoding and text decoding are predicted by syntactic comprehension, thus confirming the role of (morpho)syntactic competence on written decoding (Casani, in preparation) beyond written comprehension. For words, the effect is specific to stress errors, thus showing the interrelation between syntactic and lexical competence, which is confirmed

by the strong relations of syntactic comprehension with both receptive ($\rho = 0.540, p = 0.000$) and productive ($\rho = 0.554, p = 0.000$) vocabulary. The correct decoding of Italian non-paroxytone words requires children to resort to their lexical skills. If children have never heard the word, they will probably read it as a paroxytone.

Syntactic comprehension is also the main predictor of reading comprehension, at both the literal and inferential levels.

Both word dictation and text dictation are predicted by spoonerism, which is particularly revealing of phonological errors. Phonological competence in writing is then increased by high-level sound manipulation skills, thus showing an interactive relation between written language and metalanguage. Phonological errors in text writing are also predicted by productive vocabulary. Productive vocabulary seems to have a longitudinal relationship with literacy. Lee (2011) found that, after controlling for gender, birth order, ethnicity, and socioeconomic status, vocabulary size at age 2 can significantly predict subsequent language and literacy achievement up to fifth grade. Duff et al. (2015a) showed that a latent factor of infants' receptive and expressive vocabulary was a significant predictor of later vocabulary, phonological awareness, reading accuracy, and reading comprehension. Other studies concentrated on the longitudinal relation between infant vocabulary and literacy achievement (e.g., Duff et al., 2015b).

Our results show that productive vocabulary in 2nd grade predicts, in particular, phonological errors in text dictation, thus highlighting the bridge existing between children's lexical and phonological competence. In particular, a more accessible writing task, such as word writing, is in 2nd grade predicted by phoneme blending, whereas the more demanding text writing task is predicted by more specific abilities, such as refined sound manipulation (spoonerism) and productive vocabulary. The effect of vocabulary is not only due to the possibility that a wider word repertoire can easier activate children's writing attempts (under dictation) by reference to word meaning and pronunciation (for the same effect on decoding, see Dyson

et al., 2017) thus making writing more fluent, but seems to have a deeper impact on phonological coding. According to the self-teaching hypothesis of reading (see 1.4), phonology works as a “built-in teacher”, as it activates pre-existing words in children’s (phonological) lexicon, thus creating the orthographic entries that increase and strengthen the network of letter-sound connections (Ziegler et al., 2020). At the same time, learners’ vocabulary has a relevant role in the choice of the right candidate, by providing a *set of variability* that allows the reader for the correct approximation (Tunmer & Chapman, 2012). The predictive effect of productive vocabulary on phonological errors in text writing might be explained through the self-teaching hypothesis, which seems to be valid also in a shallow orthography like the Italian.

These data generally show that different aspects of literacy have different predictors, which develop interactively, thus confirming the multicomponent nature of reading and writing.

4.5. Interim summary 2

To address our research questions, we found positive effects of some particular metaphonological skills on decoding accuracy and a sort of preliteracy developmental pattern (from recognition to manipulation). These abilities interact to different extents with working memory and other abilities. Syllable blending in kindergarten predicts sublexical decoding accuracy in 1st grade. Between the end of the 1st grade and the 2nd grade, final syllable deletion longitudinally predicts text decoding accuracy, in association with color RAN speed and accuracy. Phonemic verbal fluency predicts word decoding accuracy and, to a lower extent, speed.

RAN speed of digits confirms to be a good predictor of text decoding speed and non-word decoding speed between the end of the 1st grade and the middle of the 2nd grade, whereas RAN speed of pictures and colors does not predict speed.

Between kindergarten and 1st grade, RAN correlates moderately with global and analytical metaphonological skills, whereas in the 2nd grade, its relations with analytical metaphonological skills completely disappear. No relations were found with the visual recognition of capital letters, whereas very moderate relations were found between KG and 1st grade between picture RAN speed and sublexical decoding.

Between the end of the 1st grade and the middle of the 2nd grade, RAN speed and accuracy of colors predict text decoding accuracy, and RAN accuracy of pictures predicts word decoding accuracy, in particular word stress errors. RAN speed of colors predicts word decoding attentional errors.

An effect of RAN on reading comprehension was found, as RAN speed of digits, as well as initial syllable deletion and syntactic comprehension, predicts reading specific comprehension. This has been explained as a secondary effect through the Simple View of Reading (Gough & Tunmer, 1986),

For what concerns reading, in sum, we confirm different contributions from different layers of metaphonological competence to decoding accuracy and speed, and from different RAN semantic categories to both decoding speed and accuracy, and, in one case, reading specific comprehension.

Similar but more dynamic patterns emerged for writing. In kindergarten, the only predictor of 1st-grade word writing is working memory span. Between the 1st and 2nd grade, working memory span becomes part of an interactive model, where phonological skills, phoneme blending, and digit RAN speed contribute to different extents in the implementation of the phonological route of writing. In this framework, RAN is interpreted as a measure of automaticity (Norton & Wolf, 2012). As long as children's abilities increase, the contribution of working memory is not prominent anymore in word dictation, due to the automatization of processes, and phoneme blending remains the main predictor of word writing. Text writing general score in 2nd grade is predicted only by picture RAN errors, whereas different grades

of contribution emerged thanks to the error analysis. In particular, phonological and non-phonological errors are mainly regulated by metaphonological manipulation skills, whereas phonological working memory directly intervenes only on stress/geminate errors and minor errors. The effect of phonological skills (non-word repetition) on stress/geminate errors encourages us to interpret those as a type of phonological errors. The effect of non-word repetition on minor errors, in turn, is interpreted in the framework of interaction between capacity and componential models of short-term memory, in which those errors are generated by an overload of the central executive. This encourages us to consider the non-word repetition test as a phonological test involving a prominent role of phonological short-term memory (see 2.1.3).

Bilingualism does not determine any significant difference in the patterns described, thus revealing no differences in literacy acquisition by monolingual and monoliterate bilingual children in a shallow orthography like Italian.

5. Study 3

Effects of metaphonological, RAN, and graphic interventions on the development of preliteracy and literacy in monolingual and mono-literate bilingual children in an Italian school setting

5.1. Aims 3

This study investigates the effects of a metaphonological (henceforth META), a RAN, and a graphic (henceforth GRAPH) intervention on the development of preliteracy and early literacy skills in a sample of monolingual and bilingual children with typical development in an Italian school.

We expect the META intervention to improve decoding accuracy (Bernhardt & Major, 2004; Gillon, 2019; Hulme et al., 2012) and the RAN intervention to increase decoding speed (Pecini et al., 2016; Wolff, 2014).

We wonder whether the relation of RAN with phonological awareness and grapheme recognition (see section 4.3 as well as Decker et al., 2013) might determine any overlap, namely an improvement of decoding speed in the META group, and of decoding/writing accuracy in the RAN group.

We also intend to explore if reading comprehension is directly or indirectly affected by different types of intervention (Wolff, 2014).

Furthermore, we investigate the possible effects of interventions on writing skills, which have received less attention from the literature so far (Zoccolotti et al., 2005).

The GRAPH intervention was conceived as a placebo, but we do not exclude it might have some impact on the executive aspects of writing.

Possible differences in the intervention effect on monolingual and bilingual children are eventually explored (Kuo et al., 2016).

A further methodological question concerns the use of Generalized Linear Mixed Models to longitudinally explore the intervention effects. With this aim, we wonder whether their use with small samples gives us converging results.

5.2. Method 3

5.2.1. Participants

Participants were the same as in Studies 1 and 2. A core group of 34 children (18 females + 16 males; 19 monolinguals + 15 bilinguals) in the last year of kindergarten was randomly divided into three samples to attend three different training programs, namely a META training (N = 12), a RAN training (N = 11), and a placebo GRAPH training (N = 11). After the first screening in the middle of the last year of KG (henceforth KG-1), parametric/non-parametric Anovas showed no significant difference between the 3 samples in any ability ($p > 0.05$).

At the end of the last year of kindergarten (KG-2) and in the middle of the 1st grade (1-1) a wider control group (henceforth, CG) that did not attend any training was included (N = 48, 24 monolinguals + 24 bilinguals).

An additional group was included in the CG for the word-dictation test in administration 1-1 (N = 46 children (26 MONO + 20 BI)), the word-dictation and reading comprehension tests

in administration 1-2 (N = 34 (17 MONO + 17 BI)), and the text dictation and reading comprehension tests in administration 2-1 (N = 37 (18 MONO + 19 BI)).

During the course of the project, one bilingual child of the RAN group was diagnosed with specific Language Impairment (henceforth, LI⁴⁷) and two monolingual children of the GRAPH group with LI in association with mild intellectual disability. Another case of specific LI and two cases of LI with mild intellectual disability in the CG were also excluded from the global analyses. These children's performances are analyzed in Study 4.

Two monolingual children (1 META + 1 RAN), who transferred schools in the 1st grade, one bilingual META child who transferred schools in the middle of the 1st grade, and 2 bilingual children (1 META + 1 GRAPH), who were absent for more than one month in the 1st grade, were also excluded from global analyses.

Table 21 displays the exact number of monolingual and bilingual participants for each group in each testing session.

5.2.2. Materials

Testing materials (see 3.2.2) were the same as in Study 1 and Study 2. Detailed information about testing materials and scoring is in sections 2.1.2 – 2.1.14.

Adapted printed materials, digital (Brignola & Perrotta, 2010; Brignola et al., 2012; 2013; Dutto, 2014; Mei et al., 2013; Perrotta & Brignola, 2000, 2010 ; Stella, 2013; Zoccali & Mauro, 2006), and ad hoc materials were employed for training.

⁴⁷ For the use of the acronym LI (Language Impairment) instead of LD (Language Disorder), see footnote 22 (section 3.2.1).

5.2.3. Procedures

Testing procedures were the same as in Study 1 and Study 2. Children were tested twice per year, in the middle (KG-1) and at the end (KG-2) of kindergarten, in the middle (1-1) and at the end (1-2) of the 1st grade, and in the middle of the 2nd grade (2-1).

The abilities tested in each test session as well as the exact number of monolingual and bilingual participants for each group in each testing session are in Table 21.

	KG-1				KG-2				1-1				1-2				2-1			
	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG
forward/backward digit span forward (Bisiacchi et al., 2005)					12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)					9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
word span (Spinnler & Tognoni, 1987)					12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)					9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
non-word repetition (Cornoldi et al., 2019)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)								
metaphonology (syllable awareness) (Marotta et al., 2008)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)												
metaphonology (phonemic awareness) (Marotta et al., 2008)									12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
RAN of pictures (Terreni et al., 2002)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)												
RAN of colors, pictures, and digits (De Luca et al., 2005)									12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
graphic reproduction of series of 1 and 2 semicircles oriented in different fashions (Cornoldi et al., 2019)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	12 (7+5)	9 (5+4)	7 (4+3)									
graphic reproduction of series of 3 semicircles oriented in different fashions (Cornoldi et al., 2019)									12 (7+5)	9 (5+4)	7 (4+3)									
visuo-spatial recognition of capital letters (Cornoldi et al., 2019)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	12 (7+5)	9 (5+4)	7 (4+3)									
syntactic comprehension (Bishop, 2009)	12 (7+5)	9 (5+4)	7 (4+3)		12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)									9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
syntactic production of direct object clitic pronouns (Arosio et al., 2014)	12 (7+5)	9 (5+4)	7 (4+3)																	
syntactic production of direct object clitic pronouns (Casani & Cardinaletti, 2019)					12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)												
syntactic production of direct object clitic pronouns (Casani et al., in preparation)												48 (24+24)					9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
receptive vocabulary (Dunn et al., 2000)					12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)				48 (24+24)					9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
productive vocabulary (Kaplan et al., 1983)												48 (24+24)					9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
sublexical decoding (Calgaro et al., 2018);									12 (7+5)	9 (5+4)	7 (4+3)	48 (24+24)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)

	KG-1				KG-2				1-1				1-2				2-1			
	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG	META	RAN	GRAPH	CG
sublexical (non-word) decoding (Traficante et al., 2017);																	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
word decoding (Traficante et al., 2017)																	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
text decoding (Cornoldi & Carretti, 2017)													9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
reading comprehension (Cornoldi & Carretti, 2016)													9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
reading comprehension (Cornoldi & Colpo, 2011)													9 (6+3)	8 (4+4)	7 (4+3)	34 (17+17)	9 (6+3)	8 (4+4)	7 (4+3)	37 (18+19)
word dictation (Stella & Apolito, 2004)									12 (7+5)	9 (5+4)	7 (4+3)	94 (50+44)	9 (6+3)	8 (4+4)	7 (4+3)	34 (17+17)				
word/non-word dictation (Sartori et al. 2017)																	9 (6+3)	8 (4+4)	7 (4+3)	6 (3+3)
text dictation (Tressoldi et al., 2013)																	9 (6+3)	8 (4+4)	7 (4+3)	37 (18+19)

Table 21. Number of monolingual and bilingual subjects (N = MONO + BI) in each group for each testing session and each ability tested.

Two intervention programs were performed, the former in the second half of the last year of kindergarten, between test sessions KG-1 and KG-2 (40 dates, 3 times a week = 20 hours), and the latter in the second half of the 1st grade of primary school, between test sessions 1-1 and 1-2 (14 dates, every day = 11 hours). Both training programs took place in small blended groups with 3 / 4 participants.

In KG, the META group was mainly trained in syllable awareness through activities of increasing difficulty (syllable blending and segmentation; initial and final syllable recognition; rhyme recognition; word and non-word minimal pair recognition; syllable puzzles), whereas in the 1st grade, the training focused on phonemic awareness and manipulation tasks (rhyme recognition and production; phonemic synthesis and segmentation; phonemic verbal fluency; initial, central, and final syllable deletion; spoonerisms).

The RAN group was trained in the rapid naming of pictures and colors in KG, and pictures, colors, and digits in the 1st grade of primary school through printed materials and a digital tachistoscope.

The GRAPH group was mainly engaged in drawing different frame sequences and in some graphic videogames.

Activities involving the use of digital materials were performed on a 16-inch laptop screen.

Training sessions were held by the author and some collaborators⁴⁸ who received ad hoc regular training.

Figure 27 summarizes the study design.

⁴⁸ Collaborators were young fellows graduated in psychology who were running an internship or a collaboration program in the hosting Institute. They were selected and trained by the author.

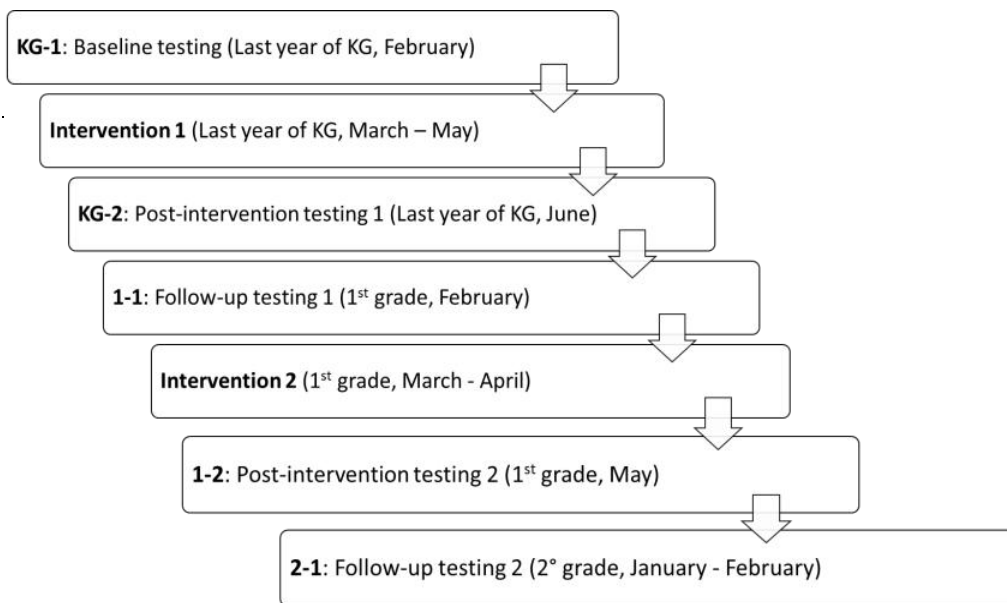


Figure 27. Study design.

5.2.4. Analyses

Given the small sample that received the intervention, the different number of participants in each testing session, the differences of distribution and variance of scores, we opted for the use of Generalized (Linear) Mixed Models (for Repeated Measures) (GMM(RM)), which allowed us to control the different variables simultaneously. Generalized estimating equations are reported as reliable in longitudinal experimental designs involving small samples (for a clinical paper supporting their use in a study with 12 participants, see Muth et al., 2016). Nevertheless, considering the very small size of subgroups (see Table 21), we also calculated the Standardized Mean Differences (SMDs) between and within groups by using Hedge's g with small-sample correction for both cross-sectional and longitudinal data. Intercrossed analyses were then performed, even considering the significance of non-parametric tests (Kruskal-Wallis; Wilcoxon).

The results of the alternative analyses are in Appendix C.

5.3. Results 3

In this section, we explore the effects of different training on children's preliteracy and literacy skills.

5.3.1. Effects of the training on pre-literacy skills

After the first screening (KG-1), the analysis of a series of Generalized (Linear) Mixed Models (GMM) with test scores as dependent variables, the 3 levels of training group (META; RAN, and GRAPH) as fixed effects, and bilingualism as a random effect did not find any effect of group ($0.109 < p < 0.986$) and bilingualism ($0.361 < p < 1.000$), showing no differences among the abilities of the three groups.

Global metaphonological skills

At the end of the last year of kindergarten, global metaphonological skills, namely syllable awareness, were retested after the first training cycle. A series of GMMRM were analyzed, with metaphonological scores as dependent variables, the combination between group level (META, RAN, GRAPH, CG) and administration time (KG-1, KG-2) as fixed effects, and bilingualism as a random effect. Outlier cases were pairwise excluded.

The models significantly predict syllable blending ($F(6, 48) = 2.359, p = 0.045$), recognition of word minimal pairs ($F(6, 45) = 2.559, p = 0.032$), recognition of rhymes ($F(6, 51) = 2.347, p = 0.045$), and recognition of the initial syllable ($F(4, 77) = 3.058, p = 0.022$), with a main combined effect of group and administration time for the first three skills, and separate effects of group ($F(3, 83) = 3.062, p = 0.033$) and administration time ($F(1, 61) = 6.155, p = 0.016$) for the recognition of the initial syllable. Figure 28 shows the estimated means of the groups for each administration.

No effect on syllable segmentation and recognition of non-word minimal pairs emerged ($p > 0.05$).

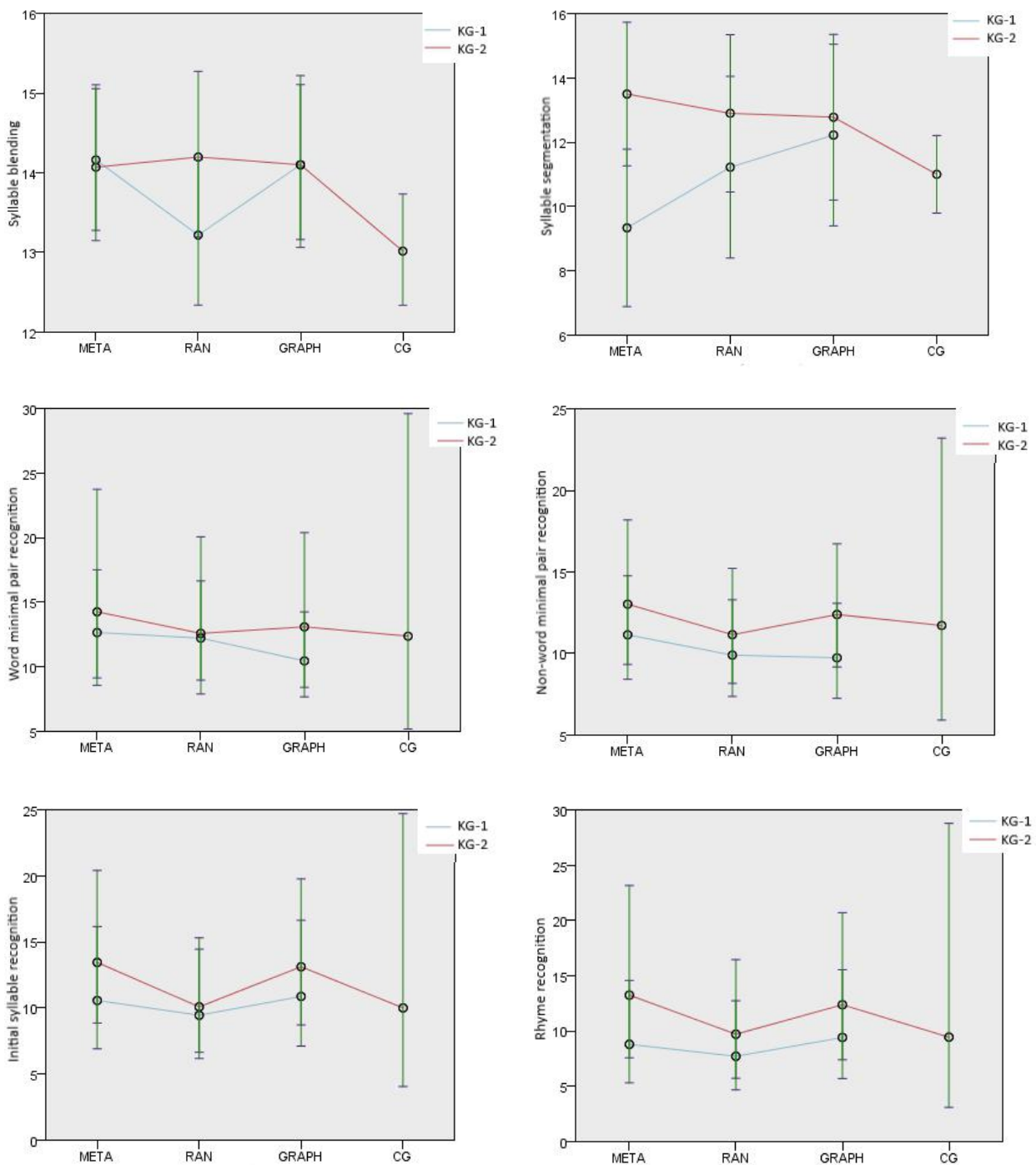


Figure 28. Estimated means of each group in the two administrations (colored lines) of global metaphonological tests in kindergarten.

Significant positive effects of the three trained groups emerged:

- on syllable blending, in both administrations by the META group, in the last administration by the RAN group, and in the first administration by the GRAPH group;

- on word minimal pair recognition, in the last administration by the META group;
- on the recognition of rhymes, in the last administration by the META group;
- on the recognition of the initial syllable, where a positive effect of the META group and a negative effect of the first administration were found.

Significant effects are in Table 22;

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Syllable blending	META * KG-1	0.084	0.033	2.557	0.014	2.512	2.620
	META*KG-2	0.078	0.036	2.142	0.036	0.005	0.150
	RAN*KG-2	0.087	0.039	2.226	0.030	0.009	0.165
	GRAPH*KG-1	0.080	0.035	2.254	0.029	0.008	0.152
Word minimal pair recognition	META*KG-2	0.142	0.047	3.029	0.003	0.049	0.236
Rhyme recognition	META*KG-2	0.339	0.124	2.744	0.008	0.093	0.586
Initial syllable recognition	META	0.208	0.083	2.517	0.014	0.044	0.373
	KG-1	- 0.205	0.083	- 2.481	0.016	- 0.371	- 0.040

*Satterthwaite approximation

Table 22. Significant predictive effects of group and administration time on global metaphonological skills in kindergarten.

The post hoc analysis showed no difference between groups in KG-1, whereas in KG-2 the META ($M = 14.069$, $SE = 0.464$, $CI [13.148, 15.058]$) and the RAN ($M = 14.197$, $SE = 0.509$, $CI [13.197, 15.272]$) group showed significantly better outcomes than the CG ($M = 13.016$, $SE = 0.256$, $CI [12.335, 13.733]$) in syllable blending. Moreover, the META group showed significantly better scores than the CG in the recognition of word minimal pairs ($Est = 1.887$, $SE = 0.667$, $t(65) = 2.830$, $p = 0.006$, $CI [0.555, 3.219]$), of the initial syllable ($Est = 2.229$, $SE = 0.912$, $t(88) = 2.445$, $p = 0.016$, $CI [0.418, 4.041]$), and of rhymes ($Est = 3.809$, $SE = 1.615$, $t(50) = 2.359$, $p = 0.022$, $CI [0.566, 7.053]$).

No significant effect on syllable segmentation ($p = 0.159$) and recognition of non-word minimal pairs ($p = 0.138$), as well as no effect of bilingualism ($0.514 < p < 1.000$) emerged.

To summarize the results of global metaphonological skills, before training, the 3 groups did not differ significantly from each other in any ability, even though a combined positive effect of group and first administration revealed a certain starting advantage by the META and the GRAPH group in syllable blending. After training, the META and the RAN group improved significantly in syllable blending, whereas the GRAPH group did not. The META group only improved significantly in the recognition of word minimal pairs and rhymes. After training, the META and RAN group had significantly higher scores than the CG in syllable blending, whereas the META group only was significantly more proficient than the CG in the recognition of word minimal pairs, initial syllables, and rhymes.

Analytical metaphonological skills

Analytical metaphonological skills were tested three times in the middle (1-1) and at the end (1-2) of the first grade, and in the middle of the 2nd grade (2-1) of primary school. The rhyme-recognition test was administered five times (in the middle and at the end of KG, in the middle and at the end of the 1st grade, and in the middle of the 2nd grade).

A series of GMMRM were analyzed, with analytical metaphonological scores as dependent variables, the combination of group (META, RAN, GRAPH, CG) and administration time as fixed effects, and bilingualism as a random effect. The models correctly predict phoneme blending ($F(11, 37) = 8.616, p = 0.000$), phoneme segmentation ($F(11, 104) = 5.545, p = 0.000$), deletion of the initial syllable ($F(11, 41) = 10.815, p = 0.000$), deletion of the final syllable ($F(11, 41) = 7.340, p = 0.000$), phonemic verbal fluency ($F(5, 71) = 4.932, p = 0.001$), and recognition of rhymes ($F(18, 42) = 4.937, p = 0.000$).

Figure 29 shows the estimated means of the groups in each administration.

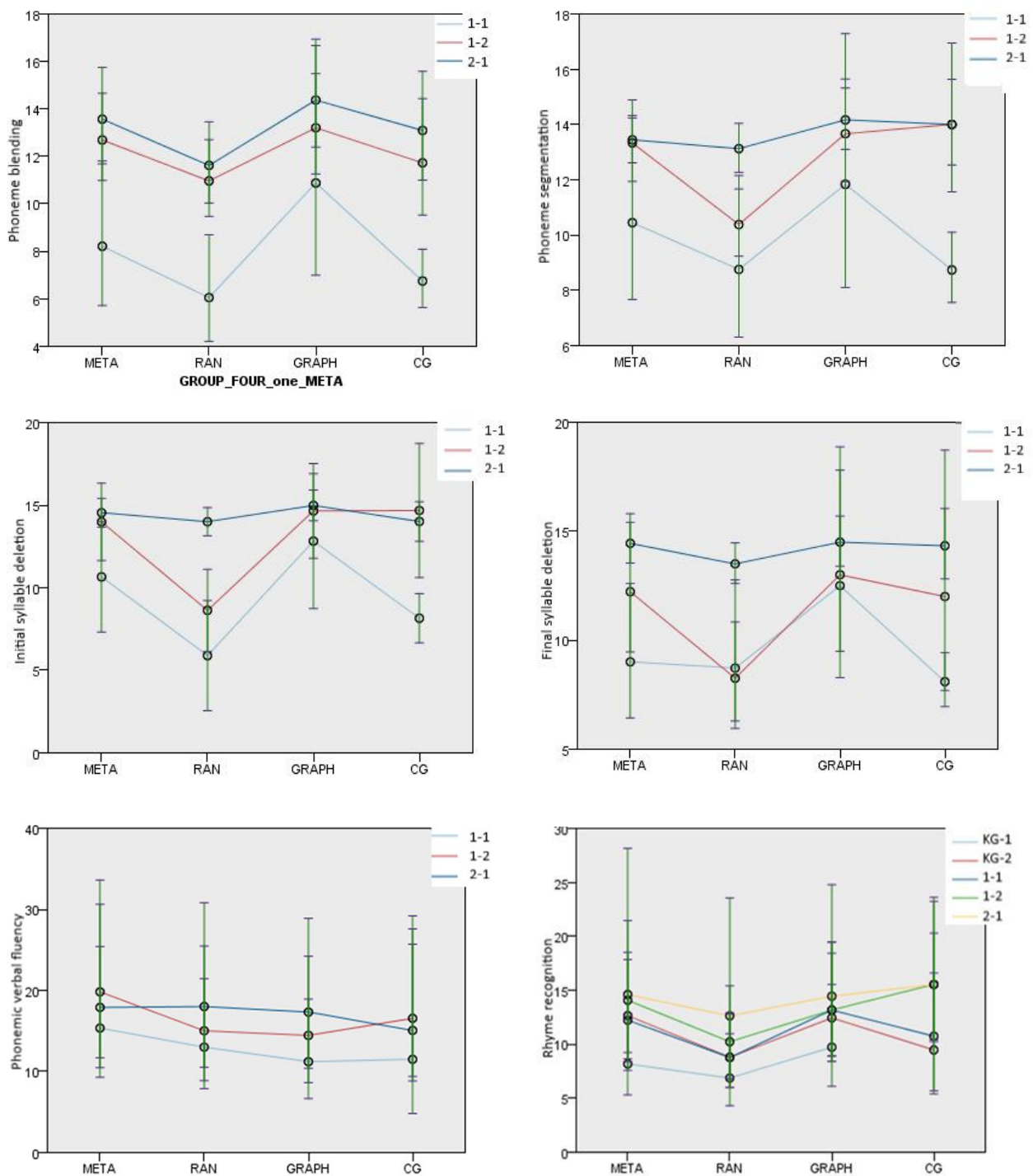


Figure 29. Estimated means of each group (x-axis) in the three administrations (colored lines) of analytical metaphonology tests (first 5 cases) and on the five administrations of the rhyme-recognition test (last case).

The following significant effects emerged:

- a negative effect on phoneme blending in administration 1-1 by the META group, the RAN group, and the CG;

- a negative effect on phoneme segmentation and deletion of the initial syllable in administration 1-1 and 1-2 by the RAN group, and in administration 1-1 by the CG;
- a negative effect on the deletion of the final syllable in administration 1-1 by the META group and the CG, and in administration 1-1 and 1-2 by the RAN group;

Phonemic verbal fluency was predicted only by administration time ($F(5, 71) = 4.932, p = 0.001$). A negative effect of administration 1-1 was found ($F(2, 57) = 4.576, p = 0.014$).

Recognition of rhymes was tested five times (in the middle and at the end of KG, in the middle and at the end of the 1st grade, and in the middle of the 2nd grade). Significant negative effects emerged in KG-1 by the META group; in KG-1, KG-2, 1-1, and 1-2 by the RAN group; in KG-2 by the GRAPH group; and in KG-2 and 1-1 by the CG. Significant predictive effects are in Table 23.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Phoneme blending	META * 1-1	- 0.467	0.194	- 2.405	0.019	- 0.854	- 0.080
	RAN * 1-1	- 0.773	0.194	- 3.984	0.000	- 1.159	- 0.387
	CG * 1-1	- 0.663	0.111	- 5.961	0.000	- 0.885	- 0.441
Phoneme segmentation	RAN * 1-1	- 0.470	0.175	- 2.688	0.008	- 0.817	- 0.123
	RAN * 1-2	- 0.300	0.081	- 3.690	0.000	- 0.461	- 0.139
	CG * 1-1	- 0.472	0.092	- 5.132	0.000	- 0.655	0.221
Initial syllable deletion	RAN * 1-1	- 8.127	1.760	- 4.617	0.000	- 11.632	- 4.623
	RAN * 1-2	- 5.837	1.320	- 4.081	0.000	- 8.080	- 2.964
	CG * 1-1	- 5.862	0.924	- 6.342	0.000	- 7.700	- 4.023
Final syllable deletion	META * 1-1	- 0.465	0.177	- 2.631	0.010	- 0.818	- 0.113
	RAN * 1-1	- 0.498	0.198	- 2.508	0.014	- 0.893	- 0.102
	RAN * 1-2	- 0.552	0.142	- 3.886	0.001	- 0.843	- 0.262
	CG * 1-1	- 0.572	0.093	- 6.119	0.000	- 0.758	- 0.386
Phonemic verbal fluency	1-1	- 0.304	0.104	- 2.920	0.005	- 0.511	- 0.096
Rhyme recognition	META * KG-1	- 0.643	0.195	- 3.297	0.002	- 1.041	- 0.245
	RAN * KG-1	- 0.817	0.216	- 3.790	0.001	- 1.259	- 0.376
	RAN * KG-2	- 0.570	0.165	- 3.645	0.001	- 0.897	- 0.243
	RAN * 1-1	- 0.573	0.155	- 3.700	0.000	- 0.881	- 0.265
	RAN * 1-2	- 0.419	0.127	- 3.289	0.002	- 0.676	- 0.162
	GRAPH * KG-1	- 0.469	0.216	- 2.175	0.038	- 0.911	- 0.028
	CG * KG-2	- 0.496	0.108	- 4.601	0.000	- 0.714	- 0.278
	CG * 1-1	- 0.369	0.104	- 3.548	0.001	- 0.581	- 0.158

*Satterthwaite approximation

Table 23. Significant longitudinal effects of group and administration time on the 3 administrations of analytical-metaphonology tests, and on the 5 administrations of the rhyme-recognition test.

To summarize the results of analytical-metaphonological tests, the CG improved along with the META and the RAN group in phoneme blending after the 1st grade, revealing no effect of interventions.

As for phoneme segmentation and deletion of the initial syllable, the RAN group appeared particularly disadvantaged, as a significant improvement was evident only in 2nd grade. It is not likely that the better result of the META group was due to the metaphonological intervention, as the GRAPH group had similar outcomes.

As for deletion of the final syllable, the improvement of the CG along with the META group after post-intervention 2, and the difficulties shown by the RAN group revealed that neither the metaphonological nor the RAN intervention had significant effects on final syllable deletion.

As for phonemic verbal fluency, the general negative effect of administration 1-1 revealed an initial difficulty by every group, showing no differences due to the training.

As for rhyme recognition, the RAN group showed an improvement only in the 2nd grade, whereas the META and the GRAPH group showed a significant improvement in KG-2 and the CG in administration 1-2. Hence, the RAN training had no effect. The META group improved significantly in Post-Intervention 1, as long as the GRAPH group, showing no particular effect of the metaphonological training.

RAN

RAN of pictures was tested twice in kindergarten, twice in 1st grade, and once in 2nd grade; RAN of colors and numbers was tested twice in the 1st grade and once in the 2nd grade.

A series of GMMRMs were analyzed, with RAN scores as the dependent variable, groups and administration times as fixed effects, and bilingualism as a random effect. The models correctly predict RAN speed ($F(18, 44) = 16.260, p = 0.000$) and errors ($F(18, 39) = 4.014, p = 0.000$) for pictures, and RAN speed for digits ($F(11, 39) = 7.377, p = 0.000$). Figure 30 shows the estimated means of the groups in each administration.

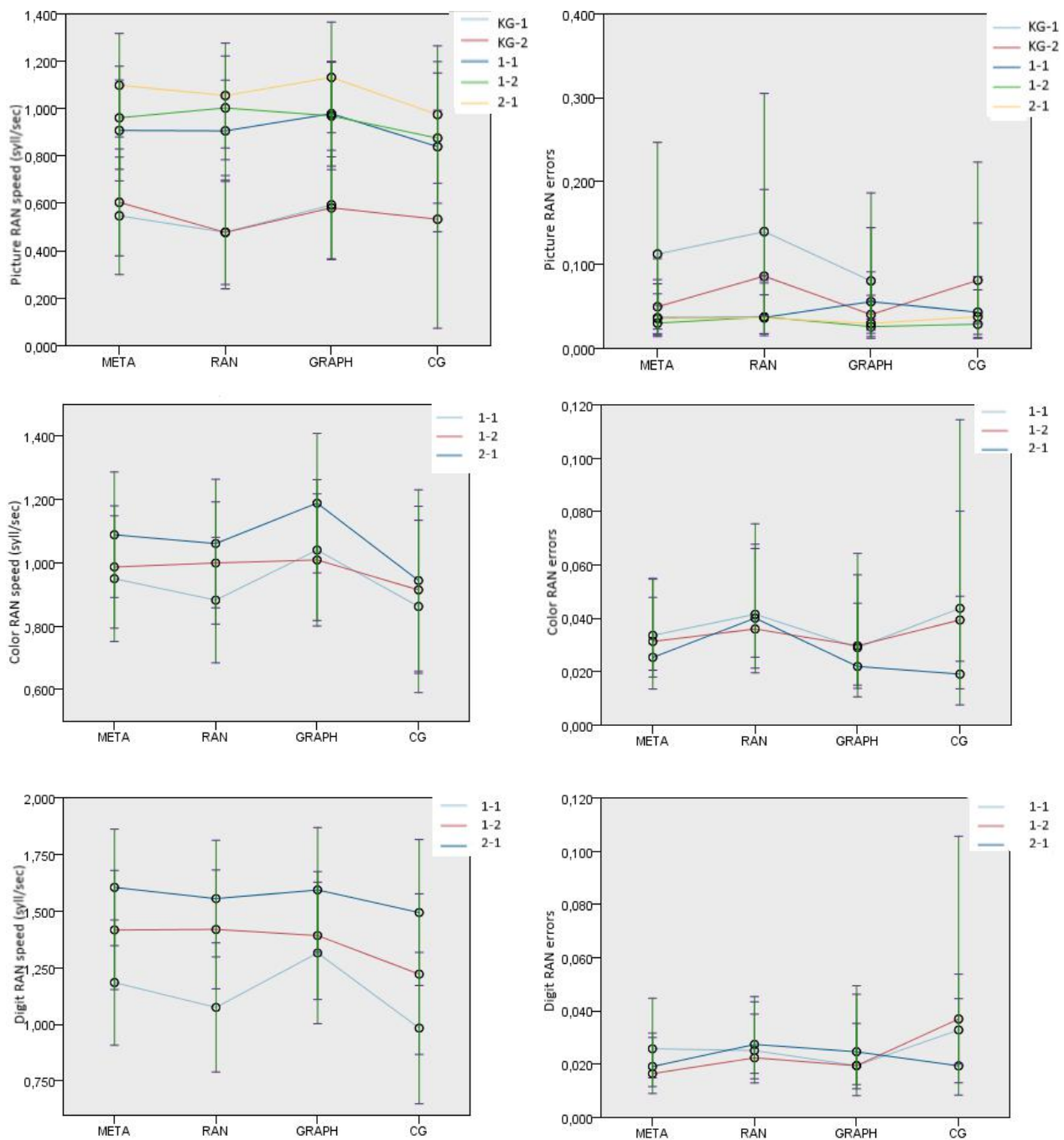


Figure 30. Estimated means of different groups in the 5 administrations of the RAN test of pictures, and in the 3 administrations of the RAN test of colors and digits. Colored lines are different administration times.

Significant negative effects of the META, RAN, and GRAPH group in administration KG-1 and KG-2, and of the CG in administration KG-2 on the RAN speed of pictures emerged. Positive effects of the META and the RAN group in administration KG-1, and of the CG in administration KG-2 on picture RAN errors emerged. Negative effects of the RAN group and CG in the first administration emerged on the RAN of digits (Table 24).

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Picture RAN speed	META * KG-1	- 0.428	0.134	- 3.186	0.004	- 0.703	- 0.153
	META * KG-2	- 0.371	0.137	- 2.712	0.011	- 0.650	- 0.091
	RAN * KG-1	- 0.497	0.135	- 3.683	0.001	- 0.774	- 0.221
	RAN * KG-2	- 0.498	0.138	- 3.614	0.001	- 0.779	- 0.217
	GRAPH * KG-1	- 0.381	0.136	- 2.801	0.009	- 0.660	- 0.103
	GRAPH * KG-2	- 0.394	0.139	- 2.828	0.008	- 0.678	- 0.110
	CG * KG-2	- 0.442	0.129	- 3.438	0.002	- 0.708	- 0.176
Picture RAN errors	META * KG-1	1.096	0.429	2.552	0.015	0.224	1.967
	RAN * KG-1	1.310	0.428	3.062	0.004	0.441	2.178
	CG * KG-2	0.769	0.342	2.250	0.034	0.063	1.475
Digit RAN speed	RAN * 1-1	- 0.419	0.184	- 2.280	0.027	- 0.787	- 0.051
	CG * 1-1	- 0.510	0.148	- 3.454	0.002	- 0.813	- 0.207

*Satterthwaite approximation

Table 24. Significant longitudinal combined effects of group and administration time on the 5 administrations of the RAN test of pictures, and on the 3 administrations of the RAN test of digits.

The post hoc analysis showed that the GRAPH group in administration 1-1 ($M = 1.315$, $StdErr = 0.151$, $CI [1.003, 1.628]$) was significantly faster than the CG ($M = 0.984$, $StdErr = 0.078$, $CI [0.649, 1.318]$) in the RAN of digits ($p = 0.028$).

There was no effect on speed ($p = 0.095$) and errors ($p = 0.286$) in the RAN of colors, and on errors in the RAN of digits ($p = 0.437$), as well as no effect of bilingualism ($0.517 < p < 0.667$).

To summarize RAN results, there were significant negative effects in the speed of every group in the two kindergarten administrations, showing no immediate effect of the RAN intervention on speed.

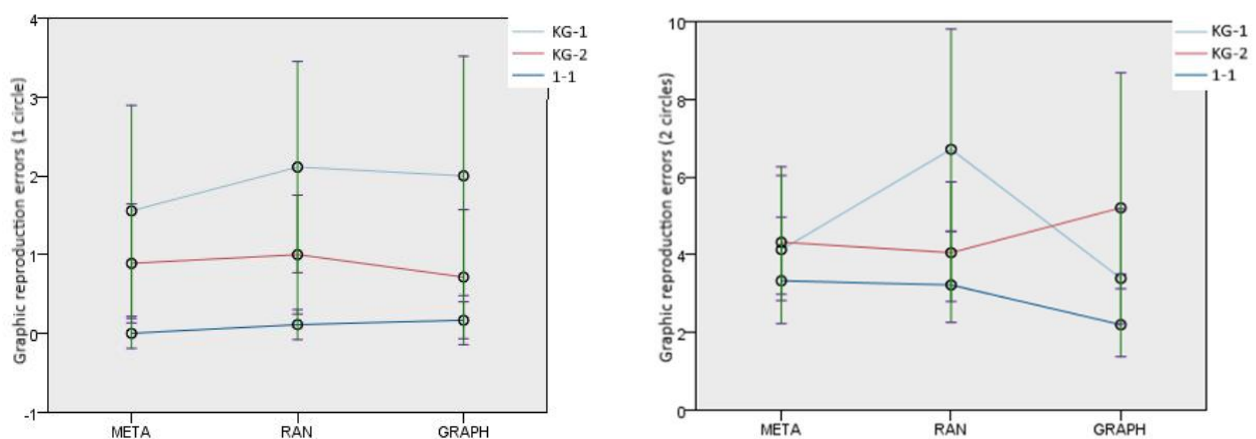
As for errors, the positive effects of the META and RAN group in KG-1, as well as the positive effect of the CG in KG-2, showed an improvement of the META and the RAN groups after Intervention 1 (KG-2). The biggest coefficient of the RAN group suggests an effect of the RAN intervention in primary school.

Visual skills

The tests of reproduction of one and two semicircles were administered 3 times, in the middle and at the end of the last year of kindergarten, and in the middle of the 1st grade; the test of visual discrimination of capital letters was administered twice, in the middle and at the end of the last year of kindergarten. A series of GMMRMs were performed with visual test scores as dependent variables, training groups and administration times as combined fixed effects, and bilingualism as a random effect.

The combination of group and administration time predicts the graphic reproduction of two semicircles ($F(8, 57) = 2.469, p = 0.023$).

Figure 31 shows the estimated means of the groups for each administration.



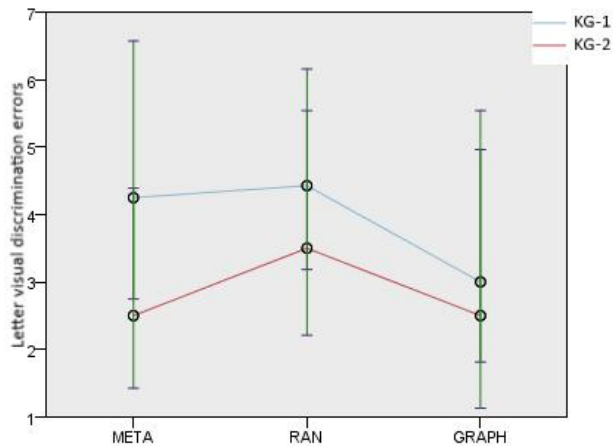


Figure 31. Estimated error means of different groups in the 3 administrations of the test of graphic reproduction of one and two semicircles, and in the 2 administrations of the visual discrimination test of capital letters.

Significant positive effects on errors in the graphic reproduction of 2 semicircles emerged in the first and second administration by the META and the RAN group, and in the second administration by the GRAPH group. Table 25 displays significant effects.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Graphic reproduction errors (2 semicircles)	META * KG-1	0.632	0.290	2.180	0.033	0.051	1.212
	META * KG-2	0.678	0.287	2.360	0.022	0.103	1.252
	RAN * KG-1	1.119	0.290	3.859	0.000	0.538	1.188
	RAN * KG-2	0.613	0.287	2.136	0.037	0.038	1.189
	GRAPH * KG-2	0.864	0.337	2.565	0.013	0.190	1.539

*Satterthwaite approximation

Table 25. Significant longitudinal combined effects of group and administration time on the 3 administrations of the test of graphic reproduction of two semicircles.

No effect of group on the graphic reproduction of one semicircle ($p = 0.117$) and on the visual discrimination of letters ($p = 0.384$), as well as no effect of bilingualism ($0.596 < p < 1.000$) emerged.

To sum up, tests of visual memory and visual discrimination in kindergarten showed no effect of the graphic intervention on children's visual skills.

Non-word repetition

The non-word repetition test was administered in the middle and at the end of the last year of kindergarten and in the middle of the 1st grade of primary school.

A GMMRM with the combination of group and administration time as a fixed effect and bilingualism as a random effect correctly predicts the number of correct syllables repeated ($F(10, 54) = 2.916, p = 0.005$).

Figure 32 shows the estimated means of the groups for each administration.

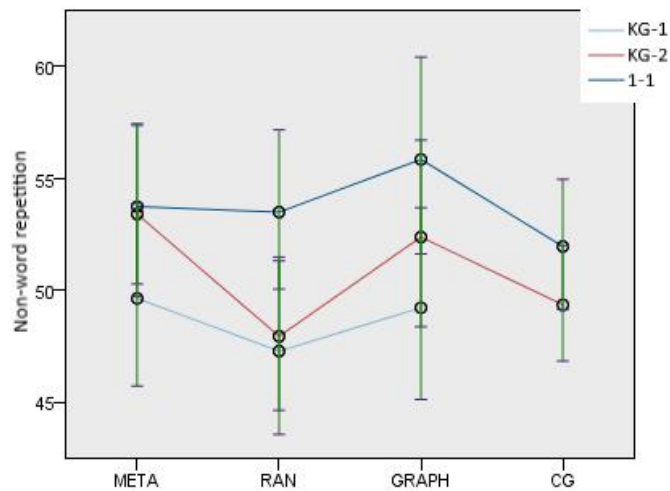


Figure 32. Estimated means of the groups in the three administrations of the non-word repetition test.

Significant negative effects of non-word repetition accuracy emerged in the first two administrations by the RAN group, and in the 2nd administration by the CG, as shown in Table 26.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	P	Conf. interval	
						lower	upper
Non-word repetition (correct syllables)	RAN * KG-1	- 0.094	0.040	- 2.359	0.026	- 0.177	- 0.012
	RAN * KG-2	- 0.080	0.036	- 2.234	0.028	- 0.152	- 0.009
	CG * KG-2	- 0.052	0.020	- 2.560	0.012	- 0.091	- 0.012

*Satterthwaite approximation

Table 26. Significant longitudinal combined effects of group and administration time on the 3 administrations of the non-word repetition test.

In administration KG-2, the mean of the META group ($M = 53.401$, $SE = 1.875$, $CI [49.713, 57.363]$) was significantly higher than that of the CG ($M = 49.346$, $SE = 0.904$, $CI [46.831, 51.996]$) ($Est = 4.055$, $SE = 1.933$, $t(66) = 2.098$, $p = 0.040$, $CI [0.196, 7.914]$).

To sum up, after Intervention 1, the META group only had a significantly higher mean than the CG, thus showing a possible effect of the metaphonological training on children's phonological skills.

Syntactic skills

In the experimental sample, syntactic comprehension was tested twice, in the last year of kindergarten and 2nd grade of primary school.

Syntactic production of DOcls was tested in the middle and at the end of the last year of kindergarten, and in the middle of the 2nd grade of primary school.

A GMMRM with the combination of group and administration time as fixed effects correctly predicts syntactic comprehension ($F(7, 30) = 3.097$, $p = 0.014$) and syntactic production of DOcl pronouns ($F(10, 34) = 2.320$, $p = 0.033$). Figure 33 shows the estimated means of different groups for each administration time.

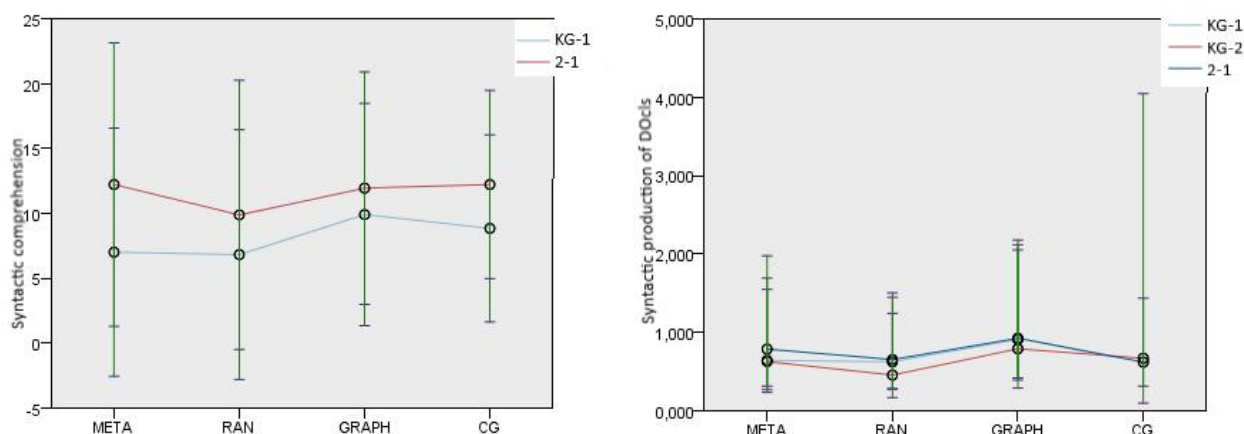


Figure 33. Estimated means of the groups in the two administrations of the syntactic-comprehension test and in the three administrations of the syntactic-production test.

As for syntactic comprehension, significant negative effects emerged in the first administration by the META and the RAN group.

As for syntactic production, significant negative effects emerged in the second administration by the RAN group and significant positive effects in the first and second administrations by the GRAPH group.

Table 27 displays significant effects.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	P	Conf. interval	
						lower	upper
Syntactic comprehension	META * KG-1	- 5.226	2.213	- 2.362	0.023	- 9.705	- 0.747
	RAN * KG-1	- 5.410	2.201	- 2.458	0.019	- 9.866	- 0.954
Syntactic production of DOcl pronouns	RAN * KG-2	- 0.310	0.138	- 2.256	0.031	- 0.591	- 0.030
	GRAPH * KG-1	0.386	0.167	2.308	0.030	0.041	0.731
	GRAPH * 2-1	0.403	0.175	2.298	0.025	0.052	0.754

*Satterthwaite approximation

Table 27. Significant longitudinal combined effects of group and administration time on the 2 administrations of the syntactic-comprehension test and on the 3 administrations of the syntactic-production test

As for syntactic comprehension, only the META group showed a significant difference ($Est = - 5.222$, $SE = 1.599$, $t(46) = - 3.266$, $p = 0,022$, $CI [- 9.411, - 2.004]$) between the syntactic-

comprehension score obtained in the first ($M = 6.998$, $SE = 2.112$, $CI [-2.576, 16.753]$) and last ($M = 12.220$, $SE = 2.038$, $CI [1.824, 23.157]$) administration.

No significant effects of bilingualism on both syntactic comprehension ($p = 0.525$) and production ($p = 0.515$) emerged.

To sum up, the META and the RAN group showed an initial disadvantage in syntactic comprehension. After the intervention, only the META group showed a significant improvement, meaning that the metaphonological training might have tapped children's syntactic competence.

Receptive and productive vocabulary

The receptive vocabulary was tested in the last year of kindergarten and 2nd grade of primary school. The productive vocabulary was tested once in the 1st grade and once in the 2nd grade.

A GMMRM did not find any significant effect of the combination of group and administration time on receptive vocabulary ($p = 0.126$), as well as no effect of group ($p = 0.272$) or bilingualism ($p = 0.434$) on receptive vocabulary scores.

The interventions, then, had no effects on children's vocabulary.

Digit and word span

In the experimental sample, the memory span was tested three times, once per year, from the last grade of KG to the 2nd grade of primary school.

A GMMRM correctly predicts digit span forward ($F(5, 107) = 3.334$, $p = 0.001$), digit span backward ($F(11, 57) = 3.987$, $p = 0.000$), and word span ($F(11, 54) = 4.300$, $p = 0.000$).

There is a main effect of administration time on digit span forward ($F(2, 55) = 12.272$, $p = 0.000$), and a combined effect of group and administration time on digit span backward ($F(11, 57) = 3.987$, $p = 0.000$), and word span ($F(11, 54) = 4.300$, $p = 0.000$).

Figure 34 displays the estimated means of the groups for each administration.

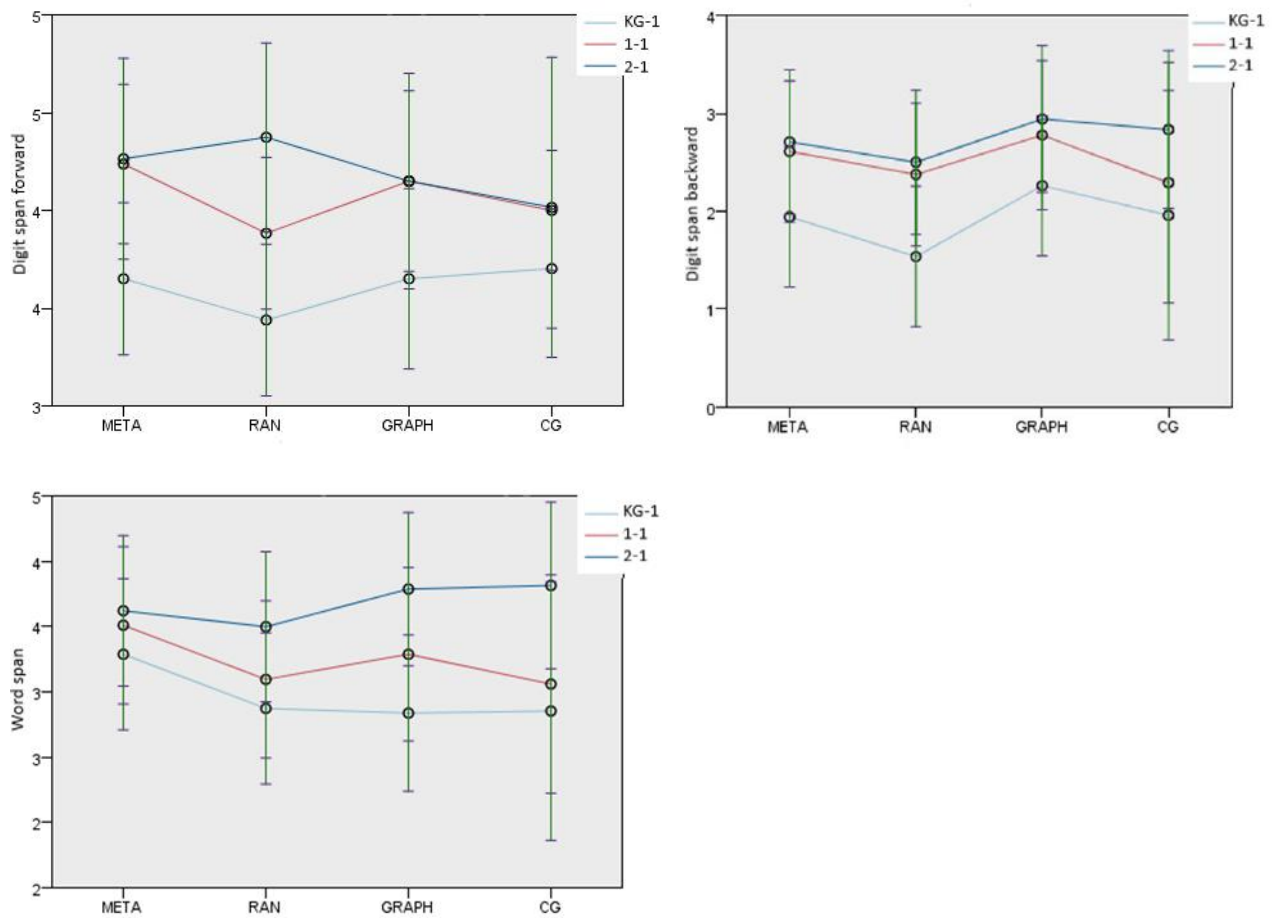


Figure 34. Estimated means of different groups in the three administrations of memory span tests.

Digit span forward showed a negative effect of administration KG-1, and digit span backward a negative effect of the META, RAN, CG in administration KG-1. Word span showed a negative effect of the RAN group and the CG in administration KG-1 and 1-1, and a negative effect of the GRAPH group in administration KG-1, as shown in Table 28.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Digit span forward	KG-1	- 0.594	0.151	- 3.492	0.000	- 0.898	- 0.290
Digit span backward	META * KG-1	- 0.892	0.381	- 2.339	0.024	- 1.660	- 0.123
	RAN * KG-1	- 1.299	0.380	- 3.416	0.001	- 2.065	- 0.532
	CG * KG-1	- 0.877	0.331	- 2.650	0.013	- 1.557	- 0.198

Word span	RAN * KG-1	- 0.943	0.325	- 2.902	0.005	- 1.595	- 0.290
	RAN * 1-1	- 0.720	0.343	- 2.098	0.040	- 1.407	- 0.033
	GRAPH * KG-1	- 0.979	0.342	- 2.862	0.006	- 1.663	- 0.294
	CG * KG-1	- 0.962	0.270	- 3.564	0.001	- 1.517	- 0.408
	CG * 1-1	- 0.756	0.275	- 2.750	0.010	- 1.318	- 0.193

*Satterthwaite approximation

Table 28. Significant longitudinal effects of group and administration time on the 3 administrations of memory span tests.

No significant effect of bilingualism emerged for either digit span forward ($p = 0.662$), digit span backward ($p = 0.531$), and word span ($p = 0.540$).

To sum up, the GRAPH group seems to be advantaged in digit span backward, as their span showed no negative effects in kindergarten, differently from that of the others. The META group had an initial advantage in word span, as it did not present the initial negative effect.

5.3.2. Effects of the training on literacy skills

Word and non-word decoding

Word decoding was tested once, in the middle of the 2nd grade.

A GLMM was analyzed to explore the combined effect of group and bilingualism, with training groups as a fixed effect nested in the bilingualism variable.

The model correctly predicts word decoding speed ($F(7, 20) = 5.909$, $p = 0.001$), word decoding generic errors ($F(7, 20) = 5.959$, $p = 0.001$), non-word decoding speed ($F(7, 20) = 2.658$, $p = 0.046$), stress errors in word decoding ($F(7, 17) = 14.505$, $p = 0.000$). Estimated marginal means are in Figure 35.

There was no effect of group and bilingualism on non-word decoding errors ($p = 0.070$), and on attentional errors in word-pair decoding ($p = 0.386$).

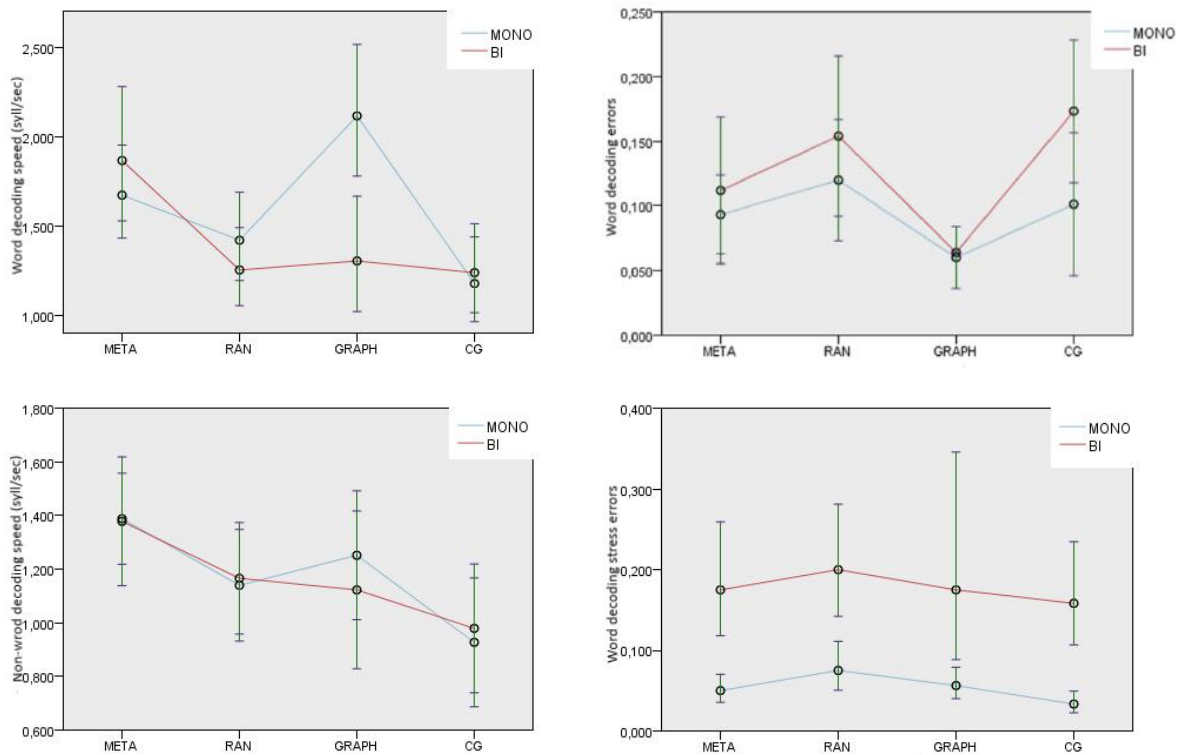


Figure 35. Estimated marginal means of word and non-word decoding speed and errors in the 2nd grade of primary school.

Table 29 shows significant effects of group and bilingualism on word and non-word decoding.

As for word decoding speed, significant positive effects emerged in the mono-META group, the mono-GRAPH group, and the bi-META group, as shown in Table 29. Among monolingual children ($F(3, 20) = 7.605, p = 0.001$), the META ($M = 1.672, SE = 0.124, CI [1.432, 1.952]$) and the GRAPH ($M = 2.115, SE = 0.176, CI [1.779, 2.516]$) group were significantly faster than the CG ($M = 1.177, SE = 0.113, CI [0.964, 1.438]$), and the GRAPH group was significantly faster than the RAN group ($M = 1.419, SE = 0.118, CI [1.194, 1.688]$). Among bilingual children ($F(3, 20) = 3.373, p = 0.039$), the bi-META group ($M = 1.866, SE = 0.179, CI [1.527, 2.279]$) was significantly faster than every other group (RAN: $M = 1.253, SE = 0.104, CI [1.054, 1.490]$; GRAPH: $M = 1.304, SE = 0.153, CI [1.020, 1.666]$; CG: $M = 1.238, SE = 0.119, CI [1.103, 1.512]$).

As for generic errors in word decoding, significant negative effects emerged by the mono-META and the mono-GRAPH group, and by the bi-GRAPH group. Among bilingual children ($F(3, 20) = 9.775, p = 0.000$), the bi-GRAPH group ($M = 0.064, SE = 0.000, CI [0.064, 0.229]$) made significantly fewer errors than the bi-RAN group ($M = 0.154, SE = 0.030, CI [0.092, 0.216]$) and the bi-CG ($M = 0.173, SE = 0.026, CI [0.118, 0.229]$), whereas it showed no significant difference compared to the bi-META group ($M = 0.112, SE = 0.027, CI [0.055, 0.169]$).

As for non-word decoding speed, significant positive effects of the META group emerged in both monolingual and bilingual children. META children were significantly faster than the CG in both the monolingual (META: $M = 1.389, SE = 0.081, CI [1.219, 1.599]$; CG: $M = 0.927, SE = 0.115, CI [0.687, 1.168]$) and bilingual (META: $M = 1.379, SE = 0.115, CI [1.138, 1.619]$; CG: $M = 0.979, SE = 0.115, CI [0.739, 1.220]$) sample.

As for stress errors in word decoding, significant negative effects emerged in every monolingual group.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Word decoding speed (syll/sec)	META (MONO)	0.213	0.096	2.225	0.038	0.013	0.414
	GRAPH (MONO)	0.536	0.127	4.221	0.000	0.271	0.800
	META (BI)	0.410	0.136	3.023	0.007	0.127	0.693
Word decoding errors	META (MONO)	- 0.080	0.030	- 2.646	0.015	- 0.143	- 0.017
	GRAPH (MONO)	- 0.113	0.029	- 3.925	0.001	- 0.174	- 0.053
	GRAPH (BI)	- 0.109	0.026	- 4.128	0.001	- 0.165	- 0.054
Non-word decoding speed (syll/sec)	META (MONO)	0.409	0.141	2.900	0.009	0.115	0.704
	META (BI)	0.399	0.163	2.450	0.024	0.059	0.739
Word decoding stress errors	META (MONO)	- 1.153	0.247	- 4.761	0.000	- 1.673	- 0.632
	RAN (MONO)	- 0.747	0.264	- 2.832	0.011	- 1.304	- 0.191
	GRAPH (MONO)	- 1.035	0.247	- 4.194	0.001	- 1.556	- 0.514
	CG (MONO)	- 1.558	0.264	- 5.906	0.000	- 2.115	- 1.002

*Satterthwaite approximation

Table 29. Significant effects of group and bilingualism on word and non-word decoding in the middle of the 2nd grade of primary school.

To sum up, monolingual 2nd graders who received the META and the GRAPH interventions, and bilinguals who received the META intervention showed significant effects of speed in decoding generic word lists. Monolinguals who received the META intervention and monolinguals and bilinguals who received the GRAPH intervention also had significant effects on accuracy.

Monolingual META and GRAPH children were significantly faster than their CG in decoding generic word lists.

Monolingual and bilingual META children were significantly faster than their respective controls in decoding non-words.

All monolinguals were significantly more accurate in decoding non-paroxytone words.

Sublexical decoding

Sublexical decoding was tested in the middle of the 1st grade through a list of strings constituting sublexical components of Italian words (Calgaro et al., 2018) and in the middle of the 2nd grade through a list of pseudowords (Traficante et al., 2017) (see 2.9).

A GMMRM was analyzed with scores obtained in the two tasks as the dependent variable, and training group and administration time as a combined fixed effect nested in the bilingualism variable.

The model correctly predicts decoding speed ($F(15, 37) = 11.259, p = 0.000$) and errors ($F(15, 30) = 2.521, p = 0.015$). Figure 36 shows the estimated means of each group in the monolingual and bilingual sample.

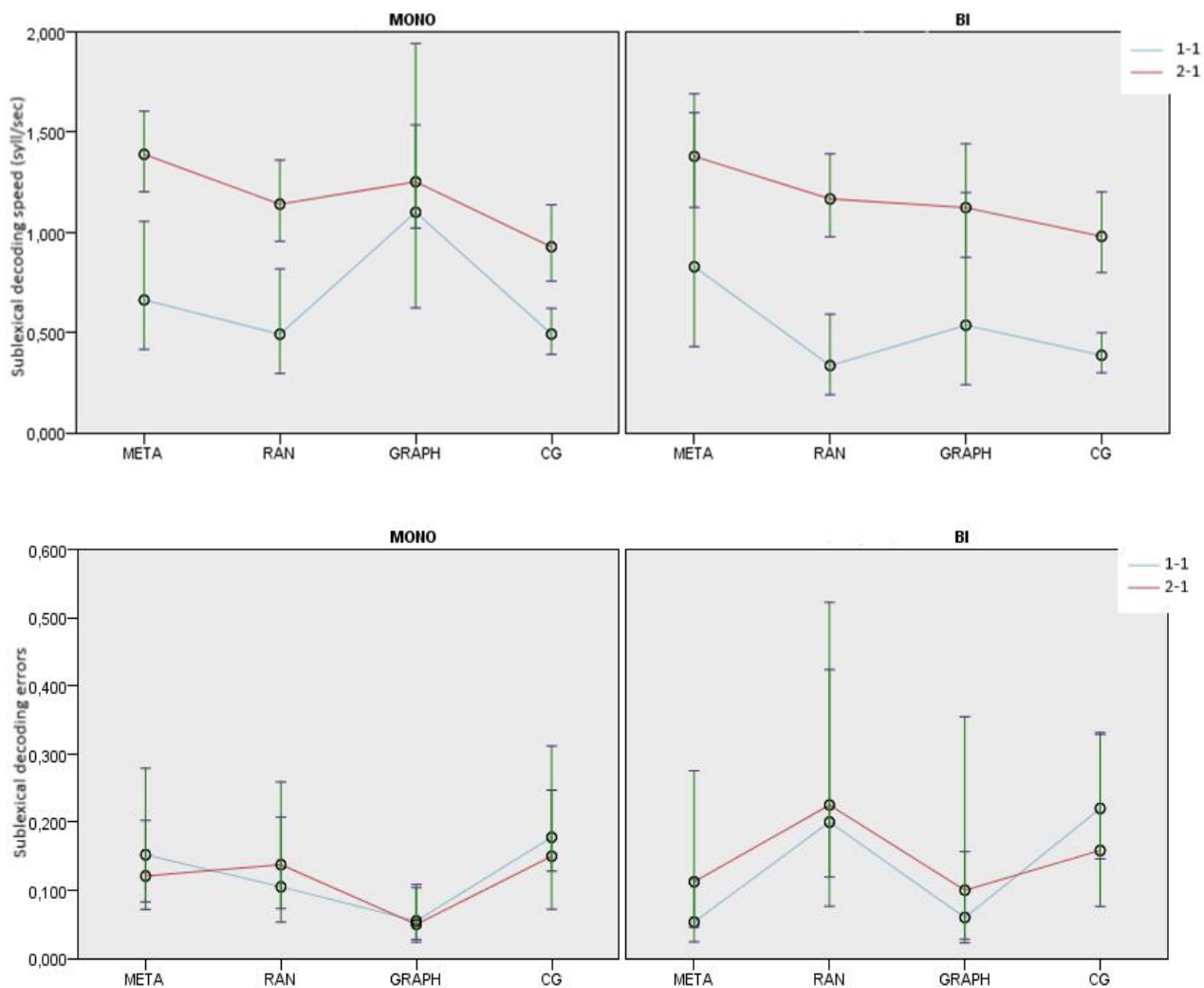


Figure 36. Longitudinal estimated means of each group in the two administrations (1-1; 2-1) of the sublexical decoding tests.

Table 30 displays the significant combined longitudinal effects of group, administration time, and bilingualism on sublexical decoding scores in the two administrations in the middle of the 1st and 2nd grade of primary school.

A significant positive effect on sublexical decoding speed emerged in the last administration by the mono-META and bi-META group, and by the mono-GRAPH group, whereas significant negative effects emerged in the first administration by the mono-RAN and bi-RAN groups, and by the mono-CG and bi-CG. The post hoc comparison revealed that, in administration 2-1, mono-META ($M = 1.389$, $SE = 0.096$, $CI [0.416, 1.054]$) and bi-META ($M = 1.379$, $SE = 0.135$, $CI [1.124, 1.690]$) children were significantly faster than the respective CGs (mono-CG: $M = 0.927$, $SE = 0.091$, $CI [0.756, 1.137]$ ($p = 0.002$); bi-CG: $M = 0.979$, $SE = 0.135$, $CI [0.709, 1.249]$ ($p = 0.002$)).

= 0.096, *CI* [0.799, 1.201] ($p = 0.025$)), and the mono-GRAPH ($M = 1.252$, $SE = 0.122$, *CI* [1.021, 1.535]) group showed a marginally significant higher speed than the mono-CG ($M = 0.927$, $SE = 0.091$, *CI* [0.756, 1.137] ($p = 0.046$)).

As for sublexical decoding errors, there is a significant negative effect in both administrations by mono-GRAPH children, and a significant negative effect in the first administration by bi-META children. The post hoc comparison revealed that, in the first administration, the bi-META group ($M = 0.053$, $SE = 0.021$, *CI* [0.024, 0.117]) makes significantly fewer errors than the bi-CG ($M = 0.178$, $SE = 0.029$, *CI* [0.128, 0.247]) ($p = 0.002$), and both the mono-GRAPH ($M = 0.055$, $SE = 0.018$, *CI* [0.028, 0.108] ($p = 0.001$)), and the bi-GRAPH ($M = 0.060$, $SE = 0.028$, *CI* [0.023, 0.157] ($p = 0.004$)) groups make significantly fewer errors than their respective CGs (MONO-CG: $M = 0.178$, $SE = 0.029$, *CI* [0.128, 0.247]).

Dependent variable	Predictor	Coeff.*	Std.Err.	<i>t</i>	<i>p</i>	Conf. interval	
						lower	upper
Sublexical decoding speed (syll./sec.)	META * 2-1 (MONO)	0.349	0.134	2.606	0.017	0.071	0.628
	RAN * 1-1 (MONO)	- 0.689	0.277	- 2.490	0.015	- 1.240	- 0.138
	GRAPH * 2-1 (MONO)	0.383	0.145	2.647	0.015	0.082	0.684
	CG * 1-1 (MONO)	- 0.688	0.159	- 4.312	0.000	- 1.006	- 0.369
	META * 2-1 (BI)	0.342	0.155	2.210	0.038	0.020	0.664
	RAN * 1-1 (BI)	- 1.072	0.305	- 3.519	0.001	- 1.679	- 0.465
	CG * 1-1 (BI)	- 0.930	0.168	- 5.546	0.000	- 1.265	- 0.596
Sublexical decoding errors	GRAPH * 1-1 (MONO)	- 1.057	0.484	- 2.187	0.034	- 2.029	- 0.085
	GRAPH * 2-1 (MONO)	- 1.153	0.492	- 2.343	0.031	- 2.186	- 0.119
	META * 1-1 (BI)	- 1.088	0.521	- 2.089	0.042	- 2.133	- 0.043

*Satterthwaite approximation

Table 30. Significant longitudinal combined effects of training group (META; RAN; GRAPH; CG), administration time (1-1; 2-1), and bilingualism (MONO; BI) on the two administrations of the sublexical decoding tests.

To sum up, monolingual and bilingual RAN and CG children, in the middle of the 1st grade, were slower, whereas the monolingual and bilingual META groups and the monolingual GRAPH group were faster in the middle of the 2nd grade. Hence, the RAN intervention

showed no effect on children's speed, whereas the metaphonological intervention positively affected sublexical decoding speed. Also, the monolingual GRAPH group was significantly faster than the respective CG, and made fewer errors in both administrations, whereas the bilingual META group made fewer errors in the 1st grade, suggesting a possible effect of META intervention 1 only on bilinguals' accuracy.

Text decoding

Text decoding was tested twice, at the end of the 1st grade and in the middle of the 2nd grade through Cornoldi & Carretti's (2016) test. Texts' length and difficulty are calibrated to the grade and moment of the year (beginning, end). Decoding speed (syll./sec.) and error proportion were computed. A longitudinal analysis of groups' results was conducted.

A GMMRM with the combination of group and administration time as a fixed effect nested in the bilingualism variable significantly predicts decoding speed ($F(15, 24) = 5.818, p = 0.000$).

Figure 37 shows the estimated means of monolingual and bilingual training groups in each administration.

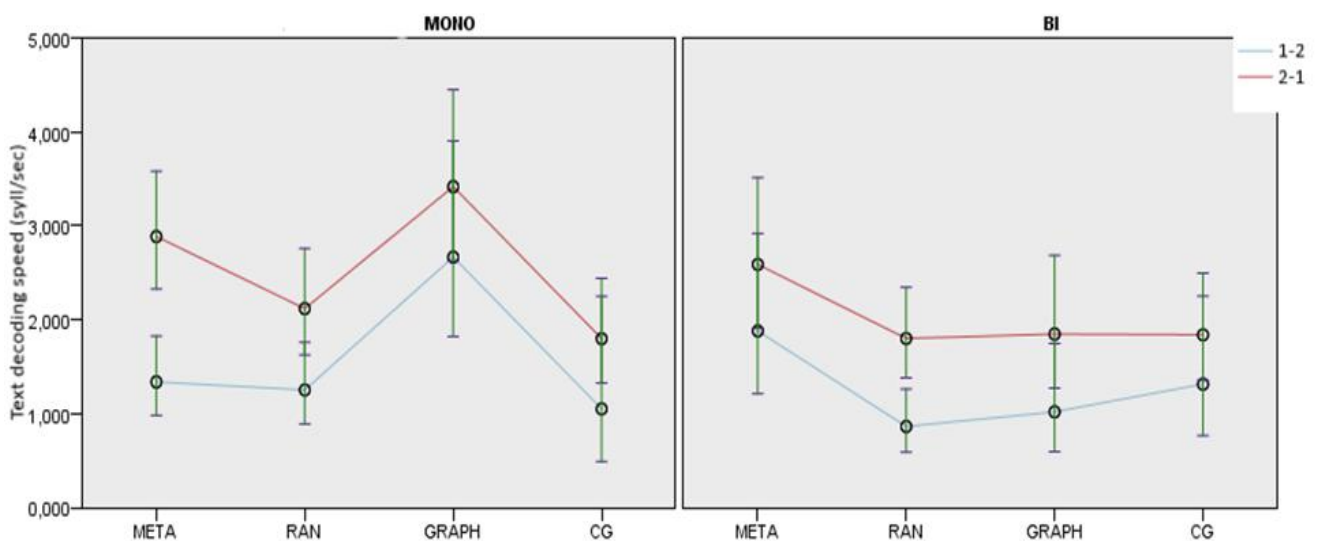


Figure 37. Estimated means of different groups in the two administrations of the text decoding test.

A significant positive effect of group emerged in the last administration by mono-META and mono-GRAPH children, and a negative effect in the first administration by bi-META children. Significant effects are in Table 31.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						lower	upper
Text decoding speed (syll/sec)	META * 2-1 (MONO)	0.450	0.180	2.506	0.021	0.077	0.824
	GRAPH * 2-1 (MONO)	0.619	0.194	3.189	0.004	0.215	1.023
	META * 1-2 (BI)	- 0.756	0.234	- 3.237	0.003	- 1.230	- 0.283

*Satterthwaite approximation

Table 31. Significant longitudinal combined effects of group (META; RAN; GRAPH; CG), administration time (1-2; 2-1), and bilingualism (mono; bi) on text decoding speed and errors.

The post hoc comparison showed that, in the first administration, the mono-GRAPH ($M = 2.666$, $SE = 0.485$, $CI [1.822, 3.902]$) group was significantly faster than the mono-CG ($M = 1.051$, $SE = 0.382$, $CI [0.491, 2.250]$) ($p = 0.017$); in the last administration, a main significant effect of monolingualism ($F(3, 21) = 4.778$, $p = 0.011$) emerged. The mono-META ($M = 2.886$, $SE = 0.299$, $CI [2.325, 3.580]$) ($p = 0.013$) and the mono-GRAPH ($M = 3.416$, $SE = 0.434$, $CI [2.622, 4.449]$) ($p = 0.004$) groups were significantly faster than the mono-CG ($M = 1.798$, $SE = 0.264$, $CI [1.325, 2.440]$).

The within-group analysis revealed main effects of group in mono-META ($F(1, 36) = 18.559$, $p = 0.000$), mono-RAN ($F(1, 38) = 6.581$, $p = 0.014$), and bi-RAN ($F(1, 36) = 11.401$, $p = 0.002$) children, with a significant decrease of errors in mono-META ($Est. = - 1.548$, $SE = 0.359$, $t(36) = - 4.308$, $p = 0.000$, $CI [- 2.277, - 0.819]$), mono-RAN ($Est. = - 0.866$, $SE = 0.338$, $t(38) = - 2.565$, $p = 0.014$, $CI [- 1.549, - 0.183]$) and bi-RAN ($Est. = - 0.937$, $SE = 0.277$, $t(36) = - 3.377$, $p = 0.002$, $CI [- 1.499, - 0.374]$) children between the first and last administration.

To sum up, monolingual GRAPH children were faster in both administrations compared to their controls, whereas monolingual META children were significantly faster in the latter administration.

There was a significant increase of accuracy between the 1st and 2nd grade in monolingual META and RAN children, and in bilingual RAN children, suggesting an effect of RAN on decoding accuracy.

Reading comprehension

In the full sample, reading comprehension was tested through the screening tests of the MT-2 battery (Cornoldi & Colpo, 1998), which were administered collectively in children's classrooms. An additional diagnostic test (Cornoldi & Carretti, 2016) was administered to the core sample. Tests took place at the end of the 1st grade, and in the middle of the 2nd grade.

A GMMRM revealed no combined effect of group and administration time in either mono- and bi- children on reading comprehension total score ($p = 0.423$), central score ($p = 0.131$), and specific score ($p = 0.389$). As the effects are not significant, the estimated means were not calculated.

Word writing

Word writing was tested three times, in the middle and at the end of the 1st grade (Stella & Apolito, 2004) and in the middle of the 2nd grade (Sartori et al., 2007) through two different dictation tests (see 2.1.14).

A GMMRM with training group and administration time nested in the bilingualism variable as fixed effects significantly predicts word errors ($F(23, 44) = 8.754, p = 0.000$).

The estimated means are in Figure 38.

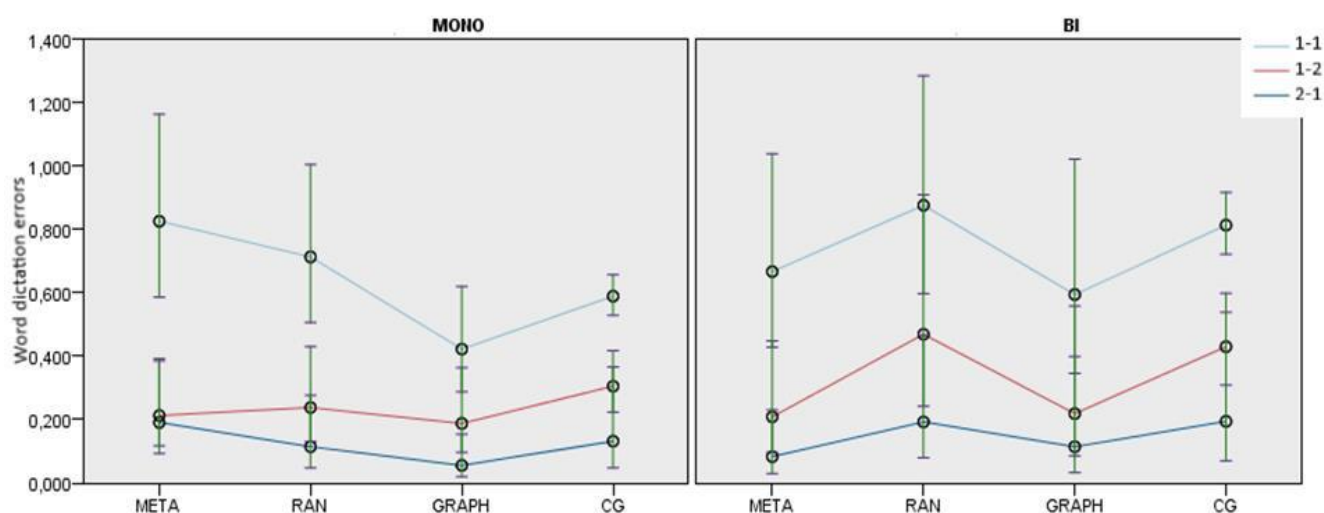


Figure 38. Estimated means of different groups in the three administrations of the word-dictation test

Significant positive effects on word writing errors emerged in administration 1-1 by the META group, the RAN group, and the CG in both mono- and bi- children, and an additional marginally significant positive effect of the GRAPH group is present in bi- children, as shown in Table 32.

Dependent variable	Predictor	Coeff.*	Std.Err.	t	p	Conf. interval	
						Lower	upper
Word dictation errors	META * 1-1 (MONO)	1.445	0.518	2.791	0.010	0.379	2.512
	RAN * 1-1 (MONO)	1.299	0.518	2.508	0.019	0.233	2.365
	CG * 1-1 (MONO)	1.108	0.491	2.256	0.035	0.085	2.132
	META * 1-1 (BI)	1.232	0.537	2.295	0.029	0.134	2.330
	RAN * 1-1 (BI)	1.505	0.525	2.865	0.008	0.426	2.583
	GRAPH * 1-1 (BI)	1.117	0.560	1.996	0.054	- 0.021	2.254
	CG * 1-1 (BI)	1.430	0.492	2.907	0.009	0.406	2.455

*Satterthwaite approximation

Table 32. Significant longitudinal effects of group and administration time on the 3 administrations of the word-dictation test.

The post hoc analysis showed that, in the second administration, bi-META children ($M = 0.209$, $SE = 0.079$, $CI [0.097, 0.448]$) made significantly fewer errors than the bi-CG ($M =$

0.430, $SE = 0.071$, $CI [0.309, 0.598]$ ($Est. = -0.221$, $SE = 0.106$, $t(49) = -2.081$, $p = 0.043$, $CI [-0.435, -0.008]$).

A significant error decrease was observed between the first and third administration in every group, and between the second and third administration in the CGs, with a marginal significance for the bi-CG. Significant effects are in Table 33.

Bilingualism	Predictor	Contrast		<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>Lower</i>	<i>Upper</i>
MONO	META	1-1	2-1	0.634	0.157	4.036	126	0.000	0.323	0.945
	RAN	1-1	2-1	0.598	0.132	4.519	125	0.000	0.336	0.860
	GRAPH	1-1	2-1	0.367	0.086	4.265	123	0.000	0.196	0.537
	CG	1-1	2-1	0.457	0.072	6.346	31	0.000	0.310	0.604
	CG	1-2	2-1	0.173	0.080	2.168	42	0.036	0.012	0.335
	CG	1-2	2-1	0.173	0.080	2.168	42	0.036	0.012	0.335
BI	META	1-1	2-1	0.583	0.154	3.783	119	0.000	0.278	0.889
	RAN	1-1	2-1	0.683	0.188	3.636	125	0.000	0.311	1.054
	GRAPH	1-1	2-1	0.479	0.176	2.721	126	0.007	0.131	0.828
	CG	1-1	2-1	0.618	0.107	5.786	32	0.000	0.401	0.836
	CG	1-2	2-1	0.235	0.118	1.988	43	0.053	-0.003	0.474
	CG	1-2	2-1	0.235	0.118	1.988	43	0.053	-0.003	0.474

Table 33. Longitudinal main effects of group on word dictation errors, and significant and marginally significant contrasts between administration times.

To summarize, there was an initial advantage of accuracy in the monolingual GRAPH group, After Intervention 2, bilingual META children were significantly more accurate than controls in word writing, thus showing a possible effect of the metaphonological intervention on bilinguals.

Non-word writing

Non-word writing was tested once, in the middle of the 2nd grade (Sartori et al., 1995/2007). A GLMM did not find any significant effect of group ($p = 0.156$) and bilingualism ($p = 1.000$) on non-word writing errors.

Text writing

Text writing was tested through the BVSCO-2 (Tressoldi et al., 2013) dictation test. The test was administered once, in the middle of the 2nd grade. Total error scores were computed, and an error analysis was conducted (see section 2).

A series of GLMM with error scores as dependent variables, and group as a fixed effect nested in the bilingualism variable significantly predict total error score ($F(7, 49) = 4.802, p = 0.000$), phonological errors ($F(7, 43) = 7.418, p = 0.000$), and non-phonological errors ($F(7, 47) = 2.493, p = 0.029$).

Figure 39 shows the estimated means of each group

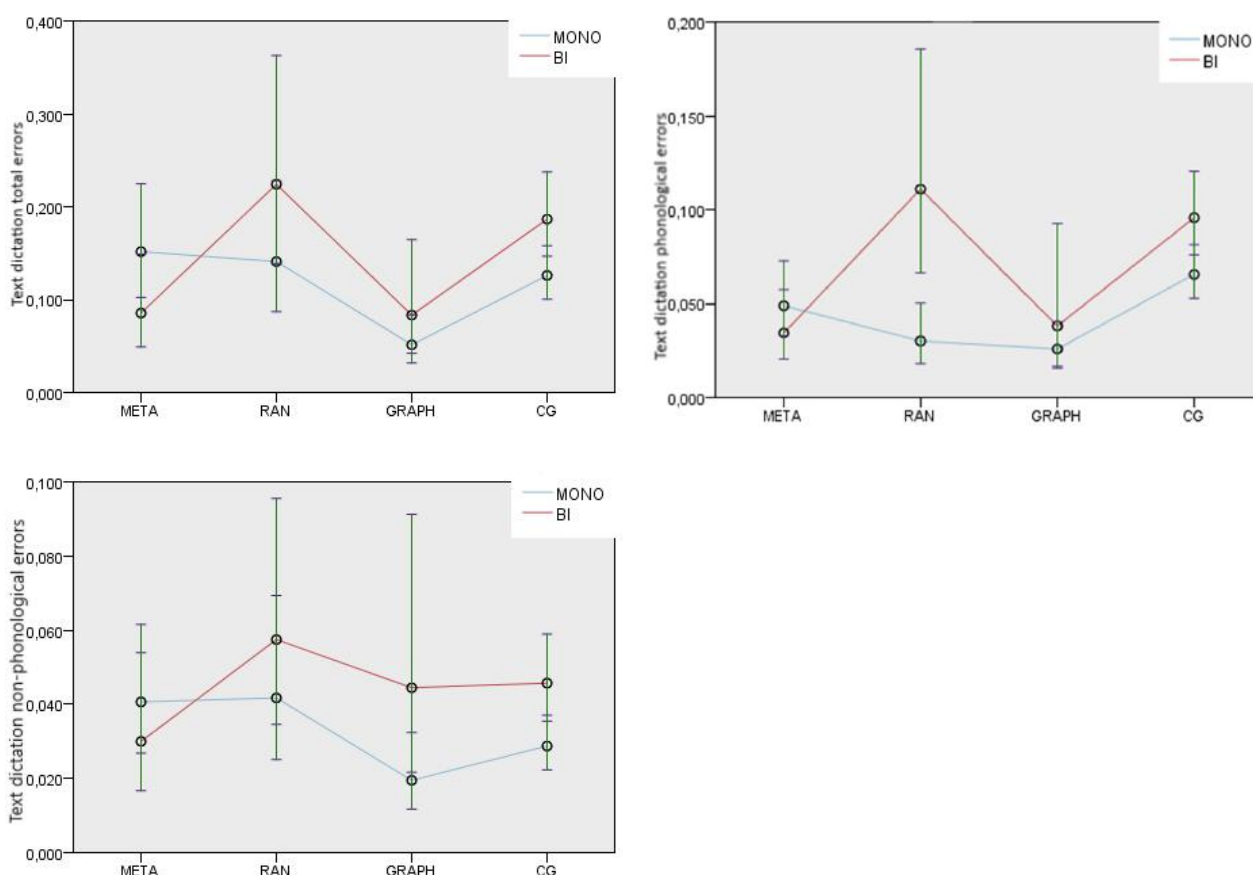


Figure 39. Estimated error means of each mono- and bi- group for text dictation in administration 2-1.

There are significant negative effects

- on total error score, in the mono-GRAPH, bi-META, and bi-GRAPH group;
- on phonological errors, in every training group except the bi-RAN group;

- on non-phonological errors, in the mono-GRAPH and mono-CG;

Table 34 displays significant predictive effects.

Dependent variable	Predictor	Coeff.	StdErr	t	p	Conf. Interval	
						Inferiore	Superiore
Text dictation total errors	GRAPH (MONO)	- 1.292	0.268	- 4.823	0.000	- 1.831	- 0.754
	CG (MONO)	- 0.393	0.165	- 2.385	0.021	- 0.724	- 0.062
	META (BI)	- 0.781	0.302	- 2.591	0.013	- 1.387	- 0.175
	GRAPH (BI)	- 0.807	0.359	- 2.244	0.029	- 1.529	- 0.084
Text dictation phonological errors	META (MONO)	- 0.674	0.228	-2,952	0.005	-1,134	- 0.214
	RAN (MONO)	- 1,160	0.280	-4,151	0.000	- 1,724	- 0,597
	GRAPH (MONO)	- 1.313	0,249	-5,279	0.000	- 1,815	- 0,812
	CG (MONO)	- 0,379	0,157	-2,421	0.020	- 0,695	- 0,063
	META (BI)	- 1,025	0,280	-3,668	0.001	- 1,589	- 0,462
	GRAPH (BI)	- 0.924	0,456	- 2,024	0.049	-1,845	- 0.003
Text dictation non- phonological errors	GRAPH (MONO)	- 0.853	0.283	- 3.017	0.004	- 1.421	- 0.284
	CG (MONO)	- 0.465	0.179	- 2.599	0.012	- 0.824	- 0.105

Table 34. Significant effects predicted by the combination of group, administration time, and bilingualism on different text-dictation errors in administration 2-1.

Post hoc analyses showed that

- mono-GRAPH children made significantly fewer errors than mono-CG children ($F(3, 49) = 7.617, p = 0.000$), and bi-META and bi-GRAPH children made significantly fewer errors than the bi-CG ($F(3, 49) = 5.257, p = 0.003$);

- mono-RAN and mono-GRAPH children made significantly fewer phonological errors than the mono-CG ($F(3, 43) = 7.339, p = 0.000$), and bi-META and bi-GRAPH children made significantly fewer phonological errors than the bi-CG ($F(3, 43) = 8.048, p = 0.000$).

Significant contrasts are in Table 35.

No differences between groups emerged in either mono- ($p = 0.086$) or bi- ($p = 0.341$) children for non-phonological errors.

CI

		Contrasts	Est.	Std.Err.	t	df	p	Lower	Upper
Text dictation total	MONO	GRAPH - CG	- 0.075	0.019	- 3.976	49	0.000	- 0.113	- 0.037
error score	BI	META - CG	- 0.101	0.033	- 3.110	49	0.003	- 0.167	- 0.036
		GRAPH - CG	- 0.103	0.036	- 2.869	49	0.006	- 0.176	- 0.031
Text dictation									
phonological errors	MONO	RAN - CG	- 0.036	0.010	- 3.420	43	0.001	- 0.056	- 0.015
		GRAPH - CG	- 0.040	0.009	- 4.400	43	0.000	- 0.058	- 0.022
	BI	META - CG	- 0.061	0.014	- 4.384	43	0.000	- 0.090	- 0.033
		GRAPH - CG	- 0.058	0.020	- 2.881	43	0.006	- 0.098	- 0.017

Table 35. Significant post hoc contrasts between training groups in MONO and BI children

No main effect emerged on stress/geminate errors ($p = 0.108$), omissions ($p = 0.293$), minor errors ($p = 0.494$), and inappropriate words ($p = 0.853$).

To sum up, the META training showed to be particularly useful for bilingual children, where it positively affected, particularly, phonological errors.

Both monolingual and bilingual GRAPH children were more accurate than the others, thus suggesting a possible effect of the GRAPH training on writing.

5.4. Discussion 3

5.4.1. Effects of the training on pre-literacy skills

Global metaphonological skills

Global metaphonological skills were tested in the middle (KG-1) and at the end (KG-2) of kindergarten, namely before and after the first training cycle. The three groups showed a general initial weakness (Table 22) in the recognition of the initial syllable. The META group showed a general advantage in this ability, which, however, did not determine a significant difference with the other groups. Before training, the 3 groups did not differ significantly from each other in any ability, even though a combined positive effect of group and first administration revealed a certain starting advantage by the META and the GRAPH group in

syllable blending. After training, the META and the RAN group showed a significant improvement in syllable blending, whereas the GRAPH group did not. Moreover, only the META group improved significantly in the recognition of word minimal pairs and rhymes. After training, the META and RAN group had significantly higher scores than the CG in syllable blending, whereas only the META group was significantly more proficient than the CG in the recognition of word minimal pairs, initial syllables, and rhymes. The groups did not show any significant difference or variation in syllable segmentation and recognition of non-word minimal pairs.

These data show that at the end of kindergarten, the META intervention has determined a significant improvement of some global metaphonological skills, namely recognition of word minimal pairs, initial syllable, and rhymes. This might be due, however, to increased familiarity with the tasks, which are similar in testing as in training for the META group.

There was a significant improvement of both META and RAN groups in syllable blending. This might be evidence that the RAN training had an indirect effect on syllable blending. Note that, in kindergarten, syllable blending and RAN are significantly but scarcely correlated (relation with RAN speed: $\rho_s = 0.337$, $p = 0.005$; with RAN errors: $\rho_s = -0.237$, $p = 0.005$) (see also Norton & Wolf, 2012), whereas RAN correlation is moderately stronger with other metaphonological skills (initial-syllable recognition: $\rho_s = 0.477$; non-word minimal pairs: $\rho_s = 0.475$) which, conversely, did not show any significant effect in the RAN group.

In Study 2, we assumed RAN as a measure of automaticity, which intervenes to different extents in various tasks (see 4.4.2). In this vein, the improvement of RAN children in syllable blending might be the result of increased automaticity of metaphonological processes, stimulated by the RAN training. In this case, we wonder why it is evident only in syllable blending. One reason might lie in that syllable blending is the most accessible among measures of global metaphonological abilities. The average score obtained by children in

syllable blending ($M = 13.133$, $StdErr = 0.239$, $CI [12.659, 13.606]$)⁴⁹, in fact, is significantly higher than those obtained in every other global metaphonological task ($F(5, 485) = 22.809$, $p = 0.000$)⁵⁰. Another possibility might be conceptualizing RAN as one of the 3 subcomponents of phonological processing, along with phonological memory and phonological awareness (Wagner et al. 1999). This construct is based on the fact that rapid naming relies on the retrieval of phonological codes (Torgesen et al. 1997). Norton & Wolf (2012), following Wolf et al. (2000), provide meta-evidence against this view, based on (1) the scarce correlation found between measures of RAN and (meta)phonology (Swanson et al., 2003); (2) the fact that regression models consistently report RAN to account for unique variance in reading (Cutting & Denckla, 2001; Katzir et al., 2006); and, eventually, on (3) recent neuroimaging studies that found different biological bases for RAN and (meta)phonology (Turkeltaub et al. 2003). As for the first argument, we showed that, in kindergarten, the relation between RAN and metaphonological measures is moderate. We found a slightly stronger relation between RAN speed and phonological memory ($\rho_s = 0.558$, $p = 0.000$). Relations of (meta)phonology with RAN accuracy were systematically weaker. As for the second argument, in Study 2, where we treated predictors of literacy, we found the goodness of fit (Akaike Information Criterion) for our models by using RAN measures as a distinct block of fixed effects. We found that RAN contributes differently to different tasks, and to the same task analyzed in different moments (see 4.3 and 4.4)⁵¹. Based on these reflections, the model hypothesized in section 4.4.2 considers RAN as a latent variable of automaticity independently contributing to the automatization of writing processes. This might explain the optional presence of the RAN deficit in poor readers (for the Double Deficit Hypothesis, see Wolf & Bowers, 1999) and the fact that the co-presence of a RAN deficit

⁴⁹ The maximum raw score is 15.000.

⁵⁰ Mean difference : $0.898 < Est < 3.816$; $0.329 < SE < 4.54$; $0.000 < p < 0.007$; $CI [0.246, 4.619]$.

⁵¹ Wagner et al. (1997) found that the contribution of RAN to reading after controlling for phonological awareness changes with age and disappears after 3rd grade.

with a (meta)phonological deficit characterizes the most impaired readers (Wolf et al., 2002). In our view, the present data suggest considering RAN (of pictures, in this case) as a measure of automaticity of phonological processes. Further reflections may emerge from the analyses of the effects of interventions on RAN abilities (see below).

Analytical metaphonological skills

Analytical metaphonological skills were tested twice in 1st grade and once in 2nd grade. The latter intervention was administered between the two test sessions in 1st grade.

As for phoneme blending, a negative effect of administration 1-1 for every group except the GRAPH group revealed an initial advantage of GRAPH children, and a significant improvement of every group except the GRAPH between administration 1-1 and 1-2, namely after the 1st-grade training. The CG improved along with the META and the RAN group, meaning that the interventions did not determine significant effects. The GRAPH group seemed to be advantaged.

As for phoneme segmentation and deletion of the initial syllable, a negative effect of RAN in administration 1-1 and 1-2, and of CG in administration 1-1 revealed a general advantage of the META and the GRAPH group. The RAN group seemed to be particularly disadvantaged, as a significant improvement was evident only in 2nd grade. The CG, after an initial disadvantage, got an improvement in administration 1-2. It does not seem plausible that the better result of the META group was due to the metaphonological intervention, as the GRAPH group had similar outcomes. The negative effects of initial syllable deletion were very robust, thus revealing a particular difficulty of this ability for the CG and, in particular, for the RAN group.

As for deletion of the final syllable, a negative effect of 1st administration for every group except GRAPH, and of 1st and 2nd administration for the RAN group, revealed a general advantage of the GRAPH group, and a particular difficulty of the RAN group, which improved

significantly only in administration 2-1, whereas the META and the CG improved significantly in administration 1-2. The improvement of the CG along with the META group after post-intervention 2 and the difficulties shown by the RAN group revealed that neither the metaphonological nor the RAN intervention had significant effects on final syllable deletion. As for phonemic verbal fluency, the general negative effect of administration 1-1 revealed an initial difficulty by every group, showing no differences due to the training. The large improvement shown by the analysis of the *SMD* (see Appendix C) suggests that the META training helped children (in particular bilinguals) develop this ability, but this was not sufficient to determine a significant difference of META children from the other groups, as confirmed also by non-parametric tests.

Rhyme recognition was tested 5 times, twice in kindergarten (i.e., before and after Intervention 1), twice in primary school (i.e., before and after Intervention 2), and once in 2nd grade. A negative effect of first administration by every group, and of the first four administrations by the RAN group revealed a significant difficulty of the RAN group, which showed an improvement only in the 2nd grade, whereas the META and the GRAPH group showed a significant improvement in KG-2 and the CG in administration 1-2. Hence, the RAN training had certainly no effect. The META group improved significantly in KG-2, namely after the former training, as long as the GRAPH group, showing that also the metaphonological training had no effect.

These data, along with those of global metaphonological skills, show that a metaphonological intervention can be effective in the last year of kindergarten, namely during the early development of analytical metaphonological skills, and before the acquisition of literacy. Several studies focused on the importance of early intervention on phonological awareness, in particular for children with language disorders (for an extensive review, see Al Otaiba et al., 2009). The most fertile time for such an intervention has been

deemed to be the transition from kindergarten to primary school when metaphonological skills are in the process of developing from the global to the analytical level (Morais, 1987).

RAN

In the 5 administrations of picture RAN, significant negative effects were found in the speed of every group in the two kindergarten administrations, showing no immediate effect of the RAN intervention on speed. The large *SMDs* between administrations KG-1, KG-2, and 1-1 (see Appendix C) suggest strong improvement in every group and, in particular, in the RAN group, but this is not sufficient to determine significant differences in RAN speed, as confirmed also by non-parametric tests.

As for errors, the positive effects of the META and RAN group in KG-1, and the positive effect of the CG in KG-2 show that the GRAPH group had a starting advantage and that the META and the RAN group improved significantly after Intervention 1 (KG-2), whereas the CG improved later, namely in the middle of the 1st grade (1-1). The most powerful coefficient is that of the RAN group, thus suggesting an effect of the RAN intervention. In this case, the parallel improvement of the META group might be explained through a possible overlapping effect of the metaphonological intervention on speed in the RAN of pictures. If we consider RAN as a measure of automaticity of phonological processes (see the section *Global metaphonological skills*), it might be possible that training of phonological awareness could improve, namely automatize, RAN accuracy. In this regard, Khodadoust et al. (2017) found that a phonological-awareness training program (30 sessions of 40 minutes) significantly improved the RAN times of 62 Persian monolingual 1st graders. Layes et al. (2015) found similar results on 20 Arabic 4th and 5th graders (10 with developmental dyslexia) and argued that the slowness shown by children with dyslexia depends on a deficit in using phonological information for lexical access. These studies support the existence of shared processes

underlying naming speed and phonological awareness (Torgesen, et al., 1997; Wagner et al., 1997), thus diverging from the view of Norton & Wolf (2012).

Our data show a possible overlap between global metaphonological skills and picture RAN accuracy, whereas the mentioned studies deal essentially with RAN times⁵². The fact that this overlap is present only at the global level and only for the more accessible metaphonological ability (namely, syllable blending) (see 3.4.2) might support the possibility that we are dealing with an early automaticity form.

Digit RAN speed shows a general advantage of the META and GRAPH group, and an improvement of the CG along with the RAN group, thus indicating no differences due to the intervention.

Visual skills

Tests of visual memory and discrimination administered in kindergarten showed that the GRAPH group tended to make more errors in the 2nd administration, thus revealing no benefit from the graphic training.

Non-word repetition

Our data show a disadvantage of the RAN group in non-word repetition in the first two administrations and a general advantage of the META and GRAPH groups.

After Intervention 1, only the META group reported a significantly higher mean than the CG, thus showing that the metaphonological intervention might have empowered children's phonology. This highlights the early relation between language and metalanguage. With this regard, Chaney (1992) found interrelations between the major domains of language awareness (phonological, word, and structural), and between these domains and overall

⁵² The mentioned studies on Persian/Arabic use the measure of time, not of speed.

linguistic skills. She reports 3-year-old children as being already able to make metalinguistic judgments and productions in structured tasks, with their overall metalinguistic performance improving in months.

Syntactic skills and vocabulary

Syntactic comprehension was tested twice (in kindergarten and 2nd grade), whereas syntactic production of direct object clitic pronouns was tested three times with tests of increasing difficulty (see 2.3).

The META and the RAN group showed a head start, as they tended to make fewer errors than other groups in the first administration. In the second administration, however, only the META group had a significant improvement, thus showing that the metaphonological training might have affected children's syntactic competence. Note that global metaphonological skills correlate moderately to strongly with syntactic comprehension ($0.398 < \rho_s < 0.540$; $0.003 < p < 0,036$). Some studies relate phonological awareness to reading comprehension (see, e.g., Cárnio et al., 2017; Elhassan et al., 2017). To our knowledge, there is however scarce literature exploring the possible effect of phonological awareness on oral syntactic comprehension. It is plausible that tapping on word structures through metaphonological tasks refines children's familiarity and sensitivity towards morphemes and functional strings, which are often discriminant in oral syntactic comprehension⁵³.

As for syntactic production, the GRAPH group showed an advantage in the first and last administration, whereas the RAN group showed a negative effect in the second administration. This suggests that the META and the CG do not have morphosyntactic problems, whereas both the RAN and the GRAPH groups showed a fall in the second

⁵³ As said before, it has been shown that morphological awareness, and more in general (morpho)syntactic competence, contribute to decoding abilities (for Italian, see, for instance, Angelelli et al., 2014; Burani, 2010; Casani, in preparation).

administration, which was certainly due to the inclusion of the more demanding conditions of gender mismatch between subject and object in the elicitation test (see 3.3.1).

Receptive vocabulary did not show any effect of group, meaning that the training did not have effects on children's lexicon. This constitutes further support to the fact that RAN is not a vocabulary task (see, for instance, Wood et al., 2017) and that vocabulary requires large-scale intervention to be improved (see 1.5).

Digit and word span

The three administrations of span tests (once per grade) revealed that the three groups have a general lower digit span forward in kindergarten, and all improve in 1st grade.

The GRAPH group seems to be advantaged in digit span backward, as their span shows no negative effects in kindergarten, differently from that of their peers.

As for word span, the RAN group and the CG showed significant negative effects in KG and 1st grade, whereas the GRAPH group only in kindergarten, thus showing an improvement already in 1st grade. The META group seems to start advantaged, as it does not present the initial negative effect. As word span was tested at the end of kindergarten, namely in Post-intervention 1, we cannot exclude that it might be tapped by the metaphonological intervention, as it is confirmed by the large effect of the difference with the GRAPH group and the CG.

5.4.2. Effects of the training on literacy skills

Word and non-word decoding

In the middle of the 2nd grade, monolingual children who received the META and the GRAPH interventions, and bilinguals who received the META intervention showed significant effects of speed in decoding generic word lists. Monolinguals who received the META intervention

and monolinguals and bilinguals who received the GRAPH intervention also had significant effects on accuracy.

Monolingual META and GRAPH children were significantly faster than their CG in decoding generic word lists. Monolingual and bilingual children who received the META intervention were significantly faster than their respective controls in decoding non-words.

All monolinguals were significantly more accurate in decoding non-paroxytone words. The reason might lie in the reduced vocabulary of bilingual children compared to monolinguals (see 3.3.1 and 3.4.1). This means that diagnoses of reading deficits affecting the lexical route of reading can be particularly difficult in the presence of bilingual children because diagnoses might be jeopardized by children's language competence, at least until the 2nd grade. For what concerns the effects of the training, we can say that META intervention enhanced non-word decoding speed. It also seemed to help word decoding accuracy, together with the GRAPH intervention. The higher presence of non-word decoding errors in the META group compared to the GRAPH group was highlighted by the computation of the *SMD* (see Appendix C) but non-parametric tests confirmed its non-significance.

Sublexical decoding

To clarify the possible effect of the GRAPH intervention on sublexical decoding, we analyzed the longitudinal word decoding performances in the 1st and 2nd grades. These revealed that the monolingual and bilingual RAN and CG, in the middle of the 1st grade, are slower, whereas the monolingual and bilingual META groups and the monolingual GRAPH group are faster in the middle of the 2nd grade. We can state that the RAN training did not give any result on children's speed. The META intervention, instead, affected children's speed, as both monolinguals and bilinguals resulted to be significantly faster than their respective controls. Also, the monolingual GRAPH group was significantly faster than the respective

CG, and the effect was marginally significant ($p = 0.046$). So, the speed of GRAPH children is rather comparable to that of META children, and the difference from controls is marginal. The monolingual GRAPH group made fewer errors in both administrations, whereas the bilingual META group made fewer errors in the 1st grade. It seems that the double META training helps speed in both monolingual and bilingual children in the long term, whereas the former META training (in kindergarten) helps bilinguals' accuracy in the middle of the 1st grade. The double GRAPH training seems to help monolinguals' speed in the 2nd grade, and monolinguals' accuracy in both administrations. Bilinguals who attended the META training resulted to be significantly more accurate than their CG, as well as monolingual and bilingual children who attended the GRAPH training were significantly more accurate than their respective controls. The brilliant performances of GRAPH children might depend on either an effect of the graphic training on reading or, more probably, on their good verbal working memory (see their digit-span-backward and word-span performances) and good syntactic-production skills (see above).

Text decoding

The longitudinal analysis of text decoding (administrations 1-2; 2-1) revealed that bilingual children who attended the META training started with a disadvantage for what concerns speed, whereas monolinguals who attended the META and the GRAPH training significantly increased their speed in the latter administration. The post hoc analysis showed that monolingual children who attended the GRAPH training were faster in both administrations compared to their controls, whereas monolinguals who attended the META training were significantly faster in the latter administration. This means that for any reason (see *Sublexical decoding*), the monolingual GRAPH group starts with an advantage. As the first administration took place at the end of the 1st grade, namely in Post-intervention 2, we cannot exclude that the GRAPH training had an unexpected effect. The activities proposed

during training were drawing frames and different geometric small figures with increasing difficulty. These activities might have stimulated children's visuospatial skills and trained their executive functions.

The within-group analysis of text decoding accuracy showed a significant increase of accuracy between the 1st and 2nd grade in monolinguals who attended the META and RAN training, and in bilinguals who attended the RAN training.

RAN seems to affect the accuracy but not speed, and text decoding but not word decoding.

Reading comprehension

Reading comprehension was not indirectly affected by any intervention.

The analyses of the *SMDs* showed effects of higher scores by the META group compared to the RAN group in the first testing session following Intervention 1. Moreover, a large effect of improvement of reading central comprehension was revealed for the only META group (monolinguals + bilinguals). Although these effects are non-significant, they might prove an indirect stimulation of comprehension through metaphonological training. Considering also the good syntactic-comprehension skills of META children (see above), these effects might be interpreted in light of the Simple View of Reading (Gough & Tunmer, 1986), according to which reading is the product of written decoding and oral language comprehension (see 1.6.2).

Word / non-word writing

The longitudinal analysis of the 3 administrations of word dictation showed an initial advantage of the monolingual GRAPH group, which is the only one without an effect of error-increase in the first administration (1-1). In the second administration, this effect disappears for every group, thus showing a general improvement. In the second administration, namely post-intervention 2, bilinguals who attended the META training were significantly more

accurate than controls, thus showing that the metaphonological intervention worked on word writing in particular for bilingual children. This is probably related to an increase of familiarity with the properties of the Italian lexicon in bilinguals favored by the metaphonological activities, as also suggested by the fact that no effect on non-word writing emerged. A significant improvement of training groups took place between the middle of the 1st and the middle of the 2nd grade, namely in the Follow-up 2, whereas controls seem to have a more gradual improvement, which might depend on the higher homogeneity of their first performance (see Figure 38) and, perhaps, on their higher numerosity.

The smaller *SMD* between the first two dictations by the GRAPH groups compared to the other groups confirms an initial advantage of these groups, which might be the effect of Intervention 1 on the executive aspects of word writing,

As mentioned, non-word writing was not affected by the interventions.

Text writing

In 2nd grade, the analysis of dictation errors showed that monolingual children who attended the GRAPH training were significantly more accurate than their controls; the GRAPH group, along with the RAN group, in particular, made significantly fewer phonological errors than monolingual controls.

Bilinguals who attended the META and the GRAPH training were significantly more accurate and made significantly fewer phonological errors than bilingual controls. Hence, the META training showed to be particularly useful for bilingual children, where it affected, particularly, phonological errors.

Even in this case, the GRAPH group was more accurate than the other groups. Since this happens in both monolinguals and bilinguals, we guess that the training had some unexpected effect related to the practiced activities on visuospatial and executive abilities. Concerning writing, the copying activities based on picture series might have trained

graphism, thus facilitating writing. The presence of fewer phonological errors in dictation might be explained through a complex model of working memory integrating a componential and a capacity approach. In 4.4.4, we hypothesized that such a model might be responsible for minor errors (capital letters and punctuation graphic errors) in dictation, cause an overload of the central system determining a sort of inhibition of the visual in favor of the phonological slave system (see Baddeley, 1996; 2007). In this case, the graphic activities might have trained children's visuospatial skills, thus facilitating a degree of automatization of writing processes that could lighten the load of visual working memory required for the task. This might result in different degrees of accuracy according to the extent of automaticity achieved.

This "slave-system competition" hypothesis might be explored through a study aiming at analyzing the longitudinal development of writing through fine-grained error analysis.

5.5. Some methodological remarks

The present study incorporated many outcomes with a variety of different distributions. There were multiple predictors and multiple measures per subject for several outcomes. We needed a kind of analysis that could model clustered data (e.g., mono-/bi-linguals receiving different interventions) with specific predictors. At the same time, the variability in the number of participants, which is common in longitudinal designs, required a modeling technique that allowed us to deal with alternative correlation structures without degrading our datasets through listwise deletions for any instance of missing data.

Mixed Models for Repeated Measures (MMRM) enable flexible modeling of the correlation structure between time points to ensure the validity of inference in treatment efficacy. Standard MMRM, however, might lead to the inflation of type I error rates for the analyses

of intervention effects in the presence of small longitudinal datasets involving missing measurements (Ukyo et al., 2019),

Small sample adjustments for linear mixed models (Kenward & Roger, 1997; Lyons & Peters, 2000; Ukyo et al., 2019) and generalized linear models (Kauermann & Carroll, 2001) have been topics of interest. McCulloch (2003) highlights the need for similar work for Generalized Linear Mixed Models (GLMM), which are also complicated by non-normal distributed data and random effects (see, in this vein, Fraser et al., 1999).

Appropriate treatment of this topic is beyond our scope and competencies. However, considering the very small literature on this topic, the present study might give empirical support to the applicability of Generalized Mixed Models to a small-sample design, provided that data distributions and variance structures are appropriately analyzed.

We decided to apply Satterthwaite approximation to face the small sample size and unstructured variances. In some cases, we had to face the presence of outliers, which might bias the outcomes or hinder models' convergence. Their elimination from small or unbalanced samples should be carefully evaluated as it might constitute an additional risk to degrade our datasets. In some cases, the analysis of robust models might be of help for convergence.

Muth et al. (2016) report both Linear Mixed Models and, in particular, Generalized Estimating Equations as effective techniques as applied to repeated measures studies involving small samples ($N = 12$). We then decided to use Generalized Mixed Models (for Repeated Measures), which allowed us to control simultaneously for all the variables implicated. Nevertheless, considering the very small size of our subsamples and the complexity of our design, we ran separate analyses of the *SMDs* by using Hedge's effect size value corrected for small samples (between and within-subjects, respectively). The results which are reported in Appendix C, were coherent with those obtained through the Generalized Mixed Models, except for minor differences (e.g., in word writing).

The separate *SMD* analyses were a useful countercheck, but their employment in such a complex design did not prove to be convenient. To be correctly interpreted, in fact, measures of between- and within-subject effect size should be crossed with each other. Moreover, if we aim at attempting a generalization of data, their significance should be also considered (we used rank tests in our alternative analyses).

To conclude, the analyses of *SMDs* with appropriate small-sample correction can be useful as a countercheck of more complex models or with easier designs. In these cases, the outcomes can also be of easier interpretation than mixed-models effect sizes, as their measure is expressed in *SDs*.

With more complex designs involving different numbers of repeated measures, Generalized Mixed Models were confirmed to be appropriate even with very small samples, provided that the variables' features were carefully analyzed.

5.6. Interim Summary 3

We longitudinally analyzed the effects of three different interventions (a metaphonological, a RAN, and a graphic training conceived as a placebo), which were administered in two cycles (the former in the last year of kindergarten and the latter in the 1st grade) to three respective groups, compared with a CG that did not attend any intervention program.

We found effects on some tapped abilities and overlapping effects between the most accessible measure of metaphonology (syllable blending) and RAN. This should not lead to interpret RAN as a subcomponent of phonological processing but, to our view, as a measure of automaticity of phonological processes. The effects on phonological awareness suggest that a metaphonological intervention is recommendable before entering primary school, namely before the development of analytical metaphonological skills. The effects of

metalinguistic activities on phonology and syntactic comprehension stress the interactive relation existing between language and metalanguage.

The metaphonological training had effects on non-word decoding speed and word decoding accuracy, as well as on text writing and, only for bilingual children, on word writing. This is probably due to the increased familiarity with structural properties of the Italian lexicon stimulated by metaphonological activities. The graphic training, which was conceived as a placebo, had no direct effects on children's visual skills but had unexpected fallout on reading and writing, probably due to the effect of the stimulation of visual-spatial skills and executive aspects of writing through the graphic activities. A particular effect of the graphic training on phonological writing errors was interpreted in light of an integrated model of working memory, which is proposed in 4.4.4.

Generalized Mixed Models (for Repeated Measures) demonstrated to be appropriate analysis tools with small samples.

6. Study 4

Individual, social, and (meta)linguistic predictors of literacy difficulties

6.1. Aims 4

This study analyzed longitudinally (from the last year of kindergarten to the 3rd grade of primary school) the individual, linguistic, and metalinguistic profile of a small group of children who revealed different literacy problems in the 2nd and 3rd grade of primary school. The aim was to explore whether and to what extent some frequent comorbidities can constitute possible predictive factors for literacy problems. We refer to language disorders (Bishop et al., 2009; Bishop & Snowling, 2004; Catts et al., 2005; McArthur et al., 2000; Ramus et al., 2013), borderline intellectual functioning (Di Blasi et al., 2014; 2019), and ADHD (Miller et al., 2013), as well as to some conditions of familiarity (such as the presence of learning disorders in their family (Grigorenko, 2004). Moreover, the role of social deprivation (Burkam & Lee, 2002), parents' educational level (Bonifacci et al., 2013), or some preliteracy skills (such as language (Raman et al., 2020), metaphonology (Ball & Blachmann, 1988; Bradley, 1988; Frith, 1985; Hulme & Snowling, 1988; Lundeberg, Frost, & Peterson, 1988), and RAN (Decker et al., 2013; Wolf & Obregon, 1992; Ziegler et al., 2010)) were considered.

We eventually investigated whether a single (in the 1st grade) or a double (in the last year of kindergarten and in the 1st grade) metaphonological intervention can help prevent the development of literacy problems, and if any differences between monolingual and bilingual children emerge in this respect.

6.2. Method 4

6.2.1. Participants

A total number of 49 children participated in this study. This consisted of

- a core group of 30 typically developing children (17 mono + 13 bilinguals) who also took part in Studies 1- 3 and were monitored from the last year of kindergarten to the 2nd grade of primary school;
- a group of 19 children (6 monitored from the last year of kindergarten to the 2nd grade + 13 monitored from the 1st to the 3rd grade) who did not participate in the previous studies because of different kinds of reading and writing difficulties.

Description of the impaired sample

At the end of the project, 19 children (12 monolinguals + 7 bilinguals) in the 2nd (N = 6) and 3rd (N = 13) grade of primary school showed some reading or writing problems, as they performed 2 or more *SDs* under the norm. These children including all those who got diagnoses of developmental disorders during the project were excluded from the analyses shown in Studies 1 - 3.

Among those who entered the project in the last year of kindergarten and were monitored until the 2nd grade (2 mono- + 4 bilinguals), one bilingual was diagnosed with specific language impairment (LI), and one monolingual with LI associated with mild ID. Four of them (1 mono- + 3 bilinguals) lived some conditions of social deprivation, and 2 (1 mono- + 1 bilingual) had some relatives with Learning Disorders (LDs).

In the parents' questionnaires, their parents declared to have both attended the first school cycle⁵⁴ (N = 1); one the first cycle and the other one the secondary school (N = 1); both the secondary school (N = 2); and one the secondary school and the other one the University (N = 2).

Children who entered the project in the 1st grade (10 mono- + 3 bilinguals) belonged to a small group that was reported by teachers for problems in the acquisition of literacy. Among them, 4 monolinguals were diagnosed with LI (1 monolingual in association with mild ID, and 3 monolinguals in comorbidity with ADHD)⁵⁵. Among these children, 3 monolinguals had conditions of social deprivation and other 3 monolinguals had also some relatives with LDs. Among the others, 4 (2 monolinguals + 2 bilinguals) had some conditions of social deprivation. One of them (bi) had an additional mild ID and familiarity with LDs. One monolingual had a diagnosis of ADHD and Oppositional Defiant Disorder (ODD); and one monolingual had familiarity with LDs.

Their parents declared to have both attended the 1st educational cycle (N = 7); one the 1st cycle and the other one the secondary school (N = 1); both the secondary school (N = 3); one the secondary school and the other one the University (N = 1); and both the University (N = 1).

Table 36 summarizes the demographic features of the impaired sample.

Subject	Grade	Bilingualism	Diagnosis	Familiarity with LD	Social depr	Family education
I-1	3	mono				Secondary
I-2	2	mono		X	X	1st cycle
I-3	3	mono		X		1st cycle + secondary
I-4	3	mono				Secondary
I-5	2	bi			X	Secondary
I-6	2	bi	LI			Secondary
I-7	3	bi	ID	X	X	Secondary + university

⁵⁴ 5 years of primary school + 3 years of first secondary school.

⁵⁵ Diagnoses were released by the Italian public health system (ASL).

I-8	2	bi		X	X	Secondary + university
I-9	3	mono	ADHD + ODD		X	1st cycle
I-10	3	mono		X	X	1st cycle
I-11	2	bi			X	Secondary + university
I-12	3	bi			X	1st cycle
I-13	3	mono	LI + ADHD		X	1st cycle
I-14	3	mono	LI + ADHD	X	X	1st cycle
I-15	3	mono	LI + ADHD + IQ	X	X	1st cycle
I-16	3	mono	LI	X		1st cycle
I-17	2	mono	LI + ID			1st cycle + secondary
I-18	3	mono				Secondary
I-19	3	bi				University

Table 36. Demographic features of children with literacy problems

6.2.2. Materials

Materials were the same as in Study 1 (see 3.2.2). For detailed information about the tests, see sections 2.2 – 2.14.

6.2.3. Procedures

Testing

Testing took place over seven sessions. The 36 children monitored from kindergarten to the 2nd grade followed the 5-session procedure described in Study 1. They were tested in the middle (KG-1) and at the end (KG-2) of kindergarten, in the middle (1-1) and at the end (1- 2) of the 1st grade, and in the middle (2-1) of the 2nd grade of primary school.

The 13 children monitored from the 1st to the 3rd grade were tested 5 times starting from 1- 1, namely in sessions 1-1, 1-2, 2-1, and, in addition, at the end of the 2nd grade (2-2) and in the middle of the 3rd grade.

Interventions

A core group of 34 children was randomly divided into 4 groups receiving different interventions.

A group of 6 children (5 monolinguals + 1 bilingual) (META-1) who entered the project in the 1st grade, received one intervention of metaphonological training in the 1st grade of primary school while being monitored from the 1st to the 3rd grade.

All the other children received the double-cycle intervention described in Study 3, the former in kindergarten and the latter in the 1st grade.

A group of 9 children (6 mono + 3 bilinguals) (henceforth, META-2) received the double metaphonological training.

A group of 10 children (6 mono + 4 bilinguals) (henceforth, RAN) received the double RAN training.

A group of 9 children (7 mono + 2 bilinguals) received the double graphic intervention (henceforth, GRAPH), which was conceived as a placebo.

A group of 10 children (6 mono + 4 bilinguals) (CG), who did not receive any intervention, was used as Control Group (CG). Seven of them were monitored from the 1st to the 3rd grade, and 3 from kindergarten to the 2nd grade.

6.2.4. Analyses

In the last testing session of the project, 2nd- and 3rd-graders who were 2 or more SDs under the norm in some measures of reading and writing were included in a group considered at risk of developing some literacy impairments. Different literacy measures were assigned to the binomial category “Typical / Atypical”, and used as outcomes of a series of generalized mixed logit regressions. In Study 3, we referred to the small literature on this topic (see 5.5) to verify the applicability of Generalized (Linear) Mixed Models with small samples. The models converged. Furthermore, a countercheck by using between/within-subject *SMDs*

gave coherent results. We guess that the use of generalized mixed models might be reliable even with categorical outcomes.

We also categorized the relevant responses of parents' and teachers' questionnaires as binomial factors, as shown in Table 36 (mono/bilingual; familiar/non-familiar with developmental disorders; with/without social deprivation). The parents' educational level was multicategory. In case of different educational levels of parents, the mother's degree was considered.

The association between these categories and the presence/absence of a literacy difficulty was firstly explored through a Fisher's exact test. Children's individual and social features that resulted to be associated with the presence/absence of the difficulty were used as fixed effects of the mixed regressions. Bilingualism (mono/bi) and grade (2nd/3rd) were entered in the generalized model as random effects (see Barr et al., 2013).

After exploring the social and individual predictors of literacy problems, we analyzed the linguistic, metalinguistic, and memory predictors of the possible difficulties. We maintained the binomial outcomes for literacy skills (typical/impaired) as well as the random effects, and entered the results of different (meta)linguistic and memory tests as fixed effects. With repeated measures, we also combined the predictor with the administration time to explore possible differences between different test sessions.

We eventually analyzed the possible effects of different interventions by using the same models, with the same outcomes and random effects, but with the Group variable (META- 1, META-2, RAN, GRAPH, and CG) as fixed effect. Satterthwaite approximation was applied.

6.3. Results 4

Table 2 displays the literacy skills that resulted to be under the norm in each subject in the last test session.

Subject	Grade	Bilingualism	Dec. speed	Dec. acc.	Reading compr. tot.	Reading compr. centr.	Reading compr. spec.	Word writing	Non-word writing	Text writing
I-1	3	mono	+	-	+	+	-	-	+	-
I-2	2	mono	+	-	+	+	-	-	+	-
I-3	3	mono	+	-	-	-	-	-	+	-
I-4	3	mono	+	+	-	+	-	+	+	-
I-5	2	bi	-	-	+	+	+	-	-	-
I-6	2	bi	+	-	+	+	+	-	+	-
I-7	3	bi	-	-	-	-	-	-	-	-
I-8	2	bi	-	-	+	-	+	-	-	-
I-9	3	mono	-	-	-	+	-	-	-	-
I-10	3	mono	+	+	-	-	-	+	+	-
I-11	2	bi	+	-	+	+	+	-	+	-
I-12	3	bi	+	-	-	-	-	-	-	-
I-13	3	mono	-	-	+	-	+	-	-	-
I-14	3	mono	+	+	+	+	+	-	+	-
I-15	3	mono	+	+	-	-	-	-	-	-
I-16	3	mono	-	-	-	-	-	-	-	-
I-17	2	mono	+	-	+	+	+	-	-	-
I-18	3	mono	+	-	-	-	-	-	+	-
I-19	3	bi	+	+	+	+	+	+	+	-

Table 1. Abilities that resulted to be 2 or more SDs under the norm (-) in each subject in the last testing session (2nd and 3rd grade)

Among children who entered the project in kindergarten and were monitored until the 2nd grade, 2 were slower, and all of them (N = 6) were less accurate than the norm in text decoding in administration 2-1. None of them had general reading comprehension problems, although one bilingual presented problems in the comprehension of central questions, and one monolingual problems in the comprehension of specific questions. All of them were not accurate in word and text writing, and three of them (2 bilinguals and 1 monolingual) in non-word writing.

Among children who entered the project in the 1st grade and were monitored until the 3rd grade, 4 (3 mono + 1 bilingual) were slow in text decoding speed in the middle of the 3rd grade, and 8 (including those with speed problems) were not accurate. Nine children (7 mono + 2 bilinguals), 6 of whom with decoding accuracy under the norm, were poor in reading comprehension. Ten children (8 mono + 2 bilinguals) were poor in word writing, and 6 (4 mono + 2 bilinguals) in non-word writing. Every child (10 mono + 3 bilinguals) was poor in text writing. Everyone being poor in word or non-word writing was also poor in text writing but the opposite was not true, as 3 children (2 mono + 1 bilingual) showed no problems in word writing and 7 (6 mono + 1 bilingual) in non-word writing. Their schoolmates who did not deviate 2 or more *SDs* from the norm were assumed as typically developing controls (N = 30).

6.3.1. Individual and social predictors of literacy difficulties

A Fisher's exact test did not reveal any significant association between any single child and a particular impaired ability or individual characteristic, so everyone was included in the analysis.

A series of Fisher's exact tests were performed to explore the relations between each poor ability and individual characteristics, namely bilingualism, LI, ID, ADHD, social deprivation, familiarity with LD, and parents' educational level, based on responses to parents' surveys. There were significant associations

- of ADHD with poor decoding speed, reading global comprehension, word writing, and non-word writing;
- of social deprivation with poor decoding speed, decoding accuracy, reading central and specific comprehension, word writing, non-word writing, and text writing;
- of LD-familiarity with the poor decoding speed and decoding accuracy, reading global, central, and specific comprehension, word writing, non-word writing, and text writing;

- of parents' educational level with poor decoding accuracy, reading comprehension, and word writing;
- of bilingualism with poor reading specific comprehension;
- of LI and ID with poor word writing and non-word writing.

Fisher's and post hoc statistics are in Appendix D.

A series of logit Generalized Linear Mixed Models (GLMM) were analyzed with presence/absence of different difficulties as outcomes, individual characteristics correlating with each difficulty as fixed effects, and grade (2nd / 3rd) as a random effect.

The models predict poor decoding speed ($F(1, 48) = 6.343, p = 0.033, \text{acc.} = 88\%$), decoding accuracy ($F(2, 47) = 3.998, p = 0.025, \text{acc.} = 78\%$), word writing ($F(2, 46) = 6.185, p = 0.004, \text{acc.} = 83.7\%$), and non-word writing ($F(1, 47) = 11.053, p = 0.002, \text{acc.} = 85.7\%$). A main effect of social deprivation emerged on decoding speed, decoding accuracy, and non-word writing, and a main effect of social deprivation + LI on word writing, whereas no effect on reading comprehension and text writing emerged.

Figure 40 shows the estimated means of individual characteristics predicting different difficulties. Text decoding accuracy estimated means ($F(1, 47) = 7.226, p = 0.010; \text{Est} = 0.431, SE = 0.160, t(47) = 2.688, p = 0.010, CI [0.109, 0.754]$) as well as non-word writing means ($F(1, 4) = 11.368, p = 0.024; \text{Est} = 0.548, SE = 0.161, t(4) = 3.411, p = 0.024, CI [0.114, 0.982]$) are significantly higher in children without social deprivation compared to those with social deprivation.

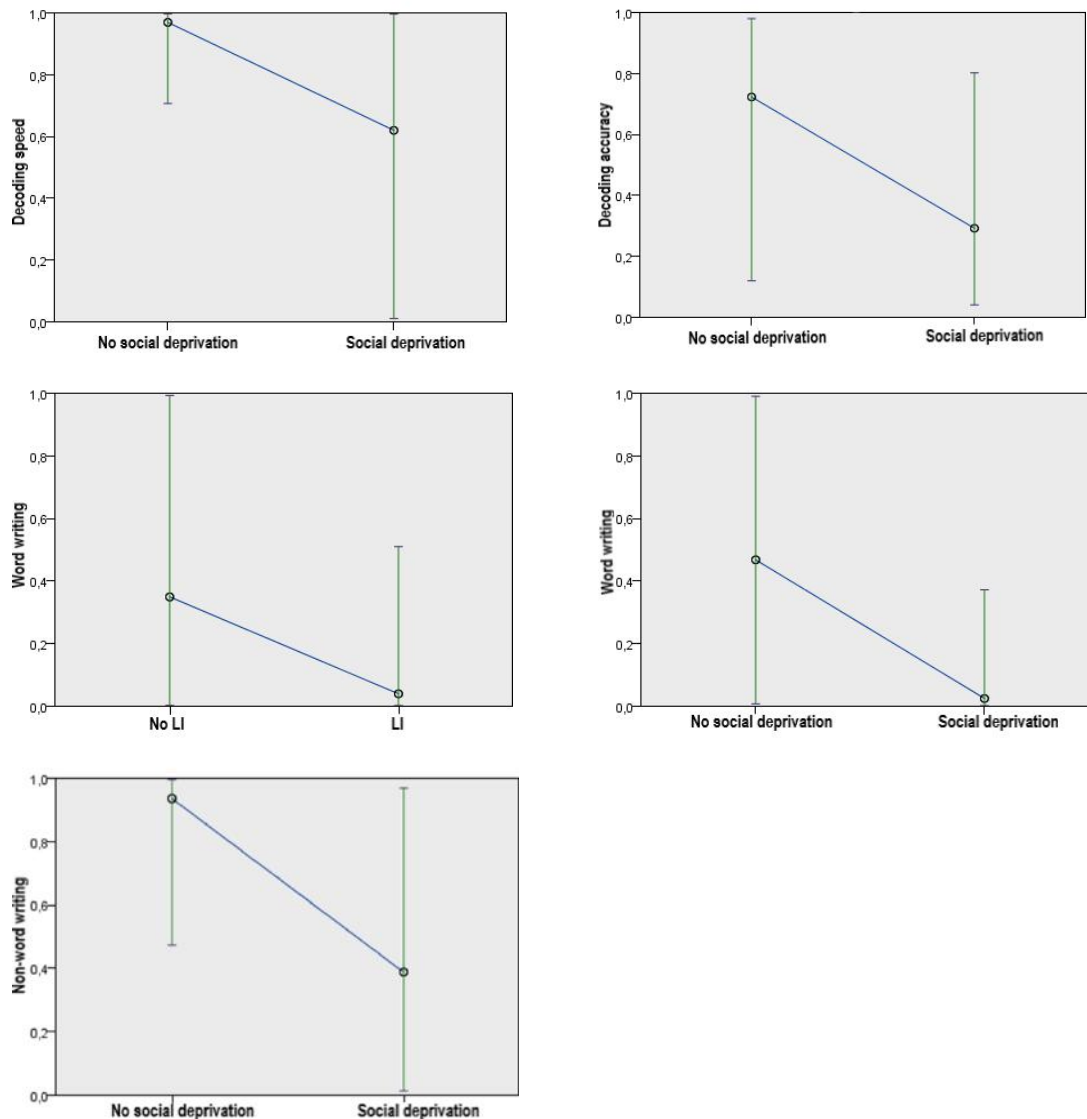


Figure 40. Contrasts between the estimated means of significant individual predictors of poor literacy skills.

Social deprivation predicted

- 100% of cases without impaired decoding speed ;
- 88.2% of cases without impaired decoding accuracy, and 56.2% of cases with that difficulty;
- 87.2% of cases without impaired non-word writing, and 80% with that difficulty;

Social deprivation and LI predicted 92.9% of cases without impaired word writing and 71.4% with that difficulty.

Figure 41 shows significant predictive effects. Complete statistics are in Appendix D.

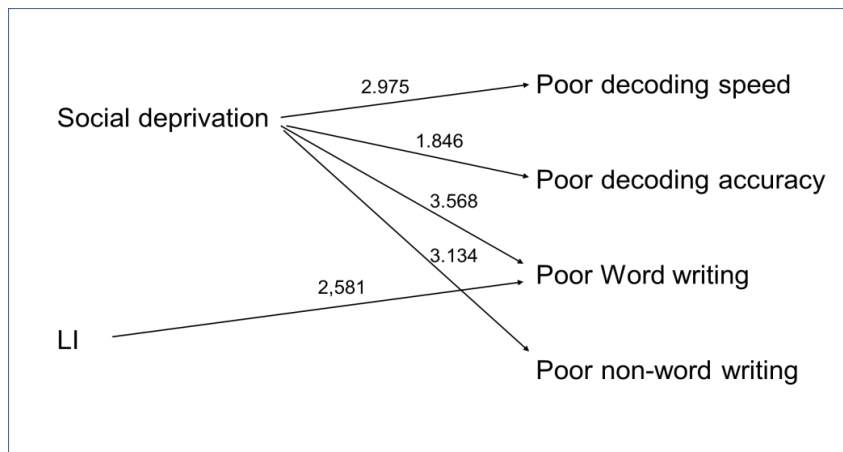


Figure 41. Individual predictors of different literacy difficulties in 2nd and 3rd grade. Coefficients are on the respective arrows. Complete statistics are in Appendix D (Table D.3).

No effect of school grade on any difficulty emerged ($0.536 < p < 1.000$).

6.3.2. Memory, linguistic, and metalinguistic predictors of literacy difficulties

Short-term memory, phonological, and visual skills

A series of logit GLMM for repeated measures (GLMMRM) were analyzed with the presence/absence of different difficulties in 2nd and 3rd grade as dependent variables, scores obtained in the 4 longitudinal administrations of digit and word span, phonological and visual short-term memory as two distinct blocks of fixed effects, and grade and bilingualism as random effects. A backward stepwise procedure was adopted to isolate significant variables of each block and obtain fitting sensitivity and specificity values for each predictive variable. Poor decoding speed was predicted by digit span forward and backward ($F(2, 131) = 10.943, p = 0.000$). Poor decoding accuracy was predicted by digit span forward and word span ($F(2, 109) = 9.192, p = 0.000$), as well as by the visual discrimination of capital letters ($F(1, 102) = 8.436, p = 0.005$).

Poor reading comprehension was predicted by digit span backward (reading compr. total score: $F(1, 132) = 6.139, p = 0.014$; central score: $F(1, 132) = 12.374, p = 0.001$; specific score: $F(1, 111) = 8.137, p = 0.005$). Moreover, poor global ($F(1, 102) = 5.826, p = 0.018$) and central ($F(1, 102) = 4.900, p = 0.029$) comprehension were predicted by the graphic reproduction of two semicircles, and poor central comprehension by non-word repetition ($F(1, 112) = 26.962, p = 0.000$).

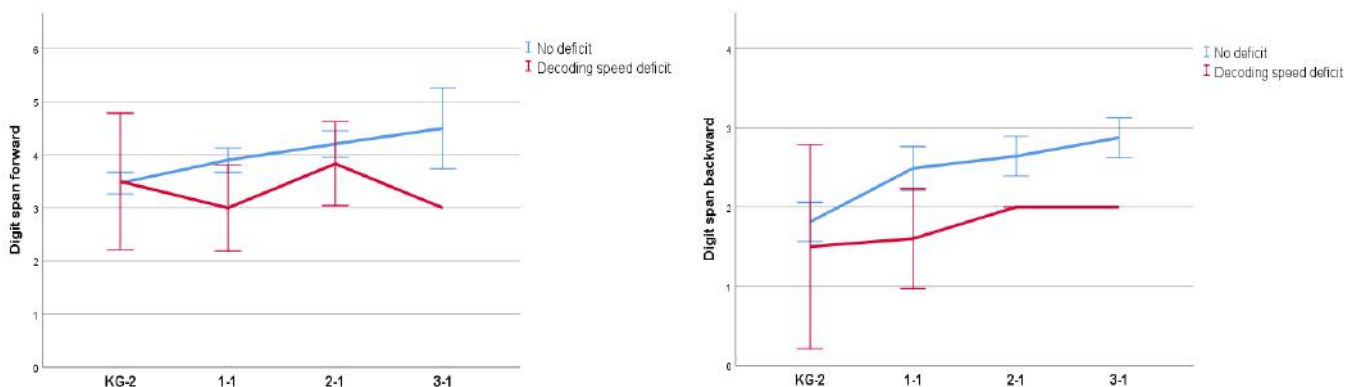
Poor word writing was predicted by digit span forward and word span ($F(2, 47) = 12.385, p = 0.000$), non-word repetition ($F(1, 112) = 5.965, p = 0.016$), and graphic reproduction of two semicircles ($F(1, 86) = 9.232, p = 0.003$).

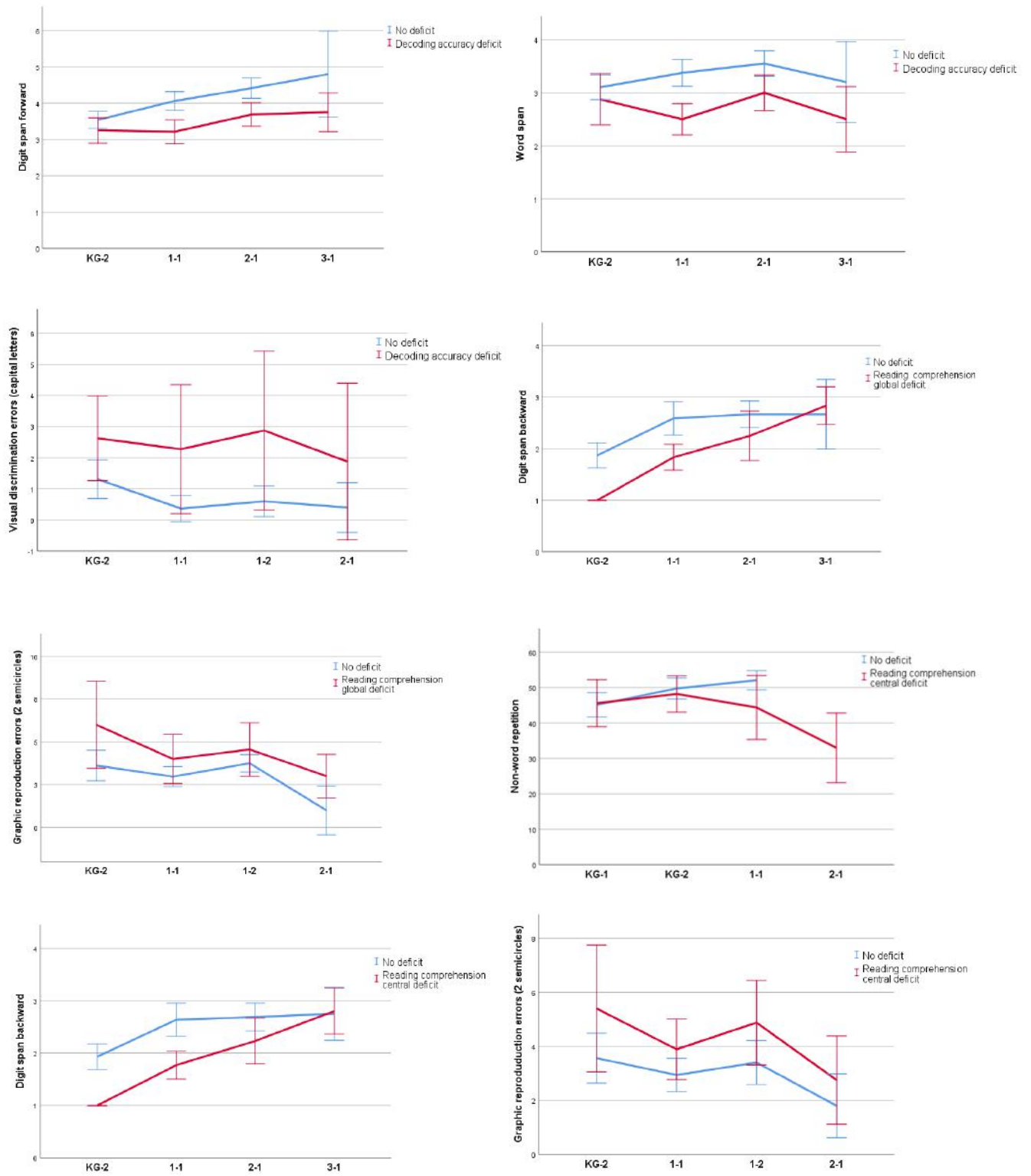
Poor non-word writing was predicted by digit span forward ($F(1, 97) = 11.354, p = 0.001$) graphic reproduction of one semicircle, and visual discrimination of capital letters ($F(2, 101) = 6.722, p = 0.002$).

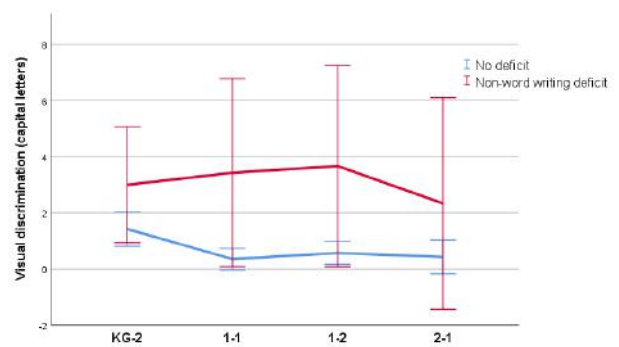
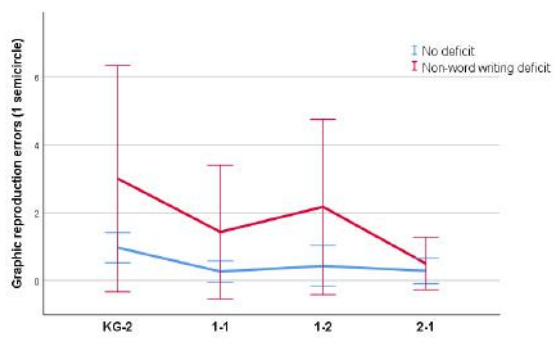
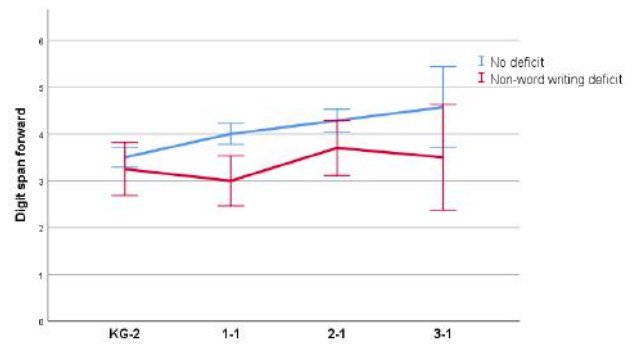
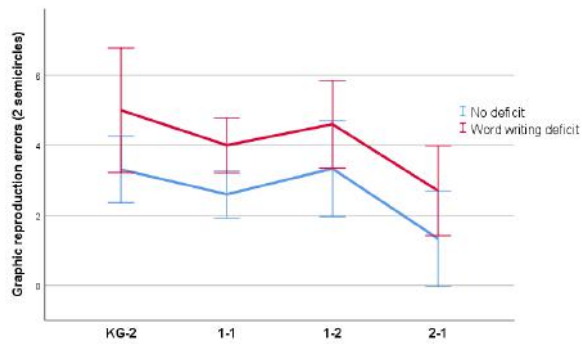
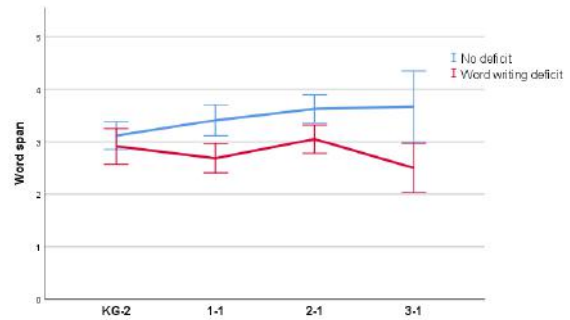
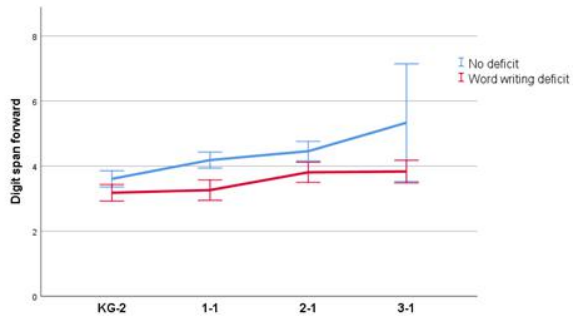
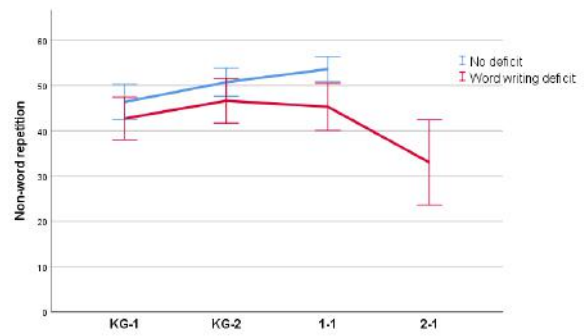
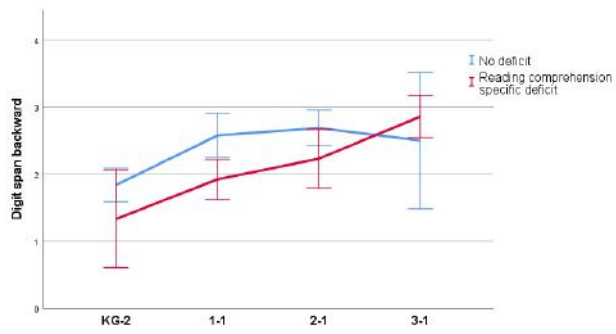
Poor text writing was predicted by digit span backward ($F(1, 132) = 10.028, p = 0.002$).

Global and partial predictive accuracy are in the Appendix (Table D.4).

Figure 42 shows the mean scores obtained by children in the predictive abilities for each test session.







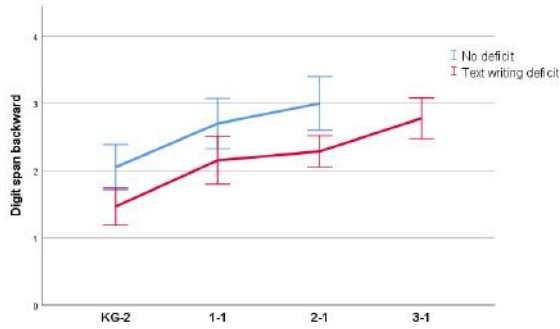


Figure 42. Longitudinal scores obtained by children with/without difficulty (lines) in each session (x-axis) testing predictive abilities (y-axis).

The predictive effect of digit span backward on decoding speed is significant in every administration (acc. = 87.3%, $F(4, 129) = 5.216$, $p = 0.001$), whereas the combined predictive effect of digit span forward and administration time on decoding speed is not significant ($p = 0.177$).

The predictive effect of digit span forward (acc. = 75.4%, $F(4, 129) = 4.007$, $p = 0.004$) and word span (acc. = 76.6%, $F(4, 132) = 4.567$, $p = 0.002$) on decoding accuracy is significant in every administration, whereas the combined effect of letter visual discrimination and administration time has no significant effect on decoding accuracy ($p = 0.134$).

The predictive effect of digit span backward on reading comprehension (acc. = 85.8%, $F(4, 55) = 2.608$, $p = 0.045$), as well as that on central comprehension (acc. = 80.6%, $F(4, 129) = 5.050$, $p = 0.001$), is significant in every administration (though the significance in 3-1 is marginal), whereas the effect on specific comprehension is significant in the first 3 administrations but not in 3rd grade (acc. = 88.1%, $F(4, 59) = 2.527$, $p = 0.050$). The combined effect of graphic reproduction of 2 semicircles and administration time on reading comprehension is not significant (global comprehension $p = 0.192$; central comprehension $p = 0.279$).

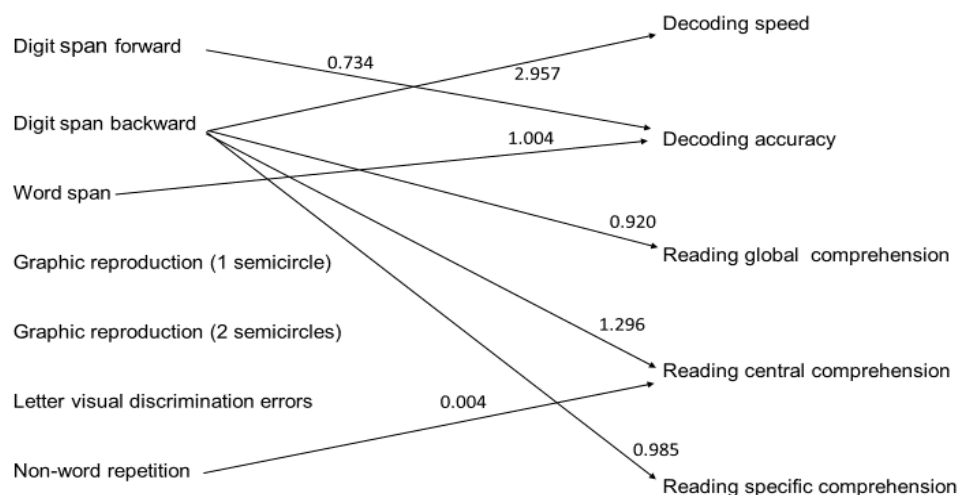
The predictive effect of digit span forward, as well as that of word span on the word writing difficulty, is significant in every administration, whereas the combined effect of non-word repetition and administration time on word writing is not significant.

The predictive effect of graphic reproduction of 2 semicircles on word writing is marginally significant in the first two administrations (KG-2 and 1-1), whereas there is no significant combined effect of this ability and administration time from administration 1-2 onward.

The predictive effect of digit span forward on non-word writing is significant in every administration, whereas that of visual discrimination of capital letters is marginally significant in the first two administrations (KG-2 and 1-1) but not from 1-2 onwards.

The predictive effect of digit span backward on text writing is significant in the first 3 administrations but not in 3rd grade.

Figure 43 summarizes the significant memory predictors of each literacy skill.



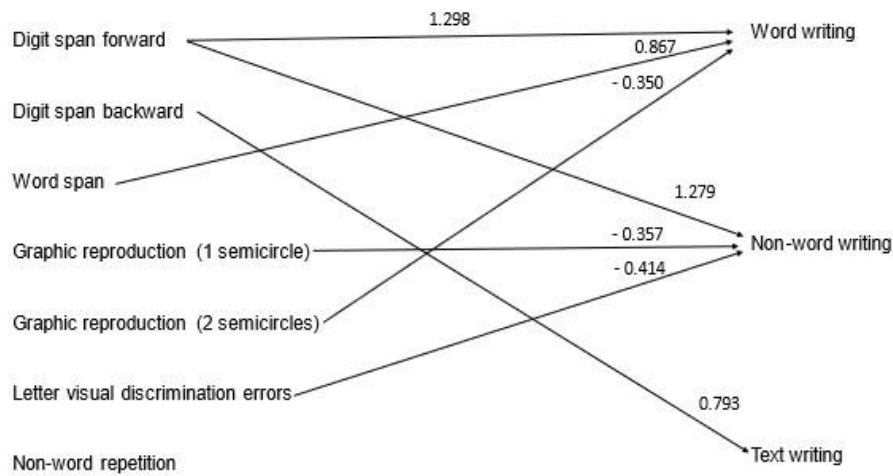


Figure 43. Memory and visual longitudinal predictors of different literacy difficulties in 2nd and 3rd grade. Coefficients are on the respective arrows. Complete statistics are in Appendix D (Table D.5).

Statistics are in Appendix D (Table D.6).

As for random effects, the following very weak but significant effects emerged:

- a negative effect of bilingualism on decoding accuracy after controlling for the main effect of visual discrimination of capital letters;
- a negative effect of bilingualism and a positive effect of monolingualism on reading global comprehension after controlling for the main effect of graphic reproduction of two semicircles;
- a negative effect of bilingualism and a positive effect of monolingualism on text writing only in 2nd grade after controlling for the main effect of digit span backward.

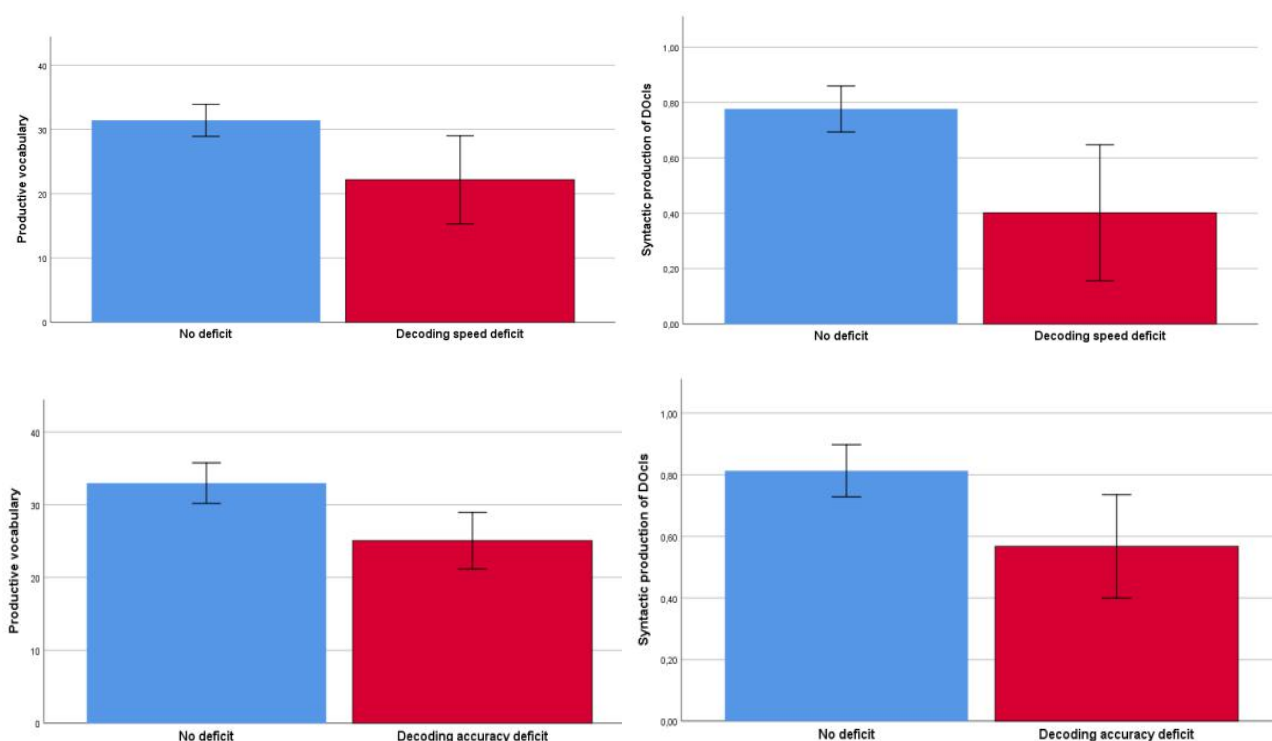
Statistics are in Appendix D (Table D.7)

Language skills

A series of logit GLMM were analyzed with the presence/absence of each difficulty as a dependent variable; receptive vocabulary, productive vocabulary, syntactic comprehension, and syntactic production scores obtained by 2nd and 3rd graders as fixed effects, and bilingualism (mono / bi) and grade (2nd / 3rd) as random effects.

A backward stepwise procedure was adopted to isolate fitting predictors, as above. Productive vocabulary resulted to be a significant predictor of every difficulty except that of global reading comprehension. Moreover, syntactic production of DOcls predicted the poor decoding speed and decoding accuracy, whereas syntactic comprehension predicted poor central reading comprehension, word writing, and non-word writing.

Figure 44 shows the scores obtained by children with/without difficulty in the abilities that are predictive of the respective difficulties.



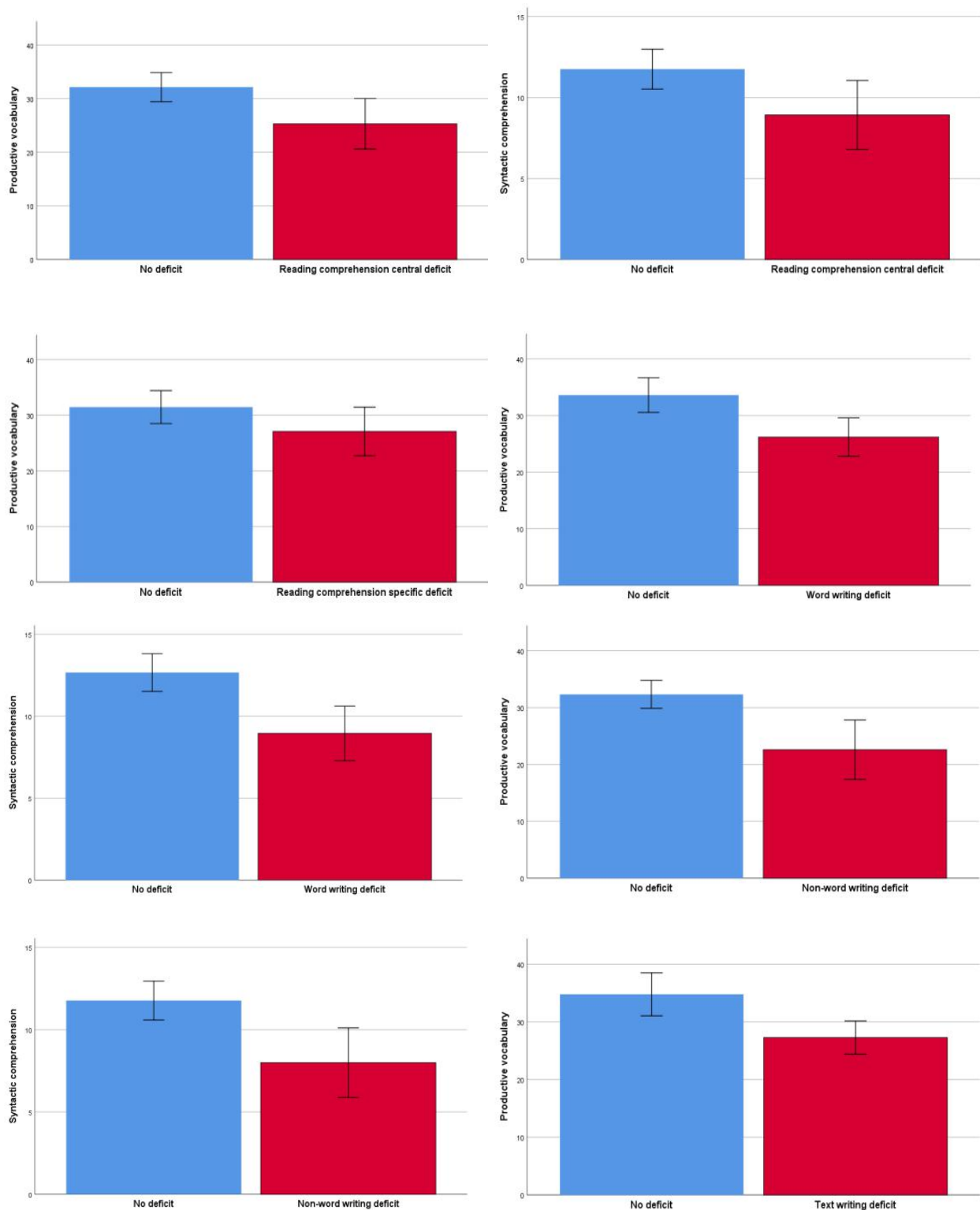


Figure 44. Scores obtained by children with/without difficulty (x-axis) in the abilities that predict the respective difficulties (y-axis) in 2nd and 3rd grade of primary school.

Figure 45 summarizes the linguistic predictors of reading and writing difficulties. The accuracy of predictive models is in Appendix D (Table D.8). Statistics of effects are in Table D.9.

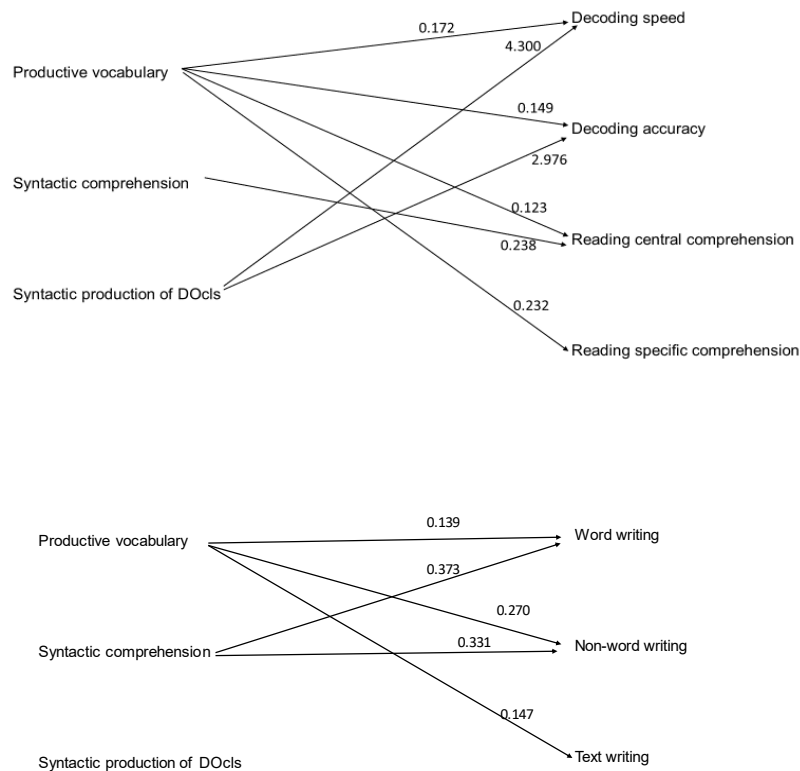


Figure 45. Linguistic predictors of different literacy difficulties in 2nd and 3rd grade. Coefficients are on the respective arrows. Complete statistics are in Appendix D (Table D.9).

There was no effect of grade ($0.383 < p < 1.000$) or bilingualism ($0.486 < p < 1.000$) on any difficulty.

Global metaphonological skills

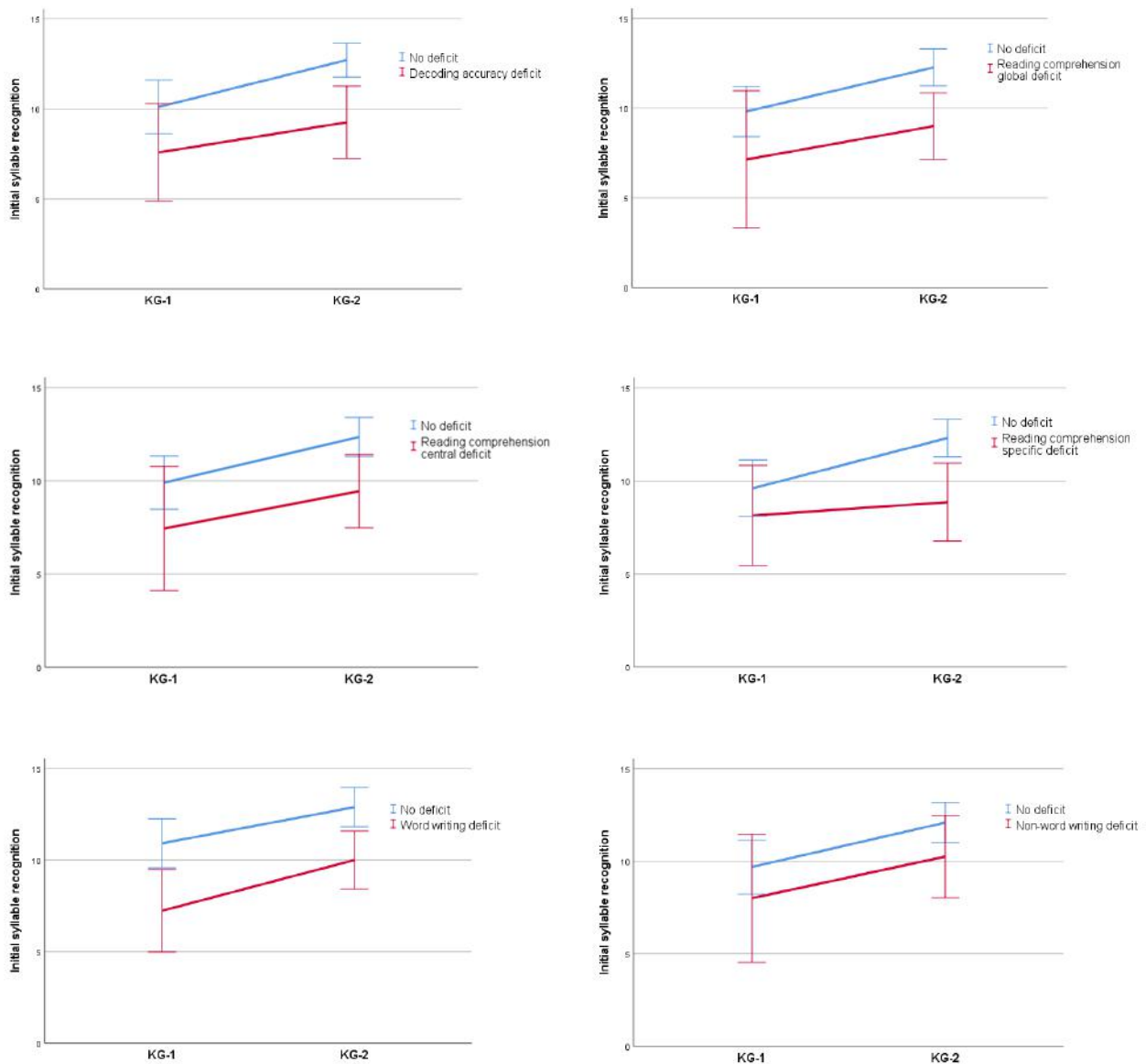
A series of logit GMMRM were analyzed with the presence/absence of the difficulty as a dependent variable, longitudinal scores obtained in the two KG administrations of global

metaphonological tests as fixed effects, and bilingualism and grade as random effects. The models significantly predict decoding accuracy, reading comprehension, word writing, non-word writing, and text writing.

A main effect of recognition of the initial syllable on every difficulty emerged.

The predictive accuracy of the model (Table D.10) and complete statistics of effects (Table D.11) are in Appendix D.

Figure 46 shows the scores obtained by children with/without difficulty in the recognition of the initial syllable during the two testing sessions in KG.



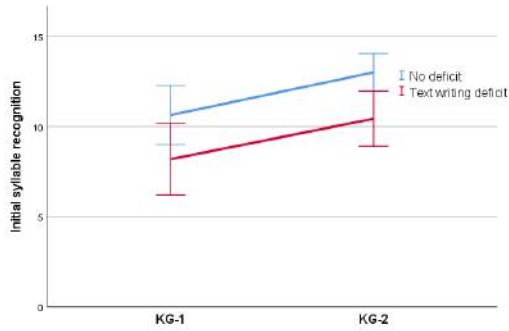


Figure 46. Longitudinal mean scores (y-axis) obtained by children with/without difficulty (lines) in the recognition of the initial syllable during the two KG testing sessions (x-axis).

Figure 47 shows the longitudinal predictive strength of initial syllable recognition on different literacy difficulties.

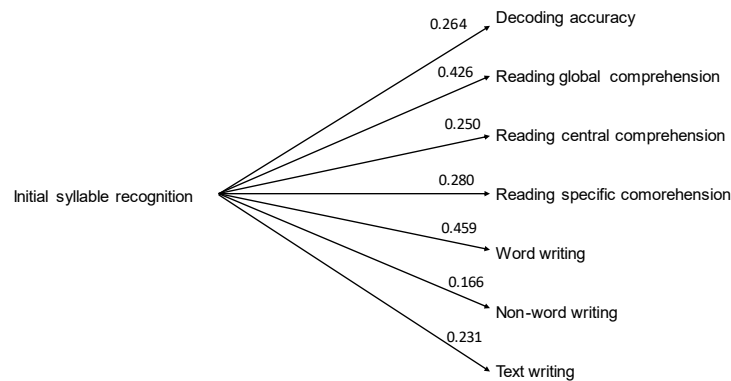


Figure 47. Predictive longitudinal effects of initial syllable recognition in kindergarten on different literacy difficulties. Strength coefficients are on the respective arrows. Complete statistics are in Appendix D (Table D.11).

As for random effects, the following very weak but significant effects emerged after controlling for the main effect of recognition of the initial syllable:

- a significant weak negative effect of bilingualism and positive effect of monolingualism on the reading global comprehension difficulty; a weak positive effect on central comprehension

was significant only in monolingual 2nd graders, and a weak negative effect only in bilingual 3rd graders, whereas no effect of bilingualism or grade on specific comprehension emerged; -a positive effect of monolingualism and negative effect of bilingualism on the word writing difficulty also emerged.

Significant random effects are in Appendix D (Table D.12).

Analytical metaphonological skills

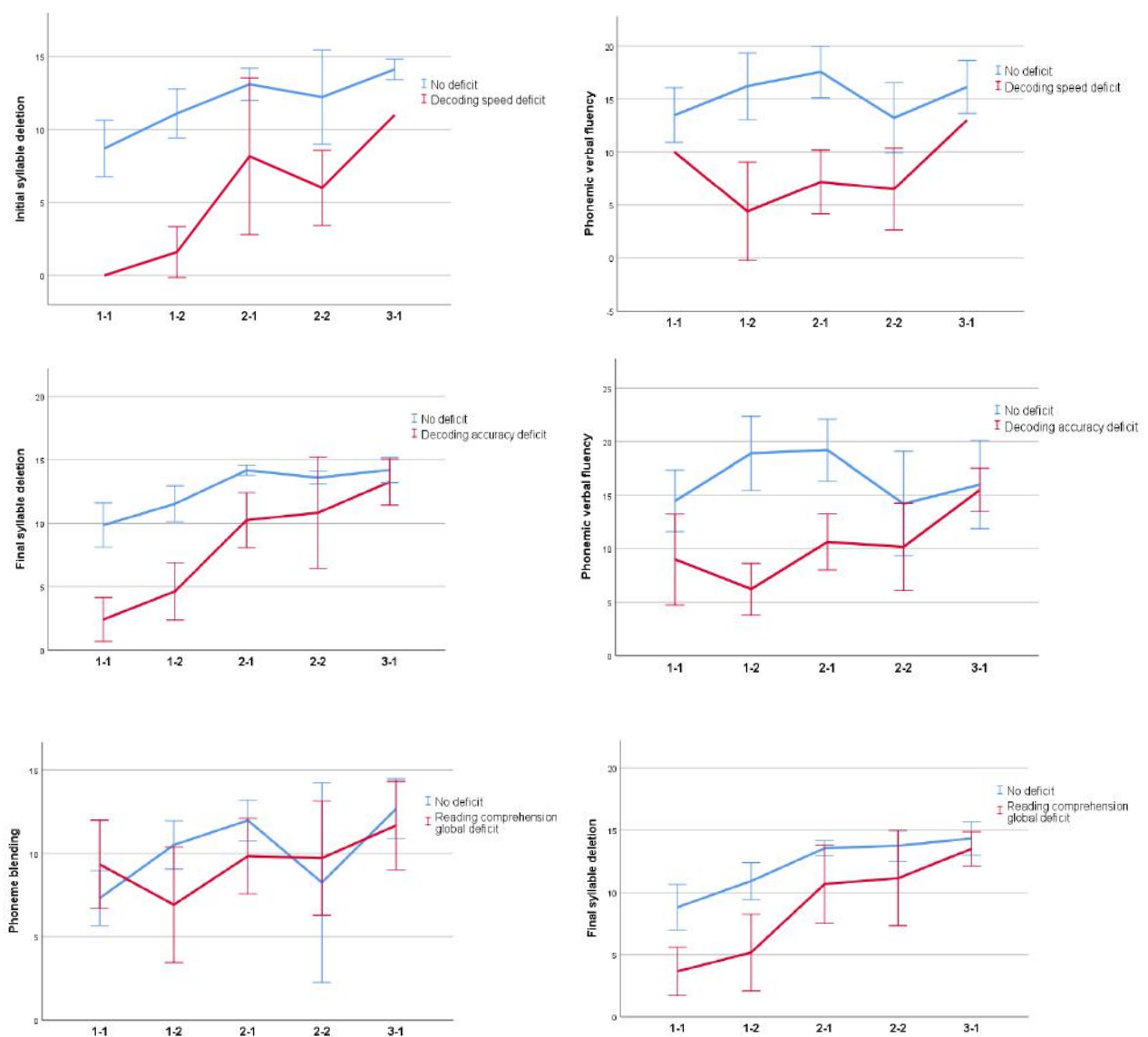
A series of logit GMMRM were analyzed with the presence/absence of each difficulty as a dependent variable, longitudinal scores obtained in the 5 five sessions testing analytical metaphonological skills as fixed effects, and bilingualism as a random effect. A stepwise backward procedure was adopted, as above.

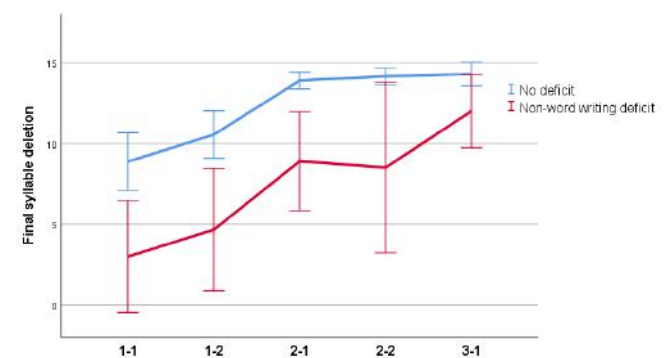
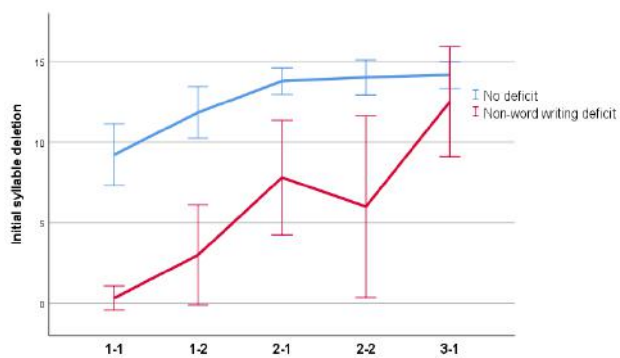
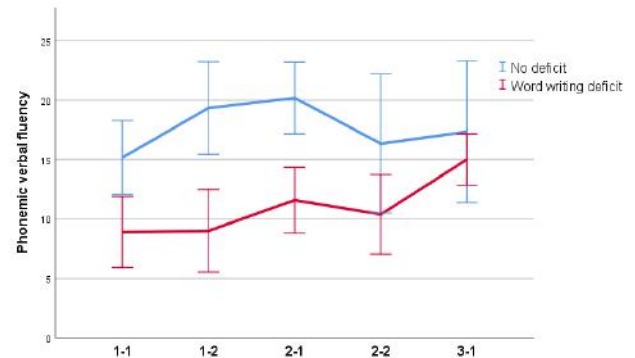
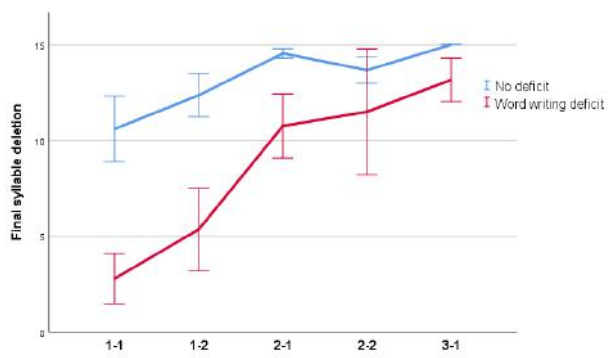
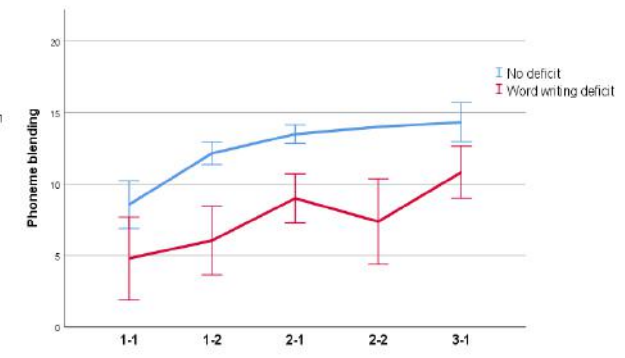
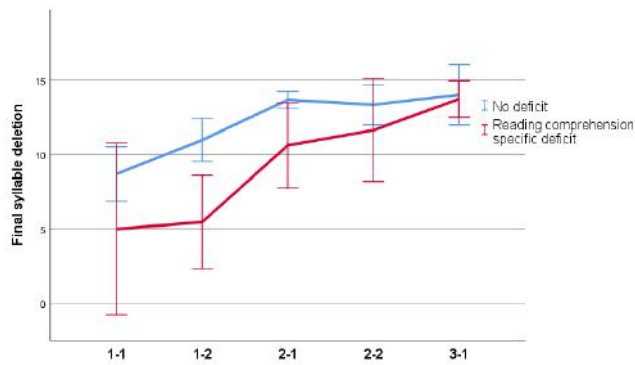
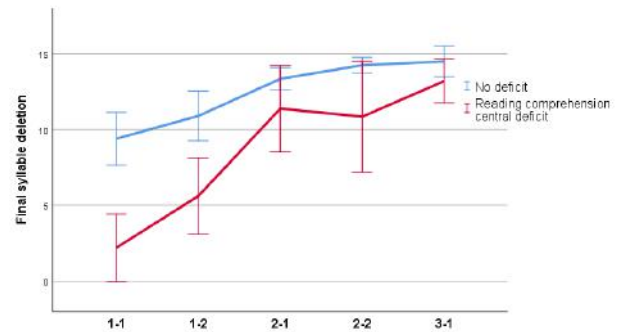
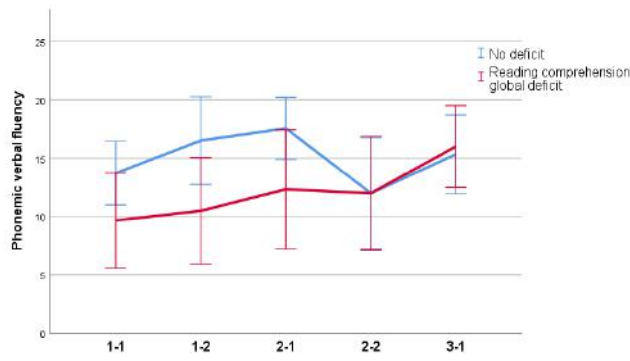
The models predict:

- the decoding speed difficulty ($F(2, 62) = 9.932, p = 0.000$), with a main effect of initial-syllable deletion ($F(1, 56) = 8.077, p = 0.006$) and phonemic verbal fluency ($F(1, 64) = 6.746, p = 0.012$);
- the decoding accuracy difficulty ($F(2, 87) = 20.734, p = 0.000$), with a main effect of final syllable deletion ($F(1, 93) = 16.272, p = 0.000$) and phonemic verbal fluency ($F(1, 80) = 19.595, p = 0.000$); the reading comprehension difficulty (global comprehension $F(3, 32) = 4.751, p = 0.008$), with a main effect of phoneme blending ($F(1, 32) = 12.846; p = 0.001$), final syllable deletion ($F(1, 32) = 9.828, p = 0.004$), and phonemic verbal fluency ($F(1, 32) = 4.199, p = 0.049$) on global comprehension, and of final syllable deletion on central ($F(1, 80) = 15.632, p = 0.000$) and specific ($F(1, 79) = 13.349, p = 0.000$) comprehension;
- the word writing difficulty ($F(3, 70) = 11.763, p = 0.000$), with main effects of phoneme blending ($F(1, 57) = 21.942, p = 0.000$), final syllable deletion ($F(1, 76) = 8.198, p = 0.005$), and phonemic verbal fluency ($F(1, 85) = 8.168, p = 0.005$);

- the non-word writing difficulty ($F(3, 61) = 8.492, p = 0.000$), with main effects of initial-syllable deletion ($F(1, 73) = 8.736, p = 0.004$), final-syllable deletion ($F(1, 72) = 4.116, p = 0.046$), and phonemic verbal fluency ($F(1, 63) = 13.328, p = 0.001$);
 - the text writing difficulty, with a main effect of phoneme blending ($F(1, 47) = 18.625, p = 0.000$).
- The predictive accuracy of each model is in Appendix D (Table D.13).

Figure 48 displays the mean scores obtained by children with and without difficulties in the predictive metaphonological abilities during the five test sessions in 1st, 2nd, and 3rd grade.





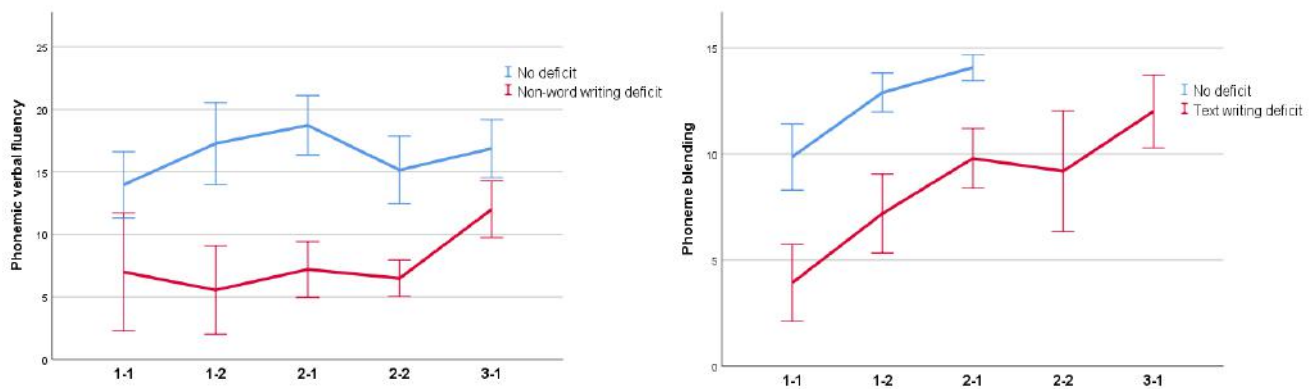


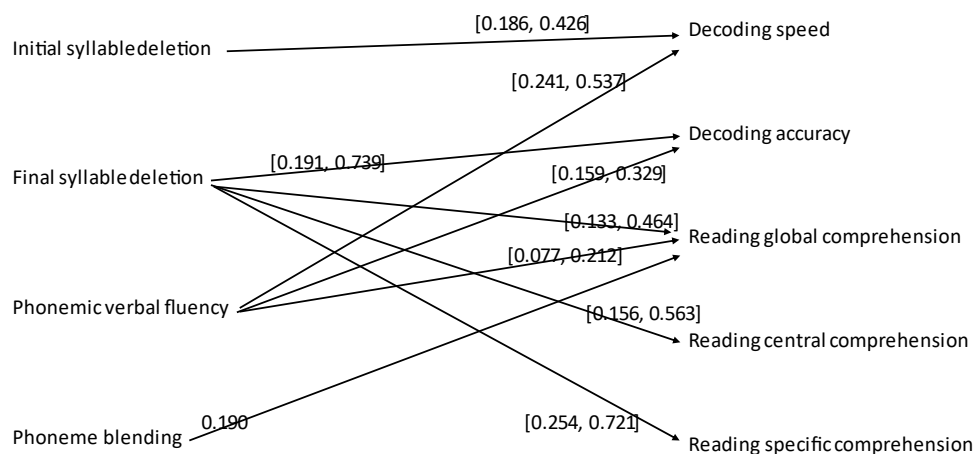
Figure 48. Longitudinal mean scores obtained by children with/without difficulties in the respective predictive metaphonological abilities during the five test sessions in 1st, 2nd, and 3rd grade.

An analysis of the association between each predictor and test-administration time as a fixed effect, and bilingualism as a random effect showed that:

- deletion of the initial syllable predicts the decoding speed difficulty in the first 4 administrations ($F(4, 31) = 5.638, p = 0.002$), with marginally significant effects in the fourth administration, whereas phonemic verbal fluency predicts the decoding speed difficulty in every administration ($F(5, 138) = 4.893, p = 0.000$);
- both final syllable deletion ($F(5, 32) = 5.682, p = 0.001$) and phonemic verbal fluency ($F(5, 40) = 7.089, p = 0.000$) predict the decoding accuracy difficulty in every administration;
- phoneme blending predicts the reading global comprehension difficulty only in the first administration, whereas its combined effect with every other administration is not significant ($F(5, 138) = 2.932, p = 0.015$);
- final syllable deletion ($F(5, 38) = 4.647, p = 0.002$) and phonemic verbal fluency ($F(5, 38) = 4.647, p = 0.002$) predict the global comprehension difficulty in the first three administrations, whereas their combined effect with administration 2-2 and 3-1 is not significant; moreover, final syllable deletion predicts the central ($F(5, 40) = 4.759, p = 0.002$) and the specific ($F(5, 41) = 5.832, p = 0.000$) comprehension difficulty in the first three administrations;

- phoneme blending predicts the word writing difficulty in every administration except 2-2 (which was excluded to make the model converge), even though the combined effect with administration 1-1 is marginally significant ($F(4, 40) = 6.260, p = 0.001$); deletion of the final syllable predicts the word writing difficulty in every administration except 2-1 (which was also excluded to make the model converge) ($F(4, 25) = 6.224, p = 0.001$); phonemic verbal fluency predicts the word writing difficulty in every administration ($F(5, 35) = 5.225, p = 0.001$).
- initial syllable deletion ($F(5, 32) = 6.431, p = 0.000$) and final syllable deletion ($F(5, 34) = 4.193, p = 0.004$) predict non-word writing in every administration,
- phonemic verbal fluency predicts non-word writing in the first 3 administrations ($F(3, 94) = 10.022, p = 0.000$), whereas the last two administrations were excluded to make the model converge;
- phoneme blending predicts the text writing difficulty in the first 3 administrations, whereas the combined effect with the last two is not significant ($F(5, 112) = 8.411, p = 0.000$).

Figure 49 summarizes analytical metaphonological skills predicting different reading and writing difficulties.



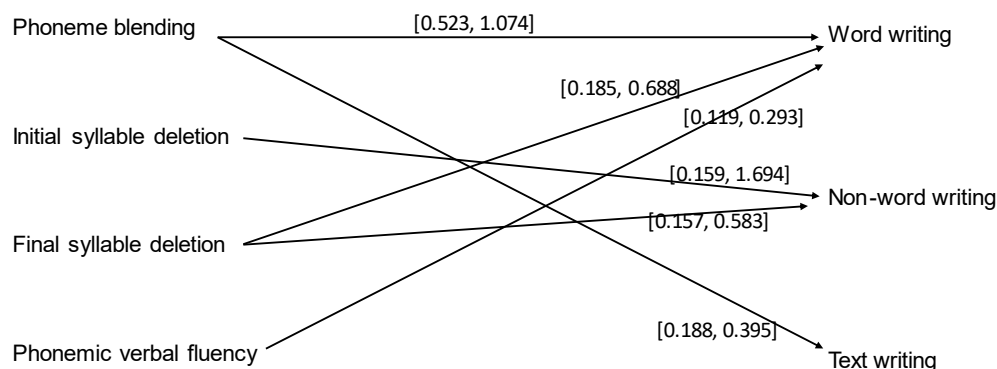


Figure 49. Significant predictive effects of analytical metaphonological skills on reading and writing difficulties. Min. and max. strength coefficients in different administrations are on the respective arrows. Complete statistics are in Appendix D (Table D.14).

There was no effect of bilingualism on any difficulty, with analytical metaphonological skills as fixed effects ($0.520 < p < 1.000$).

RAN & rhyme recognition

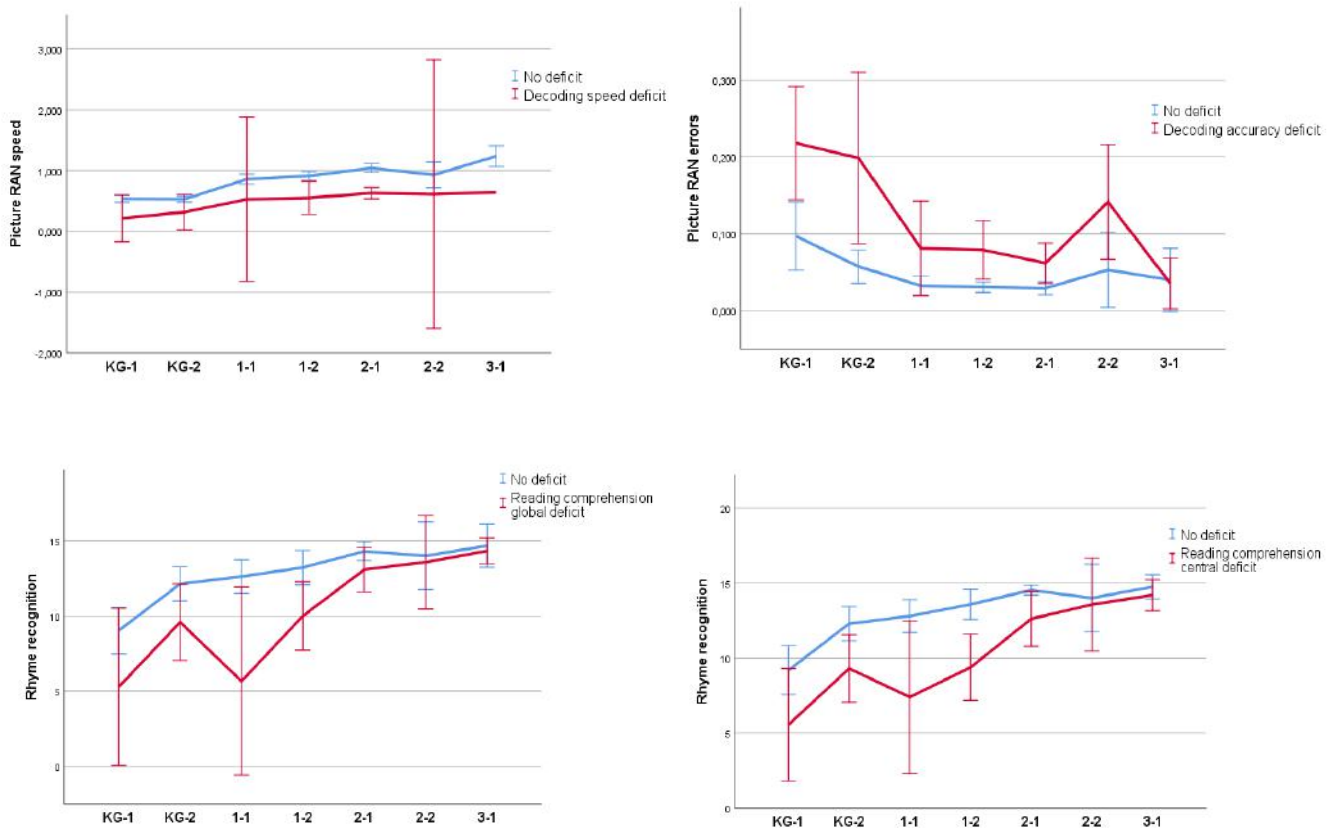
A series of logit GMMRM were analyzed with the presence/absence of difficulties as dependent variables, longitudinal scores obtained in the recognition of rhymes and RAN of pictures during the 7 test sessions from kindergarten to the 3rd grade of primary school as fixed effects, and bilingualism and grade as random effects. A backward stepwise procedure was adopted, as above.

The models significantly predict the decoding speed difficulty, with a main effect of RAN speed ($F(1, 48) = 27.672, p = 0.000$); the decoding accuracy difficulty, with a main effect of RAN errors ($F(1, 132) = 25.847, p = 0.000$); the reading comprehension difficulty, with a main effect of rhyme recognition (global comprehension $F(1, 141) = 26.905, p = 0.000$;

central comprehension $F(1, 168) = 27.863, p = 0.000$ specific comprehension ($F(1, 161) = 27.190, p = 0.000$); the word writing difficulty ($F(2, 101) = 14.477, p = 0.000$), with a main effect of RAN errors ($F(1, 94) = 20.768, p = 0.000$) and rhyme recognition ($F(1, 108) = 11.935, p = 0.001$); the non-word writing difficulty ($F(2, 73) = 10.054, p = 0.000$), with a main effect of RAN speed ($F(1, 63) = 8.754, p = 0.004$) and RAN errors ($F(81, 74) = 5.849, p = 0.018$); and the text writing difficulty, with a main effect of RAN errors ($F(1, 57) = 6.723, p = 0.012$).

Global and partial accuracy of predictive models are in Appendix D (Table D.15).

Figure 50 shows longitudinal results of children with and without difficulty in the testing of predictive abilities from kindergarten to the 3rd grade of primary school.



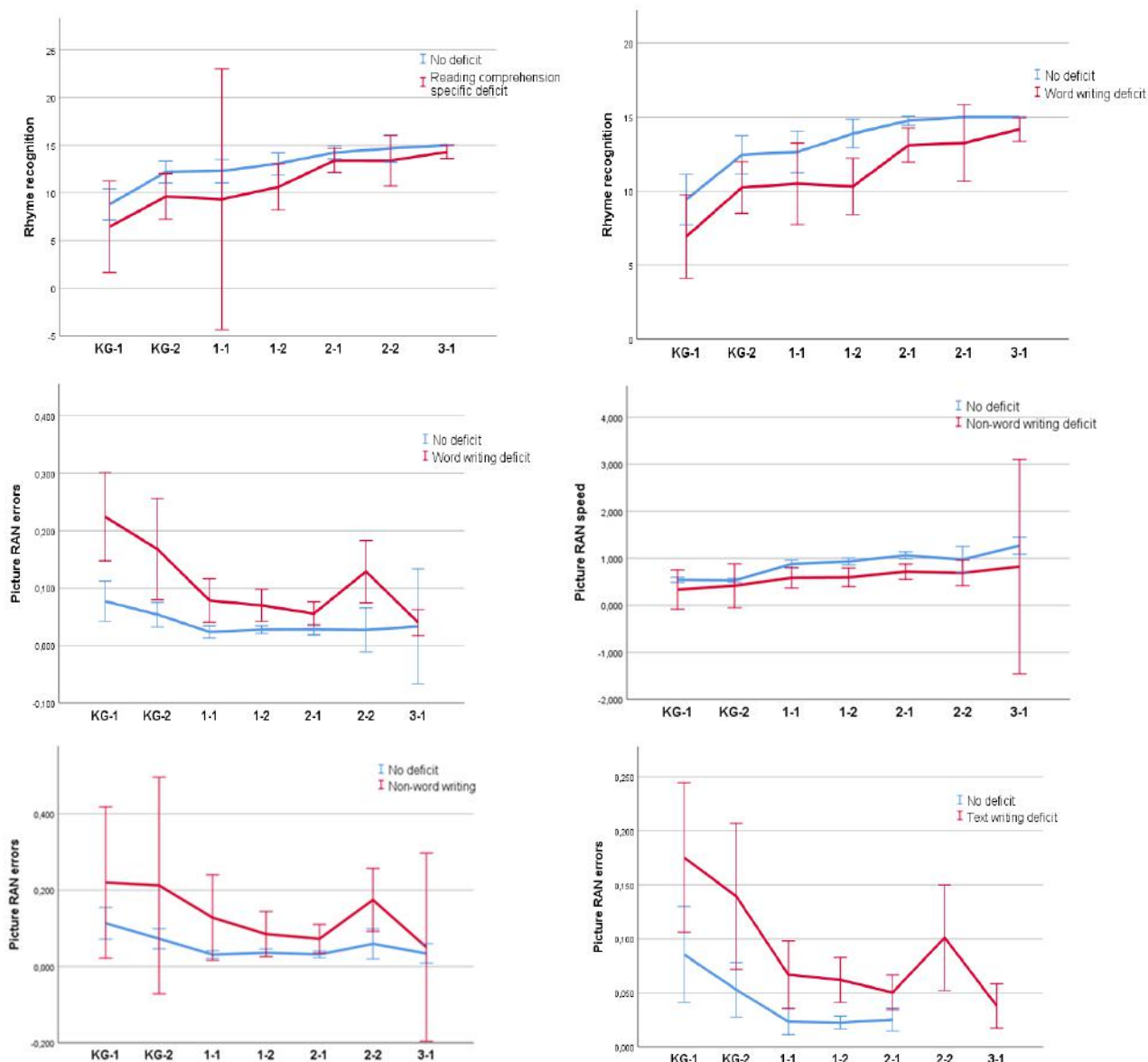


Figure 50. Longitudinal mean score of children with/without difficulty (lines) in the respective predictive picture-RAN abilities (y-axes) across the 7 testing sessions (x-axes) from KG to the 3rd grade.

An analysis of the combined effect between each predictor and the 7 levels of test-administration time as a fixed effect found:

- a positive effect of picture RAN speed on decoding speed in every administration ($F(7, 35) = 8.881, p = 0.000$);
- a negative effect of picture RAN errors on decoding accuracy in every administration except the last one ($F(7, 209) = 4.518, p = 0.000$);
- a positive effect of rhyme recognition on reading comprehension in every administration (global comprehension $F(7, 54) = 4.762, p = 0.000$; central comprehension $F(7, 62) = 4.514, p = 0.000$; specific comprehension $F(7, 47) = 4.929, p = 0.000$);

- a positive effect of rhyme recognition on word writing in every administration ($F(7, 53) = 3.011, p = 0.010$), and a negative effect of RAN errors in every administration except the last one ($F(7, 36) = 4.910, p = 0.001$);
- a positive effect of picture RAN speed on non-word writing in every administration ($F(7, 210) = 2.989, p = 0.005$), and a negative effect of picture RAN errors in every administration except the last one ($F(7, 48) = 4.499, p = 0.001$);
- a negative effect of picture RAN errors on text writing until administration 2-1, whereas no significant effect emerged in the last two administrations ($F(7, 85) = 3.706, p = 0.002$).

Complete statistics of effects are in Appendix D (Table D.16).

There was no effect of bilingualism or grade on literacy difficulties, after taking picture RAN and rhyme recognition as fixed effects ($0.305 < p < 1.000$).

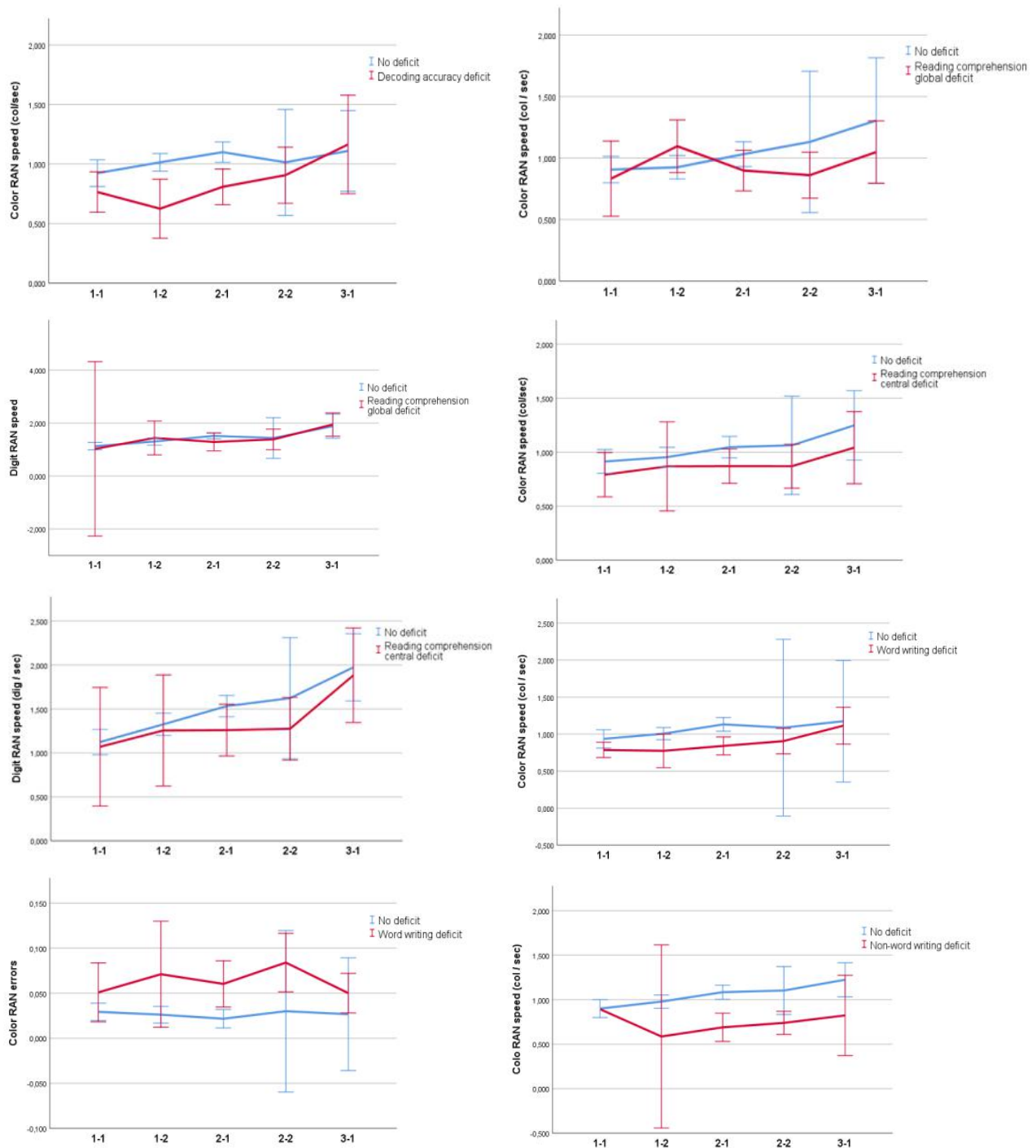
By taking speed and error scores obtained in the 5 test sessions of RAN of colors and digits in the 1st and 2nd grade as fixed factors and bilingualism and grade as random effects, the GMMRM predicts

- decoding accuracy, with a main effect of RAN speed of colors ($F(1, 129) = 19.228, p = 0.000$);
- reading global ($F(2, 126) = 3.397, p = 0.037$) and central ($F(2, 103) = 3.535, p = 0.033$) comprehension, with a main effect of RAN speed of colors (global comprehension $F(1, 126) = 6.771, p = 0.010$; central comprehension $F(1, 110) = 6.987, p = 0.009$) and digits (global comprehension $F(1, 126) = 3.881, p = 0.051$; central comprehension $F(1, 103) = 5.133, p = 0.026$);
- word writing ($F(2, 124) = 9.373, p = 0.000$), with main effects of speed ($F(1, 119) = 6.316, p = 0.013$) and errors ($F(1, 126) = 5.977, p = 0.016$) in the RAN of colors;
- non-word writing ($F(2, 79) = 8.174, p = 0.001$), with main effects of color RAN speed ($F(1, 84) = 7.666, p = 0.007$) and digit RAN errors ($F(1, 72) = 6.900, p = 0.011$).

No significant effect on decoding speed and text writing was found ($p = 1.000$).

The accuracy of predictive models is displayed in Appendix D (Table D.17)

Figure 51 shows the means obtained by children in each predictive RAN ability during the 5 test sessions in primary school.



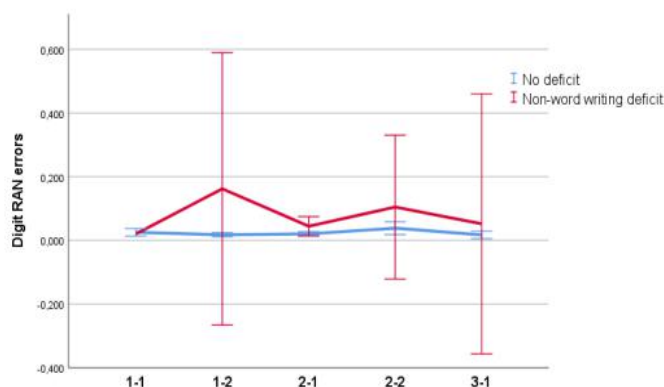


Figure 51. Longitudinal mean score (y-axes) obtained by children with/without difficulty (lines) in each predictive ability during the five test sessions (x-axes) in 1st, 2nd, and 3rd grade of primary school.

An analysis of the combination between each predictor and administration time as a fixed effect on respective difficulties found the following significant effects:

- a positive effect of RAN speed on decoding accuracy in every administration ($F(5, 125) = 4.400, p = 0.001$);
- a positive effect of color RAN speed ($F(5, 42) = 3.229, p = 0.015$) on reading central comprehension in the first 3 testing sessions, and of digit RAN speed ($F(5, 123) = 2.991, p = 0.014$) in the first two sessions, with a marginally non-significant effect in administration 2-1, whereas no effect on central comprehension emerged from administration 2-2 onwards;
- a positive effect of color RAN speed on word writing in every administration ($F(5, 41) = 3.224, p = 0.015$), and a negative effect of color RAN errors on word writing in administration 1-2, 2-1, and 2-2, with marginally non-significant effects in administration 1-1 and 3-1 ($F(5, 32) = 3.108, p = 0.022$);
- a marginally significant negative effect of digit RAN errors on non-word writing in administration 1-2 and 2-1, whereas no significant effect emerged in the first, fourth, and fifth administration ($F(5, 121) = 2.301, p = 0.049$).

Complete statistics of effects are in Appendix D (Tables D.17 – D.18).

Figure 52 summarizes the longitudinal predictive effects of RAN and rhyme recognition from kindergarten (pictures) or 1st grade (colors and digits) on different literacy difficulties found in 2nd and 3rd grade.

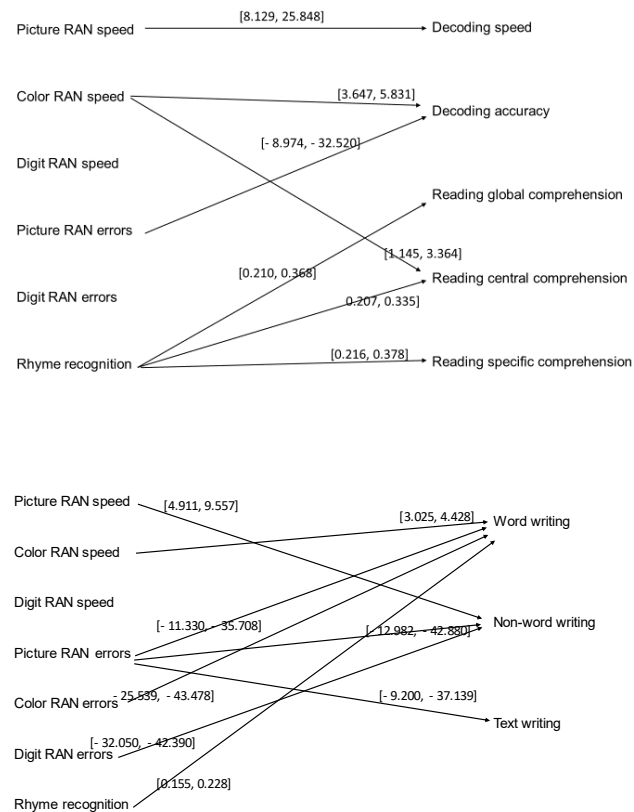


Figure 52. Significant predictive effects of RAN and rhyme recognition on different literacy difficulties. Minimum and maximum strength coefficients for the 7 administrations of picture RAN and rhymes, and for the 5 administrations of color and digit RAN are indicated on the respective arrows. Complete statistics are in Appendix D (Tables D-17 – D.18).

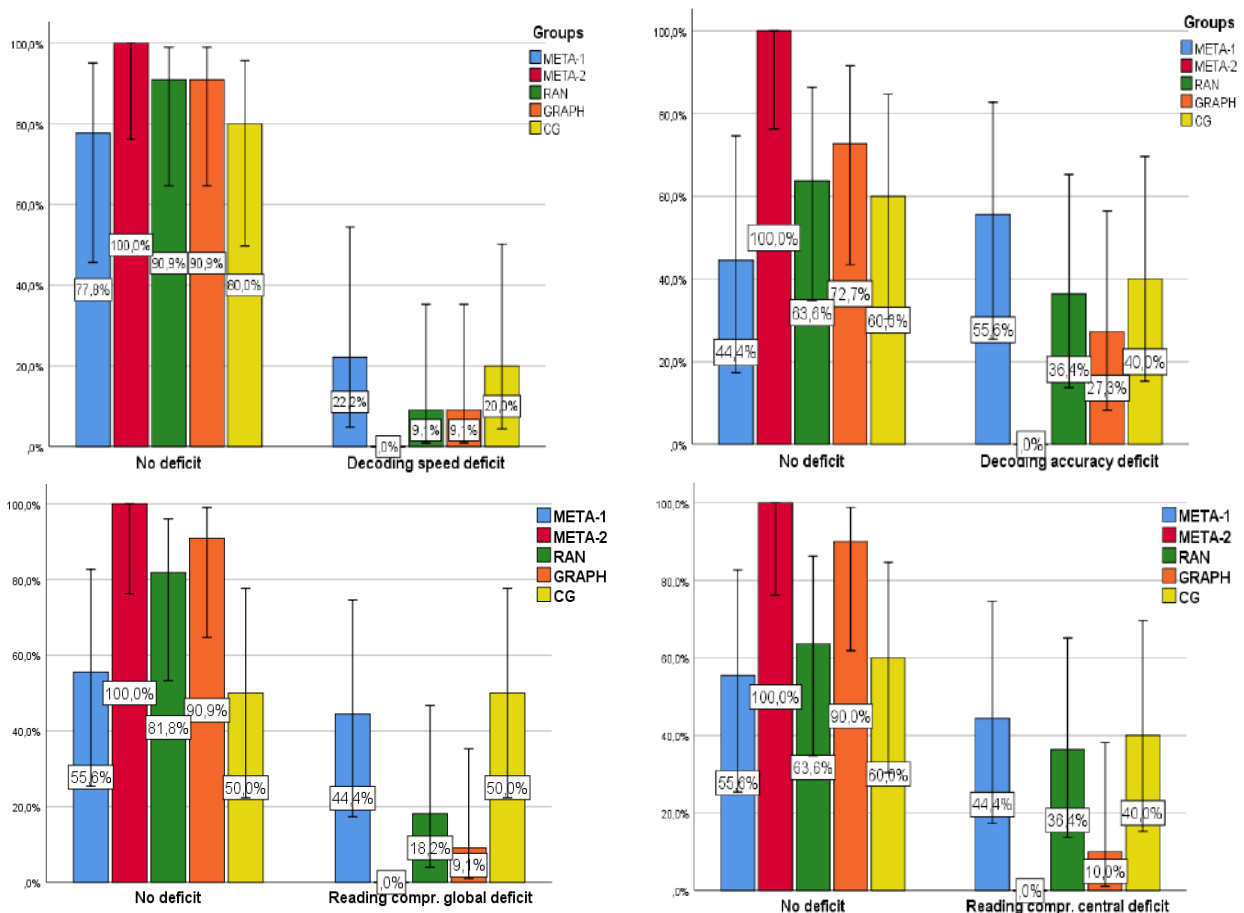
No effect of bilingualism or grade on any difficulty emerged after taking RAN and rhyme recognition as fixed effects ($0.092 < p < 1.000$).

6.3.3. Effects of the intervention on reading and writing difficulties

In this section, we explore the possible effects of different training administered in the last year of kindergarten and in the 1st grade of primary school to children who developed any reading or writing difficulty as compared to typically developing children. The META-1 group consisted of 6 children who entered the project in the 1st grade after teachers reported them to have problems in the acquisition of literacy skills. They attended one cycle of META training in the 1st grade of primary school, thus being monitored from the 1st to the 3rd grade.

Other children attended a double training, the former in kindergarten and the latter in the 1st grade, as follows. The META-2 group consists of 9 children who attended the double metaphonological training; the RAN group consists of 10 children who attended the double RAN training; the GRAPH group consists of 9 children who attended the double graphic training, which was conceived as a placebo. The CG consists of 10 children who did not attend any training. Seven of them were monitored from the 1st to the 3rd grade, and 3 from kindergarten to the 2nd grade.

Figure 53 shows the proportions of children of different groups who developed some literacy problems in 2nd and 3rd grade.



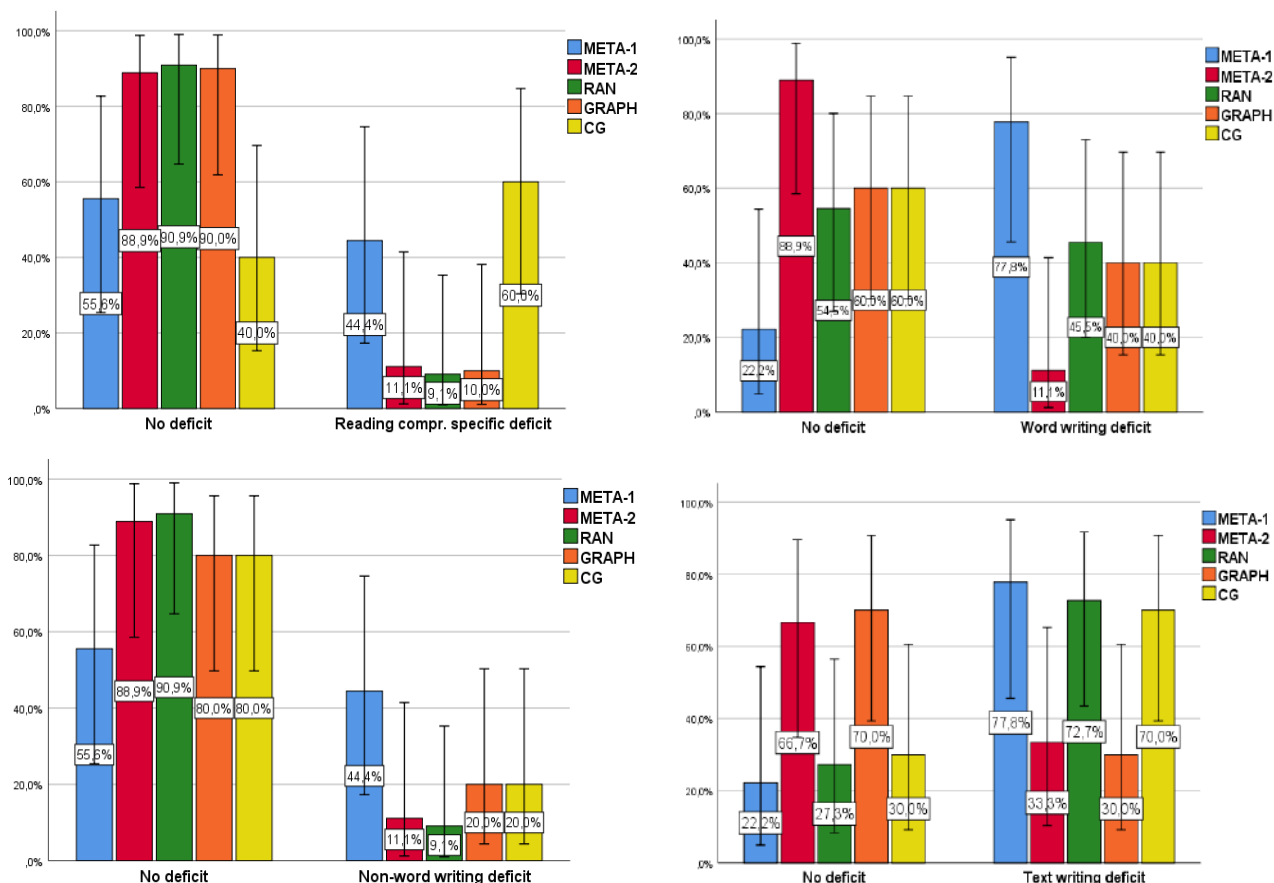


Figure 53. Proportions of children without/with literacy problems belonging to different (training) groups

A series of robust logit GLMM were analyzed to explore the predictive effect of group on different difficulties, with the presence/absence of the difficulty as a dependent variable, training group as a fixed effect, and bilingualism as a random effect. The models significantly predict the decoding accuracy (acc. = 70%) ($F(1, 46) = 4.870, p = 0.032$), the reading comprehension difficulty (global comprehension: acc. = 80%, $F(1, 45) = 9.525, p = 0.003$; central comprehension: acc. = 73.5%, $F(1, 45) = 8.106, p = 0.007$; specific comprehension acc. = 81.6%, $F(1, 44) = 20073.516, p = 0.000$); the word writing difficulty (acc. = 67.3%, $F(1, 45) = 20.952, p = 0.000$), and the non-word writing difficulty (acc. = 79.6%, $F(1, 44) = 5.930, p = 0.019$).

Table 58 displays significant predictive effects.

						Conf. Interv. (95%)	
Dependent variable	Coeff.	Est.	Std.Err.	t	p	Lower	Upper
Decoding accuracy	META-1	- 0.914	0.414	- 2.207	0.032	- 1.747	- 0.080
	META-2	2.289	0.202	14.016	0.000	2.423	3.236
Reading global comprehension	META-1	- 1.280	0.134	- 9.585	0.000	- 1.550	- 1.011
	META-2	2.016	0.079	25.674	0.000	1.858	2.175
	CG	- 1.503	0.487	- 3.086	0.003	- 2.485	- 0.522
Reading central comprehension	META-1	- 0.872	0.306	- 2.847	0.008	- 1.489	- 0.255
	META-2	2.425	0.094	25.768	0.000	2.235	2.614
Reading specific comprehension	META-1	- 1.939	0.629	- 3.084	0.004	- 3.206	- 0.672
	CG	- 2.786	0.020	- 141.681	0.000	- 2.825	- 2.746
Word writing	META-1	- 1.655	0.361	- 4.577	0.000	- 2.383	- 0.927
	META-2	1.662	0.584	2.846	0.007	0.486	2.838
Non-word writing	META-1	- 1.162	0.477	- 2.345	0.019	- 2.122	- 0.201
	RAN	0.883	0.422	2.091	0.042	0.032	1.733

Table 58. Predictive effects of group on the condition of absence of different difficulties.

The META-1 group predicts a higher presence of a difficulty in decoding accuracy, reading comprehension (both central and specific), word writing, and non-word writing. The META-2 group predicts a lower presence of the difficulty in decoding accuracy, reading central comprehension, and word writing. Furthermore, the RAN group predicts a lower presence of the non-word writing difficulty, and the CG a higher presence of the reading specific comprehension difficulty.

A logit GLMM with the decoding accuracy difficulty (present/absent) as the dependent variable, different errors in word decoding (word errors / non-word errors / word-pair errors / stress errors) as fixed effect, and bilingualism as a random effect was analyzed to explore error types that predict impaired decoding in the sample. The models correctly predict the poor decoding in both the 2nd ($F(4, 568) = 33.620, p = 0.000$) and 3rd grade ($F(4, 164) = 11.378, p = 0.000$).

Main effects of word errors ($F(1, 568) = 41.523, p = 0.000$) and word pair-errors ($F(1, 568) = 20.092, p = 0.000$) were found in the 2nd grade, and of word errors ($F(1, 164) = 6.484, p$

= 0.012), stress errors ($F(1, 164) = 24.981, p = 0.000$), and word-pair errors ($F(1, 164) = 9.145, p = 0.003$) in the 3rd grade.

In the 2nd grade, impaired decoding accuracy was predicted by errors in decoding generic words ($Coeff = -25.637, SE = 3.978, t = -6.444, p = 0.000, CI [-33.451, -17.822]$) and attentional errors in decoding word pairs ($Coeff = -41.291, SE = 9.212, t = -4.482, p = 0.000, CI [-59.384, -23.198]$), whereas in the 3rd grade by errors in decoding generic words ($Coeff = -32.529, SE = 12.774, t = -2.546, p = 0.012, CI [-57.752, -7.305]$), stress errors ($Coeff = -39.019, SE = 7.807, t = -4.998, p = 0.000, CI [-54.434, -23.604]$), and word-pair errors ($Coeff = -59.177, SE = 19.568, t = -3.024, p = 0.003, CI [-97.816, -20.537]$).

6.4. Discussion 4

6.4.1. Individual predictors of literacy difficulties

We longitudinally analyzed the individual, linguistic, and metalinguistic predictors of literacy difficulties and possible effects of metaphonological training in a sample of children who were under the norm in reading or writing in 2nd and 3rd grade.

We firstly selected individual characteristics that were significantly associated with literacy difficulties to establish if any of them had predictive power. We found significant effects of social deprivation conditions on text decoding speed, accuracy, and non-word writing. Social deprivation showed high specificity for decoding speed (100%), in particular, but also decoding accuracy (88.2%) and non-word writing (87.2%), whereas the sensitivity was null for decoding speed, quite good for non-word writing (80%), and insufficient for decoding accuracy (56.2%). The combination of social deprivation with LI predicted word writing accuracy with high specificity (92.9%) and moderate sensitivity (71.4%). According to our

data, social deprivation appears to be a prominent index of literacy difficulties, in particular of phonological problems in writing and, in combination with LI, of word writing accuracy. There is a bidirectional link between social deprivation and LI. On the one hand, socially disadvantaged children are exposed to quantitatively (Hart & Risley, 1995) and qualitatively (Peterson et al., 2003) impoverished input, which can limit language development. Locke et al. (2002)⁵⁶ showed that in some areas of the UK, more than 50% of children in the early years of primary school reveal depressed language skills compared to their general cognitive abilities and the population. Moreover, Burt et al. (1999)⁵⁷ found that metaphonological skills in preschool children are significantly affected by socio-economic status. On the other hand, limited language development hampers relationships, behavior⁵⁸, and learning (Sage, 2005). Spencer et al. (2017) investigated the school-achievement predictors of 300 teenagers (13 – 14 years old) and found that language ability, in particular vocabulary, plays an important role in educational achievement. To reduce the gap between groups with different socioeconomic backgrounds in educational attainment, they suggest on-going support for spoken language abilities throughout secondary education and a potential role for continuing provision of speech and language therapy. The issue is even more complicated in multicultural contexts, where possible social deprivation can be interspersed with insufficient language input. Among bilingual children tested in the middle of the 1st grade in 2019 within the present project, 44% reported a raw score below 50 in the Italian receptive vocabulary test (PPVT). This does not exclude that their bilingual vocabulary might be much wider and they can have sufficient pragmatic skills for satisfying communication, but this score seems too low for children who were born in Italy and

⁵⁶ Locke et al. (2002) found girls' receptive language skills to be significantly better than those of boys.

⁵⁷ Burt et al. (1999) did not find significant differences between males and females in phonological awareness.

⁵⁸ High incidence of *Speech, Language and Communication Needs* (for the definition of SLCN, see Dockrell et al., 2012) often unidentified, is reported in young offenders (Bryan, 2004; Snow & Powell, 2008), looked after children (Cross, 1999), children with conduct disorder (Gilmour et al., 2004), and other social, emotional and behavioral difficulties (Toppelberg & Shapiro, 2000).

attended the Italian kindergarten. These data, matched to those showing a significant increase of children's vocabulary only between kindergarten and the 2nd grade (see 3.3.1), encourage the support of language enhancement interventions since kindergarten in multicultural environments, in particular in areas that are socio-economically depressed. This might help the definition of more precise language profiles of children, as not to confuse a common bilingual behavior or a language need due to reduced Italian input with a developmental LI. Presently, the diagnosis of LI in bilingual children, especially with minority languages, is made by using monolingual norms in Italy, which are misleading because standardized on monolingual children, whereas we have seen that there are differences in (meta)language developmental pace of monolingual and bilingual children (see Study 1). At the moment, no standardized clinical tools are allowing a diagnosis of LI in Italian speaking bilingual children. By taking into consideration some good practices in other languages (see, e.g., Armon Lotem, 2011; Gillam et al., 2013; Fleckstein et al., 2018), the diagnosis of LI in bilingual children, possibly without testing both languages, represents an interesting challenge for researchers, which could be faced through interactive cooperation of clinicians and linguists.

The present data suggest that socially deprived children are highly likely to present a phonological difficulty in writing, which can be worsened by the presence of a language impairment. It is then important to offer complex stimuli and input to children since early infancy, to make socio-economic status “a proxy for the opportunity to learn” (Kuhl, 2011).

6.4.2. . Memory, linguistic and metalinguistic predictors of literacy difficulties

Short-term memory, phonological, and visual skills

Short-term/working memory showed an important role in every literacy difficulty, with some differences among abilities. Problems in decoding speed were predicted by the

combination of digit span forward and backward, which showed high specificity (97.5%) but almost null sensitivity (7.1%). Decoding accuracy was predicted by the combination of digit span forward and word span, with high specificity as well (91.3%) and a bit higher but insufficient sensitivity (45.2%), and by visual discrimination of capital letters (specificity = 92.8%; sensitivity = 34.3%). Although none of these predictors has the sensitivity to be considered a clinical marker, the specificity is very high, meaning that we can expect not to find a decoding difficulty in children with good memory span. Working memory seems more important for decoding speed, whereas short-term memory span, along with the ability to visually discriminate capital letters, for decoding accuracy.

Working memory also showed an important role in the reading comprehension difficulty, which was globally predicted by the digit span backward and visual working memory, with moderate sensitivity ($70\% < \text{sens.} < 79\%$) and good specificity (90%). Central comprehension seems a more complex operation, which requires the mutual support of working memory (sens = 44%; spec = 92%), phonological memory (sens. = 33%; spec. = 98%), and visual memory (sens. = 57%; spec. = 87%) whereas working memory (sens. = 72%; spec. = 95%) is sufficient to accomplish specific comprehension. These data reveal that working memory has a central role in reading comprehension and a good working memory span can guarantee a good global and specific comprehension of written texts. To explore whether the predictive effects are significant across the whole period considered, we analyzed the combination between each predictor and administration time (as a fixed effect of regression). As a result, the effect of digit span forward on decoding speed as well as that of letter visual discrimination on decoding accuracy disappeared, thus showing that the only memory predictors of written decoding that were significant from kindergarten to the 3rd grade were digit span backward for decoding speed, and the combination between digit span forward and word span for decoding accuracy. Both predictors have a similar effect in every administration, with a particularly robust effect of digit span backward on decoding

speed. This means that short-term processing abilities are more important for decoding speed, whereas short-term verbal storage is more important for decoding accuracy.

The effects of visual and phonological memory on reading comprehension disappear when considering their combined effect with time so that the main predictor of reading comprehension remains digit span backward. Its effect is significant in every administration for global and central comprehension, although the effect on global comprehension in 3rd grade is marginally significant. This is due to the effect of the central comprehension component, which is no more significant in the 3rd grade. The reason might lie in texts' nature, which gets more and more demanding as the school grade increases. In particular, the two texts proposed by Cornoldi & Carretti (2016) are both narrative until the 2nd grade, whereas one informative text is introduced in the 3rd grade. There was no significant difference, however, between global comprehension scores of the two text types in the 3rd grade. Only two children in the narrative text and one in the informative text had good global comprehension scores in 3rd grade, whereas the others placed between (– 1.5) and (– 2.4) *SDs* under the norm for the narrative text, and (- 1.2) and (- 3.1) *SDs* under the norm for the informative text. It is worth noting that the only child that was in the normal range for both texts was bilingual with Arabic mother language.

The effect of digit span forward and word span on the word writing difficulty was significant in every administration, whereas the combination of phonological working memory with time effects tended to lose its significance. In the first steps of writing acquisition, namely until the middle of the 1st grade, visual memory had also a significant effect on word writing, which disappeared from the end of the 1st grade onwards.

The predictive effect of digit span forward on the non-word writing difficulty is then significant in every administration, whereas that of capital-letter visual discrimination is marginally significant in the first two administrations (KG-2 and 1-1) and disappears from the end of the 1st grade. Working memory is the main contributor to text writing, which seems to require

higher processing abilities than word writing, which relies essentially on storage functions. Longer texts and different administration conditions (individual with repetitions for (non)word writing, and collective without repetition for text dictation) might require different loads of working memory to fulfill the dictation task (see 3.3.6). Short-term memory has a significant role in word and non-word writing, with an additional effect of word storage (word span) abilities for word writing but not for non-word writing. Visual skills significantly contribute to writing in the first steps of acquisition, namely until the middle of the 1st grade. In particular, visual memory contributes to word writing, whereas visual discrimination skills of letters marginally contribute to non-word writing. This might be evidence of the fact that the domain-general central executive of working memory is engaged in early reading acquisition, whereas verbal working memory is more and more specifically implicated in the reading performance as long as children gain more experience (Peng et al., 2017).

In Study 2, we hypothesized that, in typically developing children, the contribution of working memory to text dictation gets less and less prominent as long as children's abilities increase due to automaticity processes. Since working memory has a prominent role in the discrimination of the text writing difficulty, children with this difficulty might present problems in the automaticity of related processes.

The analysis of random variables found very small but significant negative effects of bilingualism on decoding accuracy, global reading comprehension, and text writing, matched to very small but significant positive effects of monolingualism on reading comprehension and text writing, whereas a positive effect of monolingualism on decoding accuracy was marginally non-significant. Differences between monolinguals and bilinguals in decoding accuracy, reading comprehension, and text writing were treated in chapter 3. Although the effects of bilingualism are very small, the present results suggest that such effects might interfere with the definition of possible reading and writing impairments. It seems then important, in particular in bilingual children, to assess different aspects of decoding

(namely word and non-word beyond text decoding), comprehension (namely oral beyond reading comprehension), and writing (namely word, non-word, and sentence beyond text writing) to assess the possible presence of a literacy disorder. Particular attention should be paid to the conditions in which the writing task is performed. At the same time, working memory tests can help to disentangle possible ambiguous situations. In this regard, the present data suggest that a small digit span backward, namely impaired working memory, predicts with moderate sensitivity and high specificity a reading comprehension difficulty, and with good sensitivity (but modest specificity) a text writing difficulty; a small digit span forward in combination with a small word span, namely impaired verbal short-term memory, predicts with moderate sensitivity and high specificity a word writing difficulty. As for non-word writing and decoding accuracy, the high specificity suggests the unlikelihood to find these difficulties in children presenting, respectively, good digit span forward, and good digit span forward and word span, namely good storage capacity.

Working memory is known as an established predictor of reading comprehension (for children, see Cain et al., 2004a; for adults, see Daneman & Merikle, 1996). The present outcomes are in line with studies that consider processing capacities tapped by working memory tasks rather than the storage component tapped by short-term memory tasks as particularly important in explaining reading comprehension individual variance (Cain et al., 2004a, 2004b; Daneman & Merikle, 1996). Nouwens et al. (2017), through an adaptation of a conceptual span test (Haarmann et al., 2003) and a competing language processing task (Gaulin & Campbell, 1994), found that semantic working memory explains more individual variance than all other memory measures, and both semantic storage and processing components have a role in explaining Dutch children's reading comprehension.

At the same time, we are also in line with the literature considering working memory as crucial in writing processes (Gathercole et al., 2006; Kellogg, 1996; McCutchen, 2000; Swanson & Berninger, 1996). The present data show that working memory is diversely

affected by different dictation tasks, as text writing in collective dictation relied more on working memory resources, whereas individual word and non-word writing was mainly affected by phonological storage capacity.

We are also in line with mainstream literature considering verbal short-term memory as positively related to decoding accuracy in children (for reviews, see Jorm, 1983; Stanovich, 1982) and adults (Baddeley et al., 1975; Daneman & Carpenter, 1980).

Language skills

Among linguistic predictors, productive vocabulary plays a key role, as it predicted every difficulty except that in reading global comprehension. Moreover, syntactic production of clitic pronouns predicted decoding speed and decoding accuracy problems, whereas syntactic comprehension predicted the reading central comprehension difficulty, word writing, and non-word writing difficulties. The specificity was high for every predictive effect, whereas the sensitivity had a good level for productive vocabulary as a predictor of the reading comprehension specific difficulty (sens. = 76.9%; spec. = 89%), the word writing difficulty (sens. = 81%, spec. = 80%), the text writing difficulty (sens. = 82.1%, spec. = 83.3%), and for syntactic comprehension as a predictor of the word writing difficulty (sens. = 71.4%, spec. = 79.2%). The strongest effects were those of syntactic production as a predictor of the decoding speed and decoding accuracy difficulty, and of syntactic comprehension as a predictor of word writing, non-word writing, and reading comprehension difficulties. The effects of productive vocabulary on non-word dictation and reading comprehension specific difficulty were also strong. This means that children with a wide productive vocabulary are unlikely to present a reading specific comprehension difficulty, a word writing difficulty, and a text writing difficulty, and children with good syntactic comprehension skills are also unlikely to present a word writing deficit.

We confirm an important role of (morpho)syntactic skills on literacy (Casani, in preparation) and reveal a particular influence of this ability on word and non-word writing. Further research is in progress to investigate which aspects of (morpho)syntactic competence are responsible for variance in literacy tasks, whether they concern the superficial recognition of morphemic units (Rastle et al., 2004), the semantic level (Feldman et al., 2012), or both (Feldman et al., 2009). The fact that general syntactic comprehension, beyond syntactic production, contributes to writing and, at the same time, the involvement of non-words, suggests that a combination of morpho-semantic and morpho-orthographic reasons (Feldman et al., 2009) might be plausible. Namely, children might avail of semantic strategies for reading known words and orthographic strategies for reading unknown words and pseudowords. Similar reasons, namely children's familiarity with lexicon structural chunks, might underlie the effect of vocabulary on non-word writing. In this regard, a study conducted on fourth-grade students found that expressive vocabulary is related just to exception word reading, whereas only receptive vocabulary is associated with non-word decoding skills (Ouellette & Beers, 2010). Differently, Ricketts et al. (2007) suggest that oral expressive vocabulary is related to the decoding of exception words and reading comprehension in children from 8 to 10 years old, but not to the decoding of regular meaningful words or nonwords. These studies could be indirectly affected by the opaque orthography of the English language, where previous knowledge of words that are proposed for reading is a crucial aspect. By controlling for this variable, Mitchell & Brady (2013) showed that vocabulary knowledge has a significant relationship with reading novel words with a variety of orthographic patterns. The present data stress the role of expressive vocabulary, in particular, on text decoding speed and accuracy and on word and non-word writing, beyond reading comprehension

Global metaphonological skills

To verify possible early longitudinal predictors of 2nd and 3rd-grade literacy difficulties, kindergarten global metaphonological skills were analyzed. Recognition of the initial syllable was revealed to be a significant predictor of decoding accuracy, reading comprehension (central and specific), word and non-word writing, and text writing difficulties. The specificity was high for every effect (81% < spec. < 99%), and a moderate sensitivity was found for reading comprehension (70%), word writing (64.7%), and text writing (64.3%). The strongest effects were on word dictation and global reading comprehension, followed by the other measures of reading comprehension and text writing. This means that, according to the present sample, children who are good at the recognition of the initial syllable in kindergarten are unlikely to develop a reading comprehension difficulty and word and text writing difficulties in primary school. The ability to identify word onsets constitutes the first step to the subsequent acquisition of manipulation skills.

The analysis of random effects revealed very small but significant negative effects of bilingualism and positive effects of monolingualism on reading comprehension, word writing, and non-word writing. For central reading comprehension and non-word writing, the weak positive effect of monolingualism was present only in the 2nd grade, and the negative effect of bilingualism was present only in the 3rd grade. As for reading comprehension, these data confirm the expected increasing difficulty of bilingual children in 3rd grade (see 3.4.6), probably due to the complexification of texts. As for non-word writing, data indicate a deviation from the trend analyzed in 3.3.6 and 3.4.6, where no difference emerged between monolingual and bilingual children in non-word writing. Bilingual children in 3rd grade, however, are only 2, hence the present data are not generalizable. One of them was 1.5 SDs under the norm, and the other one could not write any non-word. The latter could not read and write and presented diverse comorbidities beyond different cases of learning disabilities in her family.

In the 2nd grade, only 3 out of 17 bilingual children were in the clinical range as compared to the Italian monolingual norms. One of them ($Z = - 3.3$) who could not read and write, had social deprivation, diverse comorbidities, and different cases of learning disabilities in her family; another one ($Z = - 4.7$), who also could not read and write, lived severe conditions of social deprivation; and the last one ($Z = - 2.0$) was deprived of Italian input for being absent from school for several months during the 1st grade. All the above-mentioned cases had mother languages that were typologically distant from Italian. These data confirm that social deprivation is related to phonological difficulties in writing, which might match with already complex situations of impoverished input in which several bilingual children from migrant communities live. In these cases, the evaluation of possible learning disorders requires high attention levels, which include a thorough analysis of children's linguistic profile and social framework.

Analytical metaphonological skills

Analytical metaphonological skills in 1st and 2nd grade showed significant predictive effects of the combination between initial syllable deletion and phonemic verbal fluency on the decoding speed difficulty, and of the combination between final syllable deletion and phonemic verbal fluency on the decoding accuracy difficulty. The combination of phoneme blending, final syllable deletion, and phonemic verbal fluency also predicted reading global comprehension. The same combination predicted the word writing difficulty, whereas the combination of initial syllable deletion, final syllable deletion, and phonemic verbal fluency predicted non-word writing, and the only phoneme blending predicted the text writing difficulty. Complex metaphonological abilities, such as syllable manipulation, and retrieval abilities requiring a high level of executive control over cognitive processes, such as phonemic verbal fluency (see 4.4.4), seem to be responsible for reading and writing processes in primary school. These abilities give a combined contribution to literacy skills,

in association with working memory and automaticity processes (see 4.4.2). The mentioned combinations have high specificity (except for the effect of phoneme blending on text writing, which, conversely, presents a high sensitivity), meaning that children who show good performances in the mentioned predictive abilities are not likely (with a probability of 68% for word writing, and between 93% and 100% for the other abilities) to present these problems. In this view, the high sensitivity of the only phoneme blending for text writing should not be interpreted as an absolute predictive power of phoneme blending in detecting writing difficulties, even in consideration of the scarce specificity level (30%), but as the property of an essential component of the (text) writing process, whose absence predicts with a high likelihood (93%) the presence of a weakness. At the same time, text writing should not be considered as an easier task than word/non-word writing for showing the contribution of fewer metaphonological skills, but as a complex task that, for the quantity and quality of operations to be accomplished in limited time intervals and without repetitions (for the administration mode see 2.12; for possible effects on the performance, see previous discussions), requires highly automatized skills.

The analyses of random variables did not show any significant effect, meaning that bilingualism interferes with the prediction of some literacy difficulties like reading comprehension and word/non-word writing only in preschool years.

To have a more definite idea of the role of time on the predictive power of analytical metaphonological skills, we analyzed the combined effect of each predictor with administration time. We found that initial syllable deletion and phonemic verbal fluency were stable combined predictors of the decoding speed difficulty until the end of the 2nd grade. Later, initial syllable deletion lost its predictive power. Moreover, the combination of final syllable deletion and phonemic verbal fluency was a stable predictor of the decoding accuracy difficulty until the 3rd grade. The effect was very strong in the first administration, especially for final-

syllable deletion. It was prominent in administration 1-1 and tended to progressively decrease in the following test sessions.

The present data suggest that different manipulation skills contribute to decoding speed and accuracy, respectively. Deletion of the initial syllable, in combination with phonemic verbal fluency, contributes to decoding speed, as children rely on their familiarity with word onsets to expect whole words. At this moment, phonemic verbal fluency could intervene to quickly retrieve the lemma from the lexicon storage. The ability to recognize and manipulate the initial syllable, then, could work as a phonological hint activating the word retrieval, which is implemented by the phonemic-verbal-fluency ability. This might be the reason why decoding speed is predicted by the combination of these two skills. As long as children acquire skills allowing them to automatize these processes, the explicit role of initial-syllable manipulation abilities disappears. From the third administration in the middle of the 2nd grade, the contribution of initial-syllable deletion to decoding speed got marginally significant ($p = 0.051$), and at the end of the 2nd grade, it was not significant anymore. The role of phonemic verbal fluency, instead, was stable across time because this ability, which we have considered as a cognitive function rather than a metalinguistic skill (see 4.3.4), was kept necessary for quick lexicon retrieval. The difference might be in that more advanced readers do not rely on phonological hints to retrieve the lemma, but on more global strategies developed through automatization processes and on the possible use of the lexical route of reading.

Differently, decoding accuracy relies on the ability to manipulate final syllables, in combination with phonemic verbal fluency. This manipulation skill stably contributed to decoding accuracy across time, until the last administration in 3rd grade. Phonological hints given by word onsets could be sufficient to guarantee decoding speed but not accuracy, which, conversely, relies on the ability to recognize and manipulate word endings. In endings, readers find morphology inflections, and their ability to manipulate those makes

decoding of the whole word easier. The role played by (morpho)syntactic skills on decoding accuracy (Casani, in preparation) and speed is probably in this vein. The fact that final syllable deletion contributes to reading comprehension, which is strictly related to syntactic comprehension (see Study 2), seems to corroborate this hypothesis. Namely, the sensitivity toward word endings allows more accurate processing of morphemes, indirectly contributing to text comprehension. Final-syllable manipulation skills contribute to reading comprehension until the middle of the 2nd grade, as well as phonemic verbal fluency, meaning that silent reading aimed at comprehension and aloud reading aimed at accuracy tend to become independent processes as long as children acquire practice. In this regard, we showed (see 4.4.3) how reading comprehension in typically developing readers can be more related to the text in the first steps of literacy acquisition (age 6;7 – 9;3), and rely on different strategies in more advanced steps (age 9;5 – 11;3) (Casani, 2020c). In this vein, Kim et al. (2011) found that first-grade skilled readers benefit from context when reading connected text, differently from average readers, who may tend to focus on word decoding (see also Jenkins et al., 2003; Stanovich, 1980). Moreover, silent reading fluency was found to be a stronger predictor of reading comprehension in skilled than in average first-grade readers (Kim et al., 2011).

In our results, reading comprehension was also predicted by phoneme blending only in the middle of the 1st grade, but this effect was already automatized at the end of the 1st grade. This might indicate a progressive removal of text-related reading comprehension strategies. Phoneme blending, final-syllable deletion, and phonemic verbal fluency predicted word writing, whereas initial-syllable deletion and final-syllable deletion predicted non-word writing until the 3rd grade. The main difference is in the contribution of phoneme blending and phonemic verbal fluency in word writing but not in non-word writing. Phoneme blending, which shows the strongest effect in every administration, contributes to accurate phoneme sequencing. Phonemic verbal fluency intervenes in the retrieval of known words.

Manipulation of the final syllable might contribute to writing accuracy, perhaps as a discriminator of inflected words.

The absence of phoneme blending among the predictors of non-word writing suggests that children might process these units in chunks, possibly recurring to their orthographic lexicon to write them (for the facilitating effect of morphologically recognizable chunks on pseudoword reading, see Traficante et al., 2011).

RAN and rhyme recognition

Picture RAN and rhyme recognition tests were administered 7 times, from kindergarten to the 3rd grade, so they can give information about the development of these abilities during the whole course of the project.

Picture RAN speed confirmed to be a predictor of the text decoding speed difficulty with 90% specificity but almost null sensitivity (5%), meaning that good RAN speed is expectable in children without decoding speed difficulties, but we cannot be sure to find a RAN deficit in children with decoding speed problems. This is in line with the *double deficit hypothesis* (Wolf & Bowers, 1999), which highlights that either a naming speed deficit or a phonological awareness deficit can cause reading difficulties. In this context, a RAN deficit indicating weakness in one or more of the underlying fluency-related processes can be present or absent, possibly co-occurring with a phonological-awareness deficit⁵⁹.

Picture RAN errors predicted the decoding accuracy difficulty with high specificity (90%) but weak sensitivity (45%).

Rhyme recognition predicted global reading comprehension with high specificity (90%) and quite good sensitivity (76%). This ability reveals to be a good predictor of specific comprehension, in particular (spec. = 91%; sens. = 78%), whereas it does not have sufficient

⁵⁹ A percentage from 60% to 75% of subjects with reading or learning disorders is reported as affected by a RAN deficit (Katzir et al. 2008; Waber et al., 2004; Wolf et al., 2002).

sensitivity for central comprehension (45%). Engen & Høien (2002) found that phonological awareness had a direct impact on the reading comprehension of 1300 Norwegian children in 1st grade. They did not use rhymes, but they found different indexes of syllable and phonemic awareness indicating that phonological factors play an independent role in text processing. They explained their outcomes through the fact that vocabulary and short-term memory, which are two crucial factors in comprehension, were both partly determined by phonological abilities. They also hypothesized that phonological awareness might reflect metacognitive processes that are assumed to be involved in reading comprehension.

We have shown above that the reading comprehension difficulty is predicted by final syllable deletion, and we hypothesized that the ability to manipulate word endings, which also include inflectional morphemes, might advantage children's comprehension. As a predictor of decoding accuracy (see above), we also assumed final syllable deletion as a measure of decoding accuracy. Initial syllable deletion, conversely, was assumed as a measure of decoding speed, which operates as a phonological hint allowing the activation of quick word retrieval, whose assumed measure was phonemic verbal fluency. Rhyme recognition does not seem a measure of decoding accuracy, either for the variability of rhymes size or for the fact that the simple recognition ability, as opposed to the manipulation ability required for final-syllable deletion, might not be sufficient to guarantee accuracy⁶⁰. Nevertheless, the ability to recognize rhyme-equivalent ending chunks was a robust predictor of the reading comprehension difficulty from the last year of kindergarten to the middle of the 3rd grade, with high specificity and quite good sensitivity. This might make it an efficient predictive tool for reading comprehension. Specific and global comprehension, in particular, maintained a quite good level of sensitivity (80% and 75%, respectively) besides high specificity (91% and 90%, respectively) after controlling for the effect of time.

⁶⁰ In this regard, it could be interesting to test relations of rhyme production with decoding accuracy.

The word writing difficulty was predicted by the combination of picture RAN errors and rhyme recognition, with moderate specificity (74%) and sensitivity (70%). The non-word writing difficulty was predicted by the combination of picture RAN speed and errors, with high specificity (97.2%) but scarce sensitivity (49%). Text writing was predicted by picture RAN errors, with high specificity (91%) but quite scarce sensitivity (62%). The effects of RAN speed and rhyme recognition were present in every administration, whereas that of RAN errors disappeared in the 3rd grade. These data, in line with the literature (Ziegler et al., 2010), confirm RAN speed as a predictor of text decoding speed. Also, a predictive role of picture RAN accuracy on decoding accuracy emerged.

The other measures of RAN tested in the 5 primary-school sessions were analyzed. RAN speed of colors predicted the decoding accuracy difficulty with good specificity (88%) but scarce sensitivity (34%). In association with digit speed, it also predicted global reading comprehension with high specificity (93%) and moderate sensitivity (67%). In association with color errors, it predicted the word writing difficulty (spec. = 86%; sens. = 64%), and with digit errors, the non-word writing difficulty (spec. = 100%; sens. = 21%;).

The effect of color RAN speed on decoding accuracy and word writing was significant in every administration, whereas that on central reading comprehension was significant until the middle of the 2nd grade; the effect of digit RAN speed on central reading comprehension, in turn, was significant only in the 1st grade. The effect of color RAN errors on word writing was significant until the end of the 2nd grade. Even in this case, none of the mentioned measures has the sensitivity to be considered a clinical marker of any difficulty, but data suggest that a good and accurate performance in the RAN of colors predicts the absence of a decoding accuracy difficulty and a word writing difficulty until the end of the 2nd grade. A good performance in both color and digit RAN in the 1st grade, also predicts a normal reading comprehension performance.

These data indicate that both RAN speed and RAN accuracy impact different measures of literacy to different extents. Picture RAN is the most consistent with reading, as picture RAN speed predicts decoding speed, and picture RAN accuracy predicts decoding accuracy in every administration. The predictive role of RAN on decoding accuracy is not shared by mainstream literature, which considers RAN as mainly related to reading speed rather than reading accuracy (Ziegler et al., 2010). This might however be a bias, since research on naming has mainly focused on speed (Vaessen et al., 2009), whereas accuracy has mainly been examined as related to phonemic awareness (Martin et al., 2010; Patel et al., 2004; Talli et al. (2015), except for rare exceptions. Casani (2020b) found that a small sample of children with dyslexia in primary-school with and without developmental language disorder (age 6;7 – 11;3) had similar reading problems but differed in that language-impaired children had a naming-accuracy deficit in addition to the reading deficit. Di Filippo et al. (2005) examined 281 Italian children from the 1st to the 6th grade and found that the naming performance predicted reading speed and accuracy but not reading comprehension. Liao et al. (2015) analyzed some of the mechanisms underlying the relationship between RAN and reading accuracy and fluency in Mandarin Chinese. They found that controlling for nonverbal IQ and orthographic processing explained the relation of RAN with reading accuracy but not with reading fluency. Moreover, they revealed a reduction of the well-known disadvantage of dyslexic subjects in RAN after controlling for orthographic processing. They concluded that RAN “is related to reading accuracy (and partly to reading fluency) because children must access orthographic representations from long-term memory. Although accessing these representations is sufficient for accurate word recognition, it is not sufficient for fluent reading, which also requires efficient parafoveal processing” (Liao et al., 2015: 106).

Picture RAN contributes to word, non-word, and text writing, with a main difference. While picture RAN errors only contribute to word and text writing, both speed and errors contribute to non-word writing. Non-word writing might be a more complex task. After

controlling for grade and bilingualism, no significant difference emerged between scores obtained by the sample in word, non-word, and text writing ($p = 0.514$), meaning that children can accomplish the three tasks with the same proficiency but non-word writing requires more underlying processes.

It is worth noting that, in the present study, the effect of RAN errors on different writing difficulties had an increasing trend across time, at least until the middle of the 2nd grade. This might confirm our assumption that RAN is a measure of automaticity of phonological processes (see Study 3), which increases its role more and more until the middle of the 2nd grade. The decrease of the effect size exhibited at the end of the 2nd grade on word and non-word writing might be due to the introduction of more aware strategies, such as the use of the lexical route for word writing, and the use of orthographic chunks for non-word writing. We also found an effect of RAN speed of colors and digits on reading central comprehension. The effect of colors was present until the middle of the 2nd grade, whereas that of digits only in 1st grade. The reason might lie in the fact that, in the early stages of reading acquisition, children rely on text-related operations for achieving comprehension, to later employ more global strategies as their practice increases (see above). Li et al. (2009) analyzed RAN in a group of 37 American children with and without ADHD and found that number and letter pause variability were predictors of reading comprehension, thus supporting the analysis of intraindividual variability in subcomponents of RAN, among which response preparation had relevance to reading comprehension. In our analyses, we considered accuracy and fluency, corresponding to correctly named objects and named objects per second, respectively. In future analyses, we might decompose the latter measure to analyze pause and articulation variability.

The present data suggest that different RAN measures contribute to different portions of literacy (see also Bowers & Swanson, 1991; Cronin & Carver, 1998; Di Filippo et al., 2005).

6.4.3. Effects of the intervention on reading and writing difficulties

This last section explored the effects of different interventions on the subsample who revealed any literacy problems at the end of the project, as compared to standard measures. These children are not to be intended as subjects with a learning disability for at least 2 reasons. The first one is that a more accurate clinical evaluation is needed to get this conclusion. Particular caution is required with bilingual children, as norms on which the binomial dependent variable (with/without deficit) was modeled in this study are standardized on the Italian monolingual population, and are not adequate particularly for those literacy measures in which children showed significantly different performances in Study 1 (e.g., text decoding speed, reading central comprehension, word and text writing). In the present experiment, this variable was controlled by analyzing bilingualism as a random effect in the generalized logit mixed models. The second reason is that testing of literacy measures ended in the first days of March, thus making an assessment in 2nd-grade children rather premature.

Considering these aspects, we explored the effects of a single metaphonological training on 6 children (5 monolinguals + 1 bilingual) that had been reported by teachers for a delay in literacy acquisition and were monitored from the 1st to the 3rd grade (META-1 group). Three of these children had already received a diagnosis of learning disorder, one of them in comorbidity with developmental language disorder and ADHD, and another one in association with mild intellectual disability. Another subsample, which was monitored from the middle of kindergarten to the 2nd grade of primary school, attended double training (in kindergarten and 1st grade, respectively) as described in Study 3:

- 9 children (6 monolinguals + 3 bilinguals in 2nd grade) attended a double metaphonological training (META-2);

- 10 children (6 monolinguals + 4 bilinguals in 2nd grade) attended a double RAN training (RAN group);
- 9 children (7 monolinguals + 2 bilinguals in 2nd grade) attended a double graphic training (GRAPH group);
- 10 children (5 monolinguals + 2 bilingual in 3rd grade and 1 monolingual + 2 bilinguals in 2nd grade) did not attend any training (CG).

Figure 53 shows that the META-2 group does not include any case of decoding difficulty or reading global comprehension problems. They exhibit a small proportion of cases of reading comprehension specific problems, equivalent to those of the RAN and the GRAPH group; a small proportion of cases of word writing problems, which is smaller than that of every other group; a small proportion of cases with non-word writing accuracy problems, which is equivalent to that of the RAN group; and a moderate proportion of cases with text writing problems, which is equivalent to that of the RAN group and smaller than those of other groups.

The analysis of the logit GLMM revealed a significantly higher presence of decoding accuracy, reading comprehension (both central and specific), word writing, and non-word writing difficulties in the META-1 group.

In the META-2 group, conversely, there was a significantly lower presence of decoding accuracy, reading central comprehension, and word writing difficulties.

Furthermore, the RAN group revealed a significantly lower presence of non-word writing difficulties, and the CG a higher presence of a reading specific comprehension difficulty.

Teachers' report of children's difficulties was an effective alarm bell, as 11 out of 13 cases that were reported as problematic (the whole group monitored from 1st grade until 3rd grade) developed a reading or writing difficulty. In the observational questionnaires, 1st-grade teachers evaluated children who attended the META-1 training, in particular, as significantly worse than kindergarteners in oral comprehension ($U = 10.500$, $Z = - 3.650$, $p = 0.000$),

metacognition ($U = 13.500$, $Z = -3.371$, $p = 0.000$), and other cognitive skills ($U = 14.500$, $Z = -3.322$, $p = 0.000$). This encourages the development and use of standardized tools for the early detection of possible learning difficulties, such as the one used in the present study (Terreni et al., 2002). These tools should be promoted since kindergarten, to plan timely interventions for children with difficulties. In this regard, we have seen that a brief training of metaphonological and RAN skills in kindergarten, repeated in primary school, might have a positive effect on the development of reading and writing skills (Table 58). The inefficacy of the 1st-grade single intervention might indicate that metaphonological training programs should be planned timely, i.e. during the development of phonological awareness from the global to the analytical level, possibly before learning literacy, and should be reinforced in the first steps of the acquisition of literacy, which, as we have seen, develops interactively with phonological awareness (see Study 1). On the other hand, children who show some particular delay in the acquisition of pre-literacy and literacy should be treated with more structured interventions. Note that 4 out of 6 META-1 children attended additional rehabilitation programs outside of the project after their families were alerted about their difficulties, so the negative outcomes they obtained in this study encourage the early detection of possible prerequisite difficulties and the planning of structured interventions, which should be timely and tailored on children's specific needs.

The significant absence of non-word writing difficulties in the RAN group shows that the RAN training might help children's phonological skills in writing. RAN does not correlate with non-word writing but RAN speed correlates moderately with non-word decoding speed ($0.395 < \rho_s < 0.455$; $0.013 < p < 0.034$). Moreover, in Study 2, we found a significant longitudinal contribution of RAN in word writing. As we did not find any significant relation between RAN and non-word writing, but moderate to strong relations of RAN with non-word repetition (in kindergarten, $\rho_s = 0.558$, $p = 0.000$) and non-word decoding (see above), the interpretation of RAN as an automaticity measure of phonological processes might be confirmed. In this

vein, the RAN intervention could have helped the automatization of phonological processes involved in the sublexical route of writing (see 4.4.2, Figure 26). The fact that the META-2 intervention, which showed effects on decoding accuracy, reading comprehension, and word writing, did not affect non-word writing, might indicate that phonological awareness per se is not sufficient to increase the kind of automaticity for which RAN is helpful. An integrated intervention might then be useful to work on both constructive processes affected by phonological awareness and automaticity processes affected by RAN.

The significant presence of a reading comprehension weakness in the CG corroborates the positive effect of the META-2 intervention on reading comprehension, showing that improving phonological awareness might have a cascade effect on comprehension skills. The effect is significant for central comprehension, thus showing that the improvement of metalinguistic skills might indirectly act on mechanisms of deep comprehension. These data can be interpreted in light of the Simple View of Reading (Gough & Tunmer, 1986), according to which reading is the product of decoding and language comprehension (see 1.6.2).

As our last analysis, we explored the decoding error types determining the decoding accuracy weakness in 2nd and 3rd grade, respectively, to investigate possible different deficits in the dual-route model of reading (Friedmann & Coltheart, 2018). We found a significant presence of attentional errors in 2nd grade, which were prominent in 3rd grade, too, where a significant presence of stress errors was also found. This means that 2nd graders exhibit a prevalent deficit in the peripheral stage of the reading model, which is involved in the orthographic-visual analysis processes (Friedmann & Coltheart, 2018). Children in 3rd grade, beyond a prevalent attentional deficit, seem to be also affected by a deficit in the lexical route of reading, which could indicate the presence of surface dyslexia. These data, namely the main presence of visual difficulties in 2nd grade and the comorbidity between visual and central difficulties in 3rd grade, encourage a fine-grained analysis of

decoding errors even in a shallow orthography like Italian. Analyses of single cases could allow a more thorough interpretation of the present data. The detection of difficulties in specific segments of the dual-route model of reading might also allow for targeted interventions.

6.5. Interim summary 4

Study 4 longitudinally analyzed possible individual, social, linguistic, and metalinguistic predictors of literacy weaknesses that emerged in 2nd and 3rd graders.

Among social variables, social deprivation was a prominent index of literacy problems, in particular of phonological errors in writing, and, in combination with LI, of word writing accuracy. We encouraged the support of language enhancement interventions since kindergarten in multicultural environments, in particular in socio-economically depressed contexts, where strengthening language abilities might help disentangle situations of atypical development from *communication needs* (Dockrell et al., 2012) due to reduced exposure to the Italian input. We argued that clinician and linguists could share the challenge to implement diagnostic tools for bilingual Italian speaking children.

Among individual characteristics, short-term/working memory plays a central role in every literacy problem. Short-term processing abilities are more important for decoding speed, whereas short-term verbal storage is more important for decoding accuracy. Working memory plays a central role in reading comprehension, thus being able to ensure a good global and specific comprehension of texts. Short-term memory has a significant role in word and non-word writing, with an additional effect of word storage abilities for word writing but not for non-word writing. Visual skills significantly contribute to writing in the first steps of acquisition, namely until the middle of the 1st grade. In particular, visual memory contributes

to word writing, whereas visual discrimination skills of letters marginally contribute to non-word writing.

Bilingualism showed small but significant effects on decoding accuracy, reading comprehension, and text writing. We argued that these effects might interfere with the definition of possible literacy deficits and suggested to assess different aspects of decoding (word and non-word beyond text decoding), comprehension (oral beyond reading comprehension), and writing (word, non-word, and sentence beyond text writing) to get a consistent profile of the child. We also argued that the pressing context of collective dictation might affect children's outcomes, especially in bilinguals.

Among linguistic predictors, the key role of expressive vocabulary emerged, in particular on text decoding speed and accuracy, and word and non-word writing, beyond reading comprehension. An important role of (morpho)syntactic skills on literacy was also confirmed, in particular on word and non-word writing. We hypothesized that morpho-semantic reasons might underlie this effect.

As for global metaphonological skills, recognition of the initial syllable in kindergarten was a good predictor of decoding accuracy, non-word writing, and, in particular, reading comprehension, word writing, and text writing problems.

As for analytical metaphonological skills, different manipulation skills contributed to decoding speed and accuracy, respectively. A model was hypothesized, in which initial-syllable deletion, in combination with phonemic verbal fluency, intervenes on decoding speed, whereas final-syllable deletion intervenes on decoding accuracy. The role of initial-syllable deletion progressively disappears as long as children acquire global strategies and automaticity, and phonemic verbal fluency remains the main predictor of decoding speed and accuracy.

Phoneme blending is interpreted in a multicomponential view of literacy (Norton & Wolf, 2012), where it participates as an essential component in writing. Writing tasks require highly automatized processes, as shown by the strong effect of RAN across time.

A multicomponential framework can also explain different extents of intervention of several abilities, such as phoneme blending and phonemic verbal fluency, which contribute to word writing but not to non-word writing. Phoneme blending contributes to accurate phoneme sequencing, phonemic verbal fluency intervenes in the retrieval of known words, and manipulation of the final syllable, according to our interpretation, contributes to writing accuracy, perhaps as a discriminator of inflected words (see above). The absence of phoneme blending among the predictors of non-word writing suggests that children might possess and use an orthographic lexicon in the middle of the 2nd grade.

In this framework, different categories of RAN participate in different portions of literacy, for both speed and accuracy, as a measure of automaticity of phonological processing. The decomposition of latencies and articulation variabilities in RAN might bring further information to future analyses.

A single metaphonological intervention on 1st graders with learning difficulties did not have any effect, as this group showed a significant presence of decoding accuracy, reading comprehension (both central and specific), word writing, and non-word writing problems in 3rd grade. We suggest preventing learning disorders through the use of tools for the early detection of difficulties, and the administration of structured training interventions tailored to children's specific gaps.

The double metaphonological intervention in kindergarten and 1st grade, conversely, seemed to have effects, as this group showed a significantly lower presence of decoding accuracy, reading central comprehension, and word writing problems. The indirect effect on reading comprehension is corroborated by the fact that the CG developed problems in this area. An unexpected effect on the group undergoing the double RAN training determined a

lower presence of the non-word writing difficulty, which might be attributed to the development of automaticity in the sublexical route of writing. An early integrated intervention tapping both constructive processes affected by phonological awareness and automaticity processes affected by RAN might be advisable.

An analysis of main error categories in word/non-word decoding found a significant presence of attentional errors in 2nd grade, corresponding to an impairment in the peripheral stage of the reading model, and a significant presence of attentional and stress errors in 3rd grade, corresponding to an impairment in the peripheral and lexical route of reading. In 3rd grade, this might indicate comorbidity between attentional and surface dyslexia. We encourage a fine-grained analysis of decoding errors which might detect difficulties to specific segments of the dual-route model of reading, thus allowing ad hoc interventions.

7. General discussion

The present study longitudinally analyzed the development of preliteracy and literacy in a sample of Italian monolingual children as compared to mono-literate bilingual children with minority languages attending the same Comprehensive Institute.

The aims were to investigate how diverse preliteracy skills contribute to the development of the first steps of literacy in a language with shallow orthography like Italian; whether this development can be enhanced through a specific training of some preliteracy skills, namely phonological awareness, RAN, and graphic abilities; and whether such training has effects in case of atypical development.

Study 1 focused on the differences between typically developing monolingual and bilingual children in the acquisition of preliteracy and literacy skills from the last year of kindergarten to the 2nd grade of primary school. Study 2 analyzed how preliteracy skills intervene in the acquisition of literacy in the same sample and same period. Study 3 explored the effects of two cycles (in the last year of kindergarten and 1st grade) of three different interventions on metaphonological, RAN, and graphic skills. Study 4 eventually analyzed the individual, social, and (meta)linguistic predictors of literacy difficulties and the effects of different interventions in a subgroup of children who showed atypical acquisition of some literacy skills during the course of the project. Every study kept a constant focus on the possible differences between monolingual and bilingual children.

Study 1 showed that bilingual children have similar phonological skills as their monolingual fellows but differ in their receptive and productive vocabulary, as well as in syntactic comprehension and production skills. Non-word repetition might then be a possible testing area for both monolinguals and bilinguals, by taking into consideration that marginal differences might be possibly due to higher familiarity of monolinguals with structural properties of the Italian lexicon.

Study 2 showed that non-word repetition significantly contributes to word writing in 1st and 2nd grade. The assessment of possible word writing disorders with tools that are standardized on monolingual children might be biased by bilinguals' reduced familiarity with the Italian language. So, since non-word repetition was accessible for bilinguals as for monolinguals, it might help identify possible problems in the phonological processing of both groups. Note that non-word repetition in kindergarten was not a longitudinal predictor of 1st-grade writing, meaning that it can help define a possible phonological writing deficit in 1st and 2nd grade, but it is not an early kindergarten predictor of subsequent writing performances. In 2nd grade, non-word repetition also predicted two error categories in text dictation, namely stress/geminate errors and minor errors. The contribution to stress geminate errors encourages us to interpret them as a kind of phonological errors. Future analyses will be conducted after separating stress and geminate errors, to disentangle possible differences. The contribution of non-word repetition to minor errors, namely punctuation and capital-letter errors in text writing, was interpreted in the framework of interaction between capacity and componential models of short-term memory (Just & Carpenter, 1992; McCutchen, 1996), in which those errors are generated by an overload of the central executive (Baddeley, 1996; 2007). We hypothesized that the negative effect of non-word repetition, associated with the positive effect of graphic reproduction on minor errors might be due to the condition of cognitive overload determined by a complex activity like dictation. Under this condition, the central executive (Baddeley, 1996; 2007) governing the phonological loop and the visual sketchpad might not be able to fulfill the whole burden of functions related to the visual component, thus determining the graphic errors. As long as phonological and visual processes implied in writing are automatized, we expect a progressive reduction of minor errors and of the contribution of visual and phonological memory.

As for other language skills, we highlighted that the instructional input has a crucial role in the development of children's vocabulary size. The parallel development of monolinguals'

and bilinguals' language skills might not be evident in some complex areas, such as the production of direct object clitic pronouns. No significant improvement was observed in this area by gradually making the elicitation conditions more and more demanding. We encouraged a qualitative analysis of children's productions to detect specific areas of language competence that might need training. With this aim, we highlighted the influence of the elicitation context and developed a clitic production test balancing morphosyntactic traits (Casani, Cardinaletti, & Volpato, in preparation).

Reduced language skills affect children's working memory span, so caution should be paid in the assessment of bilinguals' verbal working memory through the common digit span tasks. Visual memory, as well as phonological memory, are more reliable measures of bilinguals' short-term/working memory because they do not suffer from a linguistic bias. In this regard, no trace of bilingual advantage was found in either visual, phonological, or executive tasks (e.g., RAN and graphic tasks).

Concerning literacy skills, bilinguals reach a comparable level of sublexical decoding as monolinguals in 2nd grade, but keep accuracy problems in decoding words with atypical stress, probably depending on their smaller vocabulary. We encourage multicomponential interventions aiming at strengthening children's vocabulary since kindergarten through social multifaceted activities. Italian language abilities in bilinguals allow them for literal reading comprehension but are not sufficient for a deeper comprehension of inferential concepts, thus highlighting a wider gap as long as texts' demand increases.

In 2nd grade, non-word writing skills are equivalent in bilinguals as in monolinguals but word and text writing are less accurate in bilinguals. Pressure conditions required by the collective dictation might affect children's writing skills. The screening of writing abilities in school should take into account this aspect and privilege either the individual assessment or other forms of writing (e.g., auto-dictation) in bilingual children.

Among global metaphonological skills, syllable segmentation was not affected by bilingualism. In kindergarten, this ability predicts sublexical decoding speed achieved in the middle of the 1st grade. This suggests the possible use of syllable segmentation for the early detection of decoding speed difficulties in bilingual children.

The majority of other metaphonological skills, except for slight differences in bilinguals' acquisition pace, achieve similar levels as in monolinguals between the middle and the end of the 1st grade, except for phonemic verbal fluency, which requires a high level of executive functions and good vocabulary.

A similar trend is observed for RAN, which shows differences between monolinguals and bilinguals in kindergarten administrations, which disappear in the 1st grade. The most accessible RAN measure is that of colors, which does not show any effect of bilingualism. Between the 1st and 2nd grade, both speed and errors in color RAN predict text decoding errors; color RAN errors predict word decoding attentional errors, whereas picture RAN errors predict word decoding errors, in particular surface errors. Furthermore, RAN accuracy besides RAN speed predicts both accuracy and speed in reading. An important trend is eventually the effect of RAN on writing. We can conclude that different measures of RAN, as well as different measures of metaphonological skills, predict different portions of literacy (see also Bowers & Swanson, 1991; Cronin & Carver, 1998; Di Filippo et al., 2005) in Italian. These portions are not completely independent from each other but RAN participates to different extents in reading and writing. The contribution of RAN is not visible in the very early steps of reading acquisition, but it takes over between the 1st and 2nd grade. Norton & Wolf (2012), following Vaessen et al. (2010) and Ziegler et al. (2003), highlight that phonological awareness has an important role in the early steps of reading acquisition, whereas as children reach good levels of accuracy, a shift occurs in which the relationship between RAN and reading becomes much stronger. The moment of this shift is determined by the language orthographic depth and corresponds to the transition from accuracy to

fluency. In Italian, which has a shallow orthography, our longitudinal analyses confirmed a developmental pattern in reading and writing, where RAN fits gradually. In kindergarten, the only predictors of 1st-grade written decoding are syllable segmentation and recognition of word minimal pairs, whereas sublexical decoding accuracy is predicted by syllable blending, meaning that global metaphonology is the only predictor. Word writing, conversely, is predicted only by working memory. Between the 1st and 2nd grade, RAN contributes to different extents, along with working memory, phonological awareness, syntactic abilities, and visual skills, in reading and writing. For this reason, we assumed RAN as a measure of automaticity of phonological processes involved in reading. LaBerge & Samuels (1974) highlight that successful reading depends on both the accuracy and automaticity of multiple cognitive and linguistic processes requiring minimal conscious effort.

RAN might be responsible for this automaticity in reading and, according to the present study, in writing, too. In our view, this is the reason why the role of working memory, which is prominent at the first onset of writing, tends to be progressively replaced by other abilities including RAN, which contribute to different extents in the writing process. In other words, the higher the automaticity, the lesser the working memory effort. In Figure 26, we presented a multicomponential model of word writing between the 1st and the 2nd grade, in which accuracy is the result of executive, working memory, metaphonological, and phonological skills. While in the very early steps of writing acquisition working memory has a prevalent role, executive and (meta)phonological skills take over as long as children's abilities develop at the analytical level. Each ability participates in a specific role for the implementation of the sublexical route of writing (Figure 26). Phonological abilities participate in the acoustic-phonological elaboration of the phonemic string, the temporary allocation to the phonological buffer, and the phonological–orthographic conversion; phoneme blending contributes to the acoustic – phonological conversion and sequential elaboration of the phoneme-grapheme conversion. Working memory (digit span backward) and automaticity (RAN) participate at

every stage as processing and automaticity measures, respectively. Coefficients in Figure 22 give us the measure of the contribution extent of each ability to the achievement of the writing model.

After the end of the 1st grade, syntactic skills contribute to literacy tasks, firstly to reading comprehension, and, in 2nd grade, to decoding accuracy, too. In 2nd grade, metaphonological skills participating in reading and writing are highly manipulatory skills (mainly spoonerism), meaning that more refined abilities replace less analytical ones as long as children acquire them. At the same time, syntactic comprehension (see 6.3.2) contributes to reading accuracy and comprehension, and productive vocabulary to phonological accuracy in text writing.

It is then possible to describe a developmental pattern of literacy predictors in the shallow orthography of Italian.

- 1) In the preliteracy stage (5 y.o.), global discriminatory and classificatory metaphonological skills (syllable segmentation, word minimal-pair recognition) predict the subsequent (1-1) sublexical decoding speed and sublexical decoding accuracy (predicted by syllable blending), whereas working memory is the only responsible for word writing under dictation.
- 2) In the first stage of literacy acquisition, between the 1st and the 2nd grade (6 – 7 y.o.), metaphonological manipulation abilities at the phoneme level (phoneme blending, final syllable deletion), as well as digit RAN speed, are predictive of text decoding speed, text decoding errors (predicted by final syllable deletion), and word dictation accuracy (predicted by phoneme blending); whereas syntactic comprehension predicts reading comprehension.
- 3) Phonemic verbal fluency progressively becomes the main measure of decoding speed (text, words, and non-words) together with other abilities (visual memory for words; digit RAN speed for non-words), whereas final syllable deletion remains a measure of text decoding accuracy, along with color RAN (speed and accuracy), and of word decoding accuracy, along with picture RAN accuracy. Surface errors in word decoding are predicted

by picture RAN accuracy, whereas attentional errors by color RAN errors. Phonemic verbal fluency, as well as syntactic production, also predicts reading comprehension global and central score, whereas visual memory predicts specific reading comprehension.

Non-word writing and phonological errors in text writing are predicted by initial syllable deletion, non-phonological errors by phoneme blending, stress/geminate and minor errors by phonological memory, with the contribution of visual memory for minor errors (see above); omission errors in text writing are predicted by syntactic comprehension.

In 2nd grade, the majority of reading and writing processes (word and non-word decoding speed, text writing phonological errors) are also predicted by a highly manipulatory skill like spoonerism; by syntactic comprehension (word and text decoding, and reading comprehension), whereas phonological errors in text writing are predicted by productive vocabulary.

These longitudinal data give us two main suggestions. Firstly, the great number of abilities involved suggests that reading and writing are multicomponential activities based on a complex range of skills, and that good achievement depends on not only accuracy but also automaticity of all these multiple cognitive and linguistic processes requiring minimal conscious effort (LaBerge & Samuels, 1974; Norton & Wolf, 2012). Another crucial aspect inferable by the great variability of the patterns described above is that there is no “specialized” ability, but a progressive adaptation of linguistic, metalinguistic, and memory skills to the increasingly demanding tasks required by literacy acquisition. Hence, the same reading or writing task might be accomplished by resorting to different (meta)linguistic and memory resources by different subjects, or even by the same subject in different moments. In this behavior, there is probably the key of the compensatory mechanisms used by dyslexic subjects to accomplish reading tasks. In the consuming attempts made during these activities, we might also find the reason why several children with developmental dyslexia can exhibit language deficits that are neither diagnosed nor treated (Guasti, 2013), or

atypical language behaviors that are located in an intermediate position between those of language-impaired and typically developing children (Casani, 2020a).

These reflections suggest that interventions focused on unique abilities might be insufficient to treat the complex processes involved in reading, especially in the presence of a reading impairment.

The 3 different interventions administered to 3 different groups in Study 3 showed effects on tapped abilities, overlapping effects between metaphonology and RAN, between metaphonology and syntactic comprehension, and fallout effects on reading and writing. The effects on treated abilities gave us the measure of the direct effect of the intervention. The overlapping effects between RAN and the most accessible metaphonological skill, namely syllable blending, confirmed our idea of RAN as a measure of automaticity of phonological processes. Moreover, the overlapping effects of metaphonology and syntactic comprehension highlighted the interactive relation existing between language and metalanguage. The graphic training, which was conceived as a placebo, had unexpected fallout on reading and writing, probably due to the effect of visual-spatial skills and executive functions tapped by the graphic and sequential activities. An effect of the graphic training on phonological writing errors was interpreted in light of a complex model of working memory integrating a componential and a capacity approach (see 4.4.4).

The overlapping results confirm the multicomponential nature of reading and writing and encourage integrated interventions tapping language, metalanguage, RAN, and graphic skills, to prepare children to face the multifaceted task of literacy acquisition. Such interventions should be timely encouraged since kindergarten, and possibly reinforced in 1st grade. Study 4 clearly illustrates the reason. In this experiment, a single belated metaphonological intervention on 1st graders with learning difficulties did not have any effect, as this group showed a significant presence of decoding accuracy, reading comprehension (both central and specific), word writing, and non-word writing difficulties in 3rd grade. The

double metaphonological intervention in kindergarten and 1st grade administered in the same study, conversely, seemed to determine a significantly lower presence of decoding accuracy, reading central comprehension, and word writing difficulties in the treated group, beyond a significant presence of reading comprehension difficulties in the CG. An unexpected effect determined a significantly lower presence of the non-word writing difficulty in the RAN group, which might be due to increased automaticity in the sublexical route of writing. This overlap also encourages integrated interventions (language + metalanguage + RAN + graphism) (for good practices in the field of multicomponential interventions in English, see Wolf et al., 2009). At the same time, it suggests that such interventions should be propaedeutic to learning to read, namely administered since kindergarten, and reinforced during the early development of reading and writing skills, as children who attended one single training after showing reading difficulties had no benefit.

8. Conclusion

The present study analyzed the longitudinal development of preliteracy and literacy skills in a sample of Italian monolingual children compared to a sample of mono-literate bilingual children with minority languages attending the same school. Children were monitored from the last year of kindergarten to the 3rd grade of primary school. Four studies were illustrated. Study 1 allowed to describe the typical development of language, metalanguage, RAN, and memory prerequisites of literacy, and the first steps of acquisition of reading and writing in Italian monolingual children and bilingual children with minority mother languages. We found similar phonological skills, but a disadvantage in vocabulary, syntactic skills, phonological awareness, RAN, and verbal working memory of bilinguals.

Study 2 investigated how the above-mentioned abilities intervene in the development of reading and writing in the same sample of typically developing children. Reading and writing revealed their multicomponential nature, in which different abilities contribute to different extents in different developmental times. This allowed to individuate a sort of developmental trend of preliteracy skills in a shallow-orthography language like the Italian one. This consists of (i) a first stage of prominent use of working memory with few global metaphonological skills; (ii) a progressive intervention of different measures of analytical metaphonology; (iii) a progressive intervention of different measures of RAN and phonemic verbal fluency; (iv) the combined use of highly manipulatory skills and general language competence, namely lexicon, and (morpho)syntax. RAN was assumed as a measure of phonological automaticity, intervening in every measure of reading, including comprehension, and writing. Different measures of RAN and phonological awareness predicted different portions of literacy.

Study 3 showed the effects of three different training interventions (metaphonological, RAN, and graphic) administered in two different cycles (in the last year of kindergarten and in the

1st grade of primary school). The META group showed positive effects on several metaphonological measures, whereas the other two groups did not increase significantly the treated abilities. The RAN group revealed a positive effect on syllable blending, thus showing the relationship of RAN with phonology. The META group and, unexpectedly, the GRAPH group, improved significantly several measures of literacy, thus showing that training phonological awareness and visual skills (the graphic training might have also tapped executive functions) has positive effects on reading and writing in a shallow orthography like the Italian.

Study 4 analyzed a small group of children who developed some literacy difficulties during the course of the project. The individual, social, linguistic, metalinguistic, and memory predictors of the difficulties were explored, showing significant effects of social deprivation, with and without language disorder, working memory, vocabulary, metaphonological skills, and RAN on different problems.

A single training administered in 1st grade to children reported for difficulties in the acquisition of literacy did not show any significant effect, as the group developed different literacy problems in 3rd grade. Children who attended the double metaphonological training as in Study 3, instead, had positive effects in several measures of literacy. Children who attended the RAN training showed positive effects on non-word writing, which were attributed to the acquisition of automaticity in the phonological route of writing.

We conclude that training intervention tapping preliteracy skills in children with and without learning difficulties should be planned timely, starting from kindergarten, and reinforced in 1st grade, during early literacy acquisition. Such interventions should be multicomponential, to face the different abilities involved in the multifaceted reading and writing processes.

The present study collected longitudinal data from a shallow orthography language like Italian on a topic that has been mainly treated in English. Writing, which has been neglected so far as compared to reading, was analyzed in its early development. Different measures

of reading and writing were considered. Phonological awareness and RAN were not contrasted as blocks of abilities but were singularly and thoroughly analyzed and discretely related to different literacy skills. Furthermore, the onset of preliteracy and literacy were studied in the bilingual population. One limit of this study lies in the fact that the bilingual sample could not be divided into different subgroups according to their mother tongues. The number variability of participants involved in different experiments is another limit, which was faced through the analysis of generalized mixed models. These models were revealed to be effective in facing a further limit of the present study, namely the small size of subsamples involved in the interventions. The high number of measures analyzed longitudinally, however, allowed us to trace an effective framework of the development of preliteracy and early literacy in the Italian shallow orthography.

Future developments will investigate possible differences related to different language groups and, possibly, the effects of biliteracy in the development of bilinguals' skills.

The effects of specific aspects of (morpho)syntactic competence on literacy will be also explored.

References

- Abutalebi, J., & Green, D.W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20, 242–275.
- Adesope, O.O., Lavin T, Thompson T, Ungerleider C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, Jun, 80 (2), 207–245.
- Al Otaiba, S., Puranik, C.S., Ziolkowski, R.A., & Montgomery, T.M. (2009). Effectiveness of Early Phonological Awareness Interventions for Students With Speech or Language Impairments. *The Journal of Special Education*, 43 (2), 107-128.
- Albuquerque, P.C. (2017). Nominación rápida: la importancia de diferentes dimensiones de lectura y ortografía. *Avances En Psicología Latinoamericana*, 35 (1), 43-60.
- Alegria, J., Pignot, E., & Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory & Cognition*, 10 (5), 451-456.
- Angelelli, P., Judica, A., Spinelli, D., Zoccolotti, P., & Luzzatti, C. (2004). Characteristics of writing disorders in Italian dyslexic children. *Cognitive and Behavioral Neurology* 17, (1), 18-31.
- Angelelli, P., Marinelli C.V., & Burani, C. (2014). The effect of morphology on spelling and reading accuracy: A study on Italian children. *Frontiers in Psychology* 5, 1373.
- Armon-Lotem, S. (2012). Introduction: Bilingual children with SLI – the nature of the problem. *Language and Cognition* 15 (1), 1–4 C.
- Arosio, F., Branchini, C., Barbieri, L., & Guasti, M. T. (2014). Failure to produce direct object clitic pronouns as a clinical marker of SLI in school-aged Italian speaking children. *Clinical Linguistics and Phonetics*, 28 (9), 639-663.

- Arosio, F., Branchini, C., Forgiarini, M., Roncaglione, E., Carravieri, E., Tenca, E., & Guasti, M. T. (2010). SLI children's weakness in morphosyntax and pragmatics. In O. Yukio (Ed.), *The Proceedings of the Tenth Tokyo Conference on Psycholinguistic*. Tokyo, JP: Hituzi Syobo Press.
- Babayigit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive-linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*, 103 (1), 169-189.
- Baddeley A.D. (1979). Working Memory and Reading. In Kolars, P.A., Wrolstad, M.E., & Bouma H. (Eds.) *Processing of Visible Language. Nato Conference Series, 13*. Boston, MA: Springer.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental psychology*, 49A (1), 5–28.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.
- Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford, UK: Oxford University Press.
- Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575-.589.
- Ball, E.W., & Blachman, B.A. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia*, 38, 208-225.
- Ball, E., & Blachman, B. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 26, 49–66.

- Belletti, A., & Guasti, M. T. (2015). The acquisition of pronominal clitics. In A. Belletti & M. T. Guasti (Eds.), *The Acquisition of Italian. Morphosyntax and its interfaces in different modes of acquisition* (pp. 81–128). John Benjamins Publishing Company.
- Bernard, J.B., Calabrèse, A., & Castet, E. (2014). Role of syllable segmentation processes in peripheral word recognition. *Vision Research*, 105, 226-232.
- Bernhardt, B., & Major, E. (2005). Speech, language and literacy skills 3 years later: a follow-up study of early phonological and metaphonological intervention. *International Journal of Language & Communication Disorders*, 40 (1), 1 – 27.
- Berninger, V.W., Abbott, R.D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, 39 (2), 141-163.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge, UK: Cambridge University Press.
- Bialystok, E., & Herman, J. (1999). Does bilingualism matter for early literacy? *Bilingualism: Language and Cognition*, 2 (1), 35-44.
- Bialystok, E., Craik, F.I.M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the simon task. *Psychology and Aging*, 19, 290-303.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems. *Scientific Studies of Reading*, 9 (1), 43-61.
- Bialystok, E., Craik, F.I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, 16 (4), 240–250.
- Biemiller, A. (2005). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research*, 2, pp. 41-51. New York: Guilford Press.

- Bishop, D. V. M. (2009). *TROG-2. Test for Reception of Grammar - Version 2*. Firenze, IT: Giunti.
- Bishop, D.V.M. & Snowling, M. J. (2004). Developmental Dyslexia and Specific Language Impairment: Same or Different? *Psychological Bulletin*, 130, 858–888.
- Bishop, D. V. M., McDonald, D., Bird, S., & Hayiou-Thomas, M. E. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development*, 80, 593–605.
- Bishop, D.V.M., Snowling, M.J., Thompson, P.A., Greenhalgh, T., & CATALISE consortium (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLoS ONE* 11 (7), 1-26.
- Bishop, D.V.M., Snowling, M.J., Thompson, P.A., Greenhalgh, T., & CATALISE consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology & Psychiatry*, 58: 1068-1080.
- Bisiacchi, P.S., Cendron, M., Gugliotta, M., Tressoldi, P.E., & Vio, C. (2005). *BNV 5-11 - Batteria di valutazione neuropsicologica per l'età evolutiva*. Trento, IT: Erickson.
- Blache, A., Parsons, C.L., & Humphreys, J.M. (1981) A Minimal-Word-Pair Model for Teaching the Linguistic Significance of Distinctive Feature Properties, *Journal of Speech and Hearing Disorders*, 46 (3), 81, 291-296.
- Blachman, B. (1994). What we have learned from longitudinal studies of phonological processing and reading and some unanswered question: A response to Torgesen, Wagner, and Rashotte. *Journal of Learning Disabilities*, 27, 287–291.
- Blachman, B. (1997). Early intervention and phonological awareness: A cautionary tale. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 409–430). Mahwah, NJ: Erlbaum.

- Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia*, 16 (4), 300-317.
- Bonifacci, P., Canducci, E., Gravagna, G., & Palladino, P. (2017). English as a foreign language in bilingual language-minority children, children with dyslexia and monolingual typical readers. *Dyslexia* 23 (2), 181-206.
- Bonifacci, P., Montuschi, M., Lami, L., & Snowling, M.J. (2014). Parents of Children with Dyslexia: Cognitive, Emotional and Behavioural Profile. *Dyslexia*, 20, 175-190.
- Borrigan, M., De Bruin, A., Havas, V., De Diego-Balaguer, R., Vulchanova, M., Vulchanov, V., & Duñabeitia, J. (2021). Differences in word learning in children: Bilingualism or linguistic experience? *Applied Psycholinguistics*, 42 (2), 345-366
- Bortolini, U. (1995). Lo sviluppo fonologico. In G. Sabbadini (Ed.), *Manuale di neuropsicologia dell'età evolutiva* (pp. 352 - 356). Bologna, IT: Zanichelli.
- Bortolini, U., & Leonard, L. B. (1996). Phonology and grammatical morphology in specific language impairment: Accounting for individual variation in English and Italian. *Applied Psycholinguistics*, 17, 85–104.
- Bortolini, U., Arfé, B., Caselli, M. C., Degasperi, L., Deevy, P., & Leonard, L.B. (2006). Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language and Communication Disorders*, 41, 695–712.
- Bortolini, U., Caselli, M. C., Deevy, P., & Leonard, L. B. (2002). Specific language impairment in Italian: First steps in the search of a clinical marker. *International Journal of Language and Communication Disorders*, 37, 77–93.
- Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading-disability: Multiple measures of a singular process. *Journal of Experimental Child Psychology*, 51 (2), 195–219

- Bradley, L. & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection, *Nature*, 301, 419-421.
- Bradley, L. (1988). Rhyme recognition and reading and spelling in young children. In R. L. Masland, & M. R. Masland (Eds.), *Pre-school prevention of reading failure* (pp. 143-162). Parkton, MD: York Press Inc.
- Brignola, M., & Perrotta, E. (2010). *Giocare con le parole. Seconda parte*. Trento, IT: Erickson.
- Brignola, M., Perrotta, E., & Tigoli, M.C. (2012). *Occhio alle parole*. Trento, IT: Erickson.
- Brignola, M., Perrotta, E., Tigoli, M.C., & Franceschi, S. (2013). *Tachistoscopio Suite. Strumenti di intervento per le difficoltà di lettura*. Trento, IT: Erickson.
- Bruck, M. (1990). Word recognition skills of adults with a childhood diagnosis of dyslexia. *British Journal of Developmental Educational Psychology*, 54, 31–39.
- Bruck, M., & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language*, 22 (2), 307–324.
- Bruggeman, L., & Cutler, A. (2019). The dynamics of lexical activation and competition in bilinguals' first versus second language. *Psychology*, 1342-1346.
- Bryan, K. (2012). Preliminary Study of the Prevalence of Speech and Language Difficulties in Young Offenders. *International Journal of Language and Communication Disorders* 39, 391-400.
- Burani, C. (2010). Word morphology enhances reading fluency in children with developmental dyslexia. *Lingue e Linguaggio*, 9, 177–198.
- Burani, C., & Arduino, L. S. (2004). Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns. *Brain and Language*, 90 (1-3), 318–325.
- Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpho-lexical reading develop in readers of a shallow orthography?. *Brain and Language*, 81, 568-586.

- Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition* 108 (1), 243-262.
- Burkam, T., & Lee, V.E. (2002). *Inequality at the starting gate. Social background differences in achievement as children begin school*. Economic Policy Institute.
- Burt, L., Holm, A., & Dodd, B. (1999). Phonological Awareness Skills of 4 Year Old British Children: an assessment and developmental data. *International Journal of Language and Communication Disorders*, 34, 311-335.
- Bus, A., & Van Ijzendoorn, M. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91, 403–414.
- Cain, K., Oakhill, J., & Bryant, P. (2004a). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96, 31–42.
- Cain, K., Oakhill, J., & Lemmon, K. (2004b). 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, 671–681.
- Calgaro, G., Toffalini, E., & Cornoldi, C. (2018). La Prova di Lettura Sublessicale – PLS per la valutazione delle abilità di lettura di bambini di prima primaria. *Dislessia*, 15 (2), 209 – 224.
- Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, 13 (1), 61–68.
- Cantalini, M. (1987). *The effects of age and gender on school readiness and school success*. Unpublished doctoral dissertation. Toronto, ON. Institute for Studies in Education.
- Cardinaletti, A., & Casani, E. (2019). Pragmatica nell'uso dei pronomi clitici diretti di terza persona singolare in bambini dislessici con e senza disturbo specifico del linguaggio.

In E. Nuzzo, & I. Vedder (Eds.). *Lingua in contesto. La prospettiva pragmatica* (pp. 9 - 27). Milano: AltLA - Associazione Italiana di Linguistica Applicata.

Cardoso-Martins, C., & Pennington, F.B. (2004). The relationship between phoneme awareness and rapid serial naming skills and literacy acquisition: The role of developmental period and reading ability. *Scientific Studies of Reading*, 8 (1), 27-52.

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264-293). Cambridge, MA: MIT Press.

Cárnio, M.S., Vosgrau, J.S., Soares, A.J.C. (2017). The role of phonological awareness in reading comprehension. *Revista CEFAC*, 19 (5), 590-600.

Casalis, S., Deacon, S. H., & Pacton, S. (2011). How specific is the connection between morphological awareness and spelling? A study of French children. *Applied Psycholinguistics*, 32 (3), 499–511.

Casani, E. (2019). Distinguishing Developmental Dyslexias. Reading errors by Italian dyslexic children with and without SLI. Poster presented at the *DISPOC Conference "Brain, Language, and Learning"*. Università di Siena, September 26 – 28, 2019. Siena, IT.

Casani, E. (2020a). Production of direct object clitic pronouns by Italian children with different acquisition modes. In F.M. Dovetto (Ed.), *I sistemi instabili. Collana "Linguistica delle differenze" n. 5*, pp. 449 - 464. Roma: Aracne.

Casani, E. (2020b). Rapid Naming, Reading, Comprehension and Syntactic Production by Italian Dyslexic Children With and Without Specific Language Impairment. In D. Levey (Ed.), *Strategies and Analyses of Language and Communication in Multilingual and International Contexts* (pp. 111 - 122). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

- Casani, E. (2020c). Distinguishing DD from SLI. Language profiles of Italian dyslexic children with and without specific language impairment. *Pragmalingüística. Investigaciones en lingüística clínica, volumen monográfico 2*, 78 - 96.
- Casani, E. (in preparation). Effects of decoding accuracy variables and (morpho)syntactic competence on the reading of Italian children with and without developmental dyslexia. *Frontiers in Psychology*.
- Casani, E., & Cardinaletti, A. (2019). Soggetto espresso e discordanza di genere soggetto-oggetto. Note per l'implementazione di un test di elicitazione di pronomi clitici diretti. Poster presented at "Giornate AIRIPA di aggiornamento sull'uso degli strumenti in Psicologia Clinica dello Sviluppo", XVI Edizione. 8 – 9 March, 2019. Bologna, IT.
- Casani, E., Cardinaletti, A., & Volpato, F. (in preparation). Test di elicitazione di pronomi clitici diretti di III persona.
- Castles, A. (2006). The dual route model and the developmental dyslexias. *London Review of Education*, 4 (1).
- Castles, A., Holmes, V.M., Neath, J., Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology*, 56 (3): 445-467.
- Castles A, & Coltheart M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, 91 (1): 77-111.
- Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders?. *Journal of Speech, Language, and Hearing Research*, 48, 1378–1396.
- Chall, J. S. (1967). *Learning to read: the great debate*. New York: McGraw-Hill.
- Chall, J. S., Conard, S. S., & Harris-Sharples, S. (1991). *Should textbooks challenge students? The case for easier or harder textbooks*. New York: Teachers College Press, Teachers College, Columbia University.,

- Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chaney, C. (1992). Language development, metalinguistic skills, and print awareness in 3-year-old children. *Applied Psycholinguistics*, 13 (4), 485-514.
- Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: effects of spoken language and orthography. *Cognition*, 81, 227 - 241.
- Chik, P.P.-m., Ho, C.S.-h., Yeung, P.-s., Chan, D.W.-o., Chung, K.K.-h., Luan, H., Lo, L.-y., & Lau, W.S.-y. (2012). Syntactic skills in sentence reading comprehension among Chinese elementary school children. *Reading and Writing: An Interdisciplinary Journal*, 25 (3), 679–699.
- Chondrogianni, V., Marinis, T., Edwards, S., & Blom, E. (2015). Production and on-line comprehension of definite articles and clitic pronouns by Greek sequential bilingual children and monolingual children with specific language impairment. *Applied Psycholinguistics*, 36(5), 1155-1191.
- Cilibrasi, L., Stojanovik, V., & Riddell, P.M. (2014). Lexical and morphosyntactic minimal pairs: Evidence for different processing and implications in language pathology treatment. *PeerJ PrePrints* 2, e197v1.
- Cisero, C. A., & Royer, J. M. (1995). The development and cross-language transfer of phonological awareness. *Contemporary Educational Psychology*, 20 (3), 275–303.
- Clarke, P, Hulme, C, & Snowling, M. (2005). Individual differences in RAN and reading: a response timing analysis. *Journal of Research in Reading*, 28, 73–86.
- Coady, J.A., & Evans, J.L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, 43: 1-40.

- Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, 18 (4), 987–1003.
- Colombo, L., Fudio, S., & Mosna, G. (2009). Phonological and working memory mechanisms involved in written spelling, *European Journal of Cognitive Psychology*, 21 (6), 837-861.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108 (1), 204–256.
- Conrad, N.J., & Levy, B.A. (2011). Training letter and orthographic pattern recognition in children with slow naming speed. *Reading and Writing*, 24 (1), 91-115.
- Cornoldi, C. & Carretti, B. (2016). *Prove MT 3 Clinica*. Firenze: Giunti.
- Cornoldi, C., & Colpo, G. (2011). *Prove di Lettura MT-2 per la Scuola Primaria*. Firenze, IT: Giunti.
- Cornoldi, C., Miato, L., Molin A., & Poli, S. (2009). PRCR-2/2009 - *Prove di Prerequisito per la Diagnosi delle Difficoltà di Lettura e Scrittura*. Firenze, IT: Giunti.
- Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296.
- Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, 94, 94–103.
- Cronin, V., & Carver, P. (1998). Phonological sensitivity, rapid naming, and beginning reading. *Applied Psycholinguistics*, 19, 447–461.
- Cross, M. (1999). Lost for Words. *Child and Family Social Work* 4 (3): 249-57.

- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33, 934–945.
- Cutting, L.E., & Denckla, M.B. (2001). The relationship of rapid serial naming and word reading in normally developing readers: an exploratory model. *Reading and Writing*, 14, 673–705.
- Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422–433.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15 (6), 254-262.
- de Jong, P.F., & Vrieling, L.O. (2004). Rapid automatic naming: easy to measure, hard to improve (quickly). *Annals of Dyslexia* 54, 65–88.
- De Luca, M., Di Filippo, G., Judica, A., Spinelli, D., & Zoccolotti P. (2005). *Test di denominazione rapida e ricerca visiva di colori, figure e numeri*. Roma: Santa Lucia. Neuroscienze e Riabilitazione.
- Decker, S. L., Roberts, A. M., & Englund J. A. (2013). Cognitive predictors of rapid picture naming. *Learning and Individual Differences*, 25, 141–149.
- Defior, S., Alegría, J., Titos, R., & Martos, F. (2008). Using morphology when spelling in a shallow orthographic system: the case of Spanish. *Cognitive Development* 23, 204–215.
- Dehaene, S. (2009). *Reading in the Brain: The Science and Evolution of a Human Invention*. New York, NY: Viking.
- Délage, H., Durrleman, S., & Frauenfelder, U. (2016). Disentangling sources of difficulty associated with the acquisition of accusative clitics in French. *Lingua*, 180, 1-24.

- Denckla, M.B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 10 (2), 186–202.
- Denckla, M. B., & Rudel, R.G. (1976). Rapid 'automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14 (4), 471–479.
- Denes, G., Cipollotti, L., & Zorzi, M. (1996). Dislessie e disgrazie acquisite. In G. Denes & L. Pizzamiglio (Eds), *Manuale di Neuropsicologia*. Bologna, IT: Zanichelli.
- Devescovi, A., Orsolini, M. & Pace, C. (1989). Consapevolezza linguistica nei bambini in eta' prescolare. *Rassegna di Linguistica Applicata*, 1-2, 153-178.
- Di Blasi, F., Savelli, E., Zingale, M., Buono, S., & Di Nuovo, S. (2014). Effects of treatment of reading skills in a child with borderline intellectual functioning. *International Journal on Disability and Human Development*, 13 (1), 25-32.
- Di Blasi, F.D., Buono, S., Cantagallo, C., Di Filippo, G., and Zoccolotti, P. (2019) Reading skills in children with mild to borderline intellectual disability: a cross-sectional study on second to eighth graders. *Journal of Intellectual Disability Research*, 63: 1023– 1040.
- Di Filippo, G., Brizzolara, D., Chilosi, A., De Luca, M., Judica, A., Pecini, C., Spinelli, D., & Zoccolotti, P. (2005). Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). *Child Neuropsychology*, 11 (4), 349-361.
- Ding, Y., Richman, L.C., Yang, L.Y., & Guo, J.P. (2010). Rapid automatized naming and immediate memory functions in Chinese mandarin-speaking elementary readers. *Journal of Learning Disabilities*, 43, 48-61.
- Dockrell, J., Ricketts, J., & Lindsay, G. (2012). Understanding speech, language and communication needs: profiles of need and provision. *Research Report DFE-RR247-BCRP4*. Department for Education, UK.

- Duff, F.J., Reen, G., Plunkett, K., & Nation, K. (2015a). Do infant vocabulary skills predict school-age language and literacy outcomes?. *Journal of Child Psychology and Psychiatry*, 56, 848-856.
- Duff, F.J., Nation, K., Plunkett, K., & Bishop, D. (2015b). Early prediction of language and literacy problems: is 18 months too early? *PeerJ*, 3, e109.
- Dunn L.M., Leota PhD., & Dunn, M. (2000). *Peabody*. Torino, IT: Omega.
- Durgunoğlu, A.Y., Nagy, W.E., & Hancin-Bhatt, B.J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, 85 (3), 453–465.
- Dutto, V. (2014). *Giochiamo con i fonemi. Attività e giochi per il consolidamento delle abilità fono-articolatorie*. Trento, IT: Erickson.
- Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, 21, 392–407.
- Edwards, E. C., Font, G., Baumann, J. F., & Boland, E. B. (2003). Unlocking word meanings: Strategies and guidelines for teaching morphemic and contextual analysis. In J.F. Baumann & E.J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp 159-178). New York, NY: Guilford Press.
- Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Scientific Studies of Reading*, 9, 167–188.
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the national reading panel's meta-analysis. *Review of Educational Research*, 71, 393–447.
- Elhassan, Z., Crewther, S.G., & Bavin, E.L. (2017). The Contribution of Phonological Awareness to Reading Fluency and Its Individual Sub-skills in Readers Aged 9- to 12-years. *Frontiers in Psychology*, 8, 533.

- Elleman, A.M., Lindo, E.J., Morphy, P., & Compton, D.L. (2009) The Impact of Vocabulary Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis, *Journal of Research on Educational Effectiveness*, 2 (1), 1-44.
- Ellis, A. W. (1984). *Reading, Writing and Dyslexia: A Cognitive Analysis*. London, UK: Lawrence Earlbaum.
- Engen, L., & Høien, T. (2002). Phonological skills and reading comprehension. *Reading and Writing* 15, 613–631.
- Farukh, A., & Vulchanova, M. (2014), Predictors of Reading in Urdu: Does Deep Orthography Have an Impact?. *Dyslexia*, 20: 146-166.
- Feldman, L.B., Kostić, A., Gvozdenović, V., O'Connor, P.A., & Moscoso del Prado Martín, F. (2012). Semantic similarity influences early morphological priming in Serbian: A challenge to form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, 19, 668–676.
- Feldman, L.B., O' Connor, P.A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review* 16 (4). 684-691.
- Fernandez-Fein, S., & Baker, L. (1997). Rhyme and alliteration sensitivity and relevant experiences among preschoolers from diverse backgrounds. *Journal of Literacy Research*, 29, 433-459.
- Fleckstein, A., Prévost, P., Tuller, L., Sizaret, E., & Zebib, R. (2018). How to identify SLI in bilingual children: A study on sentence repetition in French. *Language Acquisition*, 25 (1), 85-101.

- Fraser, D., Wong, A., & Wu, J. (1999). Regression analysis, nonlinear or nonnormal: simple and accurate p values from likelihood analysis. *Journal of the American Statistical Association*, 94: 1286-1295.
- Friedmann, N., & Coltheart, M. (2018). Types of developmental dyslexia. In A. Bar-On, & D. Ravid (Eds.), *Handbook of communication disorders: Theoretical, empirical, and applied linguistic perspectives* (pp. 721-751). Berlin, Boston: De Gruyter Mouton.
- Friedmann, N., & Gvion, A. (2003). *TILTAN: Battery for the diagnosis of dyslexias in Hebrew*. Tel Aviv, IL: Tel Aviv University.
- Friedmann, N., & Haddad-Hanna, M. (2014). Types of developmental dyslexia in Arabic. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives. Language and Literacy Series* (pp. 119-152). The Netherlands, NL: Springer.
- Friedmann, N., & Novogrodsky, R. (2008). Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI. In A. Gavarró, & M. João Freitas (Eds.), *Language acquisition and development* (pp. 205-217). Newcastle, UK: Cambridge Scholars Press.
- Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and breakdown in aphasia and developmental language impairment. In C. Boeckx & K.K. Grohmann (Eds.), *The Cambridge Handbook of Bilingualism* (pp. 350-374). Cambridge, UK: Cambridge University Press.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia. Neuropsychological and Cognitive Studies of Phonological Reading* (pp. 301-330). London, UK: Erlbaum.
- Frost, R. (2006). Becoming literate in Hebrew: the grain size hypothesis and Semitic orthographic systems. *Developmental science*, 9 (5), 439.

- Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 13 (1), 104-115.
- Fudge, E.C. (1987). Branching structure within the syllable. *Journal of Linguistics*, 23, 359-377.
- Fugate, M.H. (1997). Letter training and its effect on the development of beginning reading skills. *School Psychology Quarterly*, 12, 170–192.
- Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27 (4), 513-543.
- Gathercole, S.E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In S. Pickering (Ed.), *Working memory and education* (pp. 219 – 240). London, UK: Academic Press.
- Gaulin, C. A., & Campbell, T. F. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and Motor Skills*, 79, 55–64.
- Georgiou, G.K., Parrila, R., & Kirby, J. (2006). Rapid naming speed components and early reading acquisition. *Scientific Studies of Reading*, 10, 199–220.
- Georgiou, G. K., Parrila, R., & Liao, C.-H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing: An Interdisciplinary Journal*, 21, 885-903.
- Georgiou, G. K., Hirvonen, R., Liao, C.-H., Manolitsis, G., Parrila, R., & Nurmi, J.-E. (2011). The role of achievement strategies on literacy acquisition across languages. *Contemporary Educational Psychology*, 36, 130-141.
- Geudens, A. (2006). Phonological Awareness and Learning to Read a First Language: Controversies and New Perspectives. *Lot Occasional Series*, 6, 25-43.

- Gillam, R., Peña, E.D., Bedore, L.M., Bohman, T.M., & Mendez-Perez, A. (2013). Identification of Specific Language Impairment in Bilingual Children: I. Assessment in English. *Journal of Speech, Language, and Hearing Research* 56 (6), 1813–1823.
- Gillon, G., McNeill, B., Scott, A., Denston, A., Wilson, L., Carson, K., & Macfarlane, A.H. (2019). A better start to literacy learning: findings from a teacher-implemented intervention in children's first year at school. *Reading and Writing*, 32, 1989 – 2012.
- Gilmour, J., Hill, B., Place, M., & Skuse, D.H. (2004). Social Communication Deficits in Conduct Disorder: a clinical and community survey. *Journal of Child Psychology & Psychiatry* 45 (5), 967-978.
- Gollan, T.H., Montoya, R.I., Cera, C.M., & Sandoval, T.C. (2008). More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814.
- Gollan, T.H., Montoya, R.I., Fennema-Notestine, C., & Morris, S.K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33, 1220–1234.
- Gollan, T.H., Slattery, T.J., Goldenberg, D., Van Assche, E., Duyck W, & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140: 186–209.
- Gough, P.B., & Tunmer, W.E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7 (1): 6-10.
- Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81
- Grigorenko, E. L. (Ed.) (2008). *Educating individuals with disabilities: IDEIA 2004 and beyond*. Springer Publishing Co.
- Grundy, J.G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, Jul 33 (3), 325–40.

- Guasti, M. T. (2013). Oral skills deficit in children with Developmental Dyslexia. In S. Stavrakaki, M. Lalioti, & P. Konstantinopoulou (Eds.), *Advances in Language Acquisition* (pp. 416-424). Newcastle, UK: Cambridge Scholars Press.
- Guedens, A. (2006). Phonological awareness and learning to read a first Language: controversies and new perspectives. *LOT Occasional Series*, 6, 25–43.
- Haarmann, H. J., Davelaar, E. J., & Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *Journal of Memory and Language*, 48, 320–345.
- Hammill D.D. (1990). On Defining Learning Disabilities: An Emerging Consensus. *Journal of Learning Disabilities*, 23 (2), 74-84.
- Hancock, R., Pugh, K. R., & Hoeft, F. (2017). Neural noise hypothesis of developmental dyslexia. *Trends in Cognitive Sciences*, 21, 434–448.
- Harm, M., & Seidenberg, M.S. (1999). Reading acquisition, phonology, and dyslexia: Insights from a connectionist model. *Psychological Review*, 106, 491-528.
- Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 27 (1), 4-9.
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Hart, B., & Risley, T. (1999). *The social world of children learning to talk*. Baltimore, MD: Paul Brookes Pub.
- Heikkilä, R., Aro, M., Närhi, V., Westerholm, J. & Ahonen, T. (2013). Does Training in Syllable Recognition Improve Reading Speed? A Computer-Based Trial With Poor Readers From Second and Third Grade. *Scientific Studies of Reading*, 17 (6), 398-414.
- Hickok, G, & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8: 393–402.

- Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, 20, 921-950.
- Hsu, L.S.-J., Ip, K. I., Arredondo, M. M., Tardif, T., & Kovelman, I. (2019). Simultaneous acquisition of English and Chinese impacts children's reliance on vocabulary, morphological and phonological awareness for reading in English. *International Journal of Bilingual Education and Bilingualism*, 22 (2), 207-223.
- Huggins, A. W. F., & Bolt, B., & Newman Inc. (1977). *Syntactic Aspects of Reading Comprehension. Technical Report No. 33*. Washington D.C., USA: ERIC Clearinghouse.
- Hulme, C., & Snowling, M. (1988). The classification of children with reading difficulties. *Developmental medicine and child neurology*, 30 (3), 398-402.
- Hulme, C., & Snowling, M. J. (2009). *Developmental Disorders of Language Learning and Cognition*. Hoboken, NJ: Wiley.
- Hulme, C., Bowyer-Crane, C., Carrol, I.J.M., Duff F.J., & Snowling, M.J. (2012). The Causal Role of Phoneme Awareness and Letter-Sound Knowledge in Learning to Read: Combining Intervention Studies With Mediation Analyses. *Psychological Science*, 23 (6), 572-577.
- Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science*, 26, 1877–1886.
- Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, 91, 273–296.
- Hyams, N., & Schaeffer, J. (2008). Clitic and auxiliary omissions in Italian children's participle constructions. In J. Freitas & A. Gavarro (Eds.), *Proceeding of GALA 2007*, Cambridge. Cambridge, UK: University Scholars Press.

- Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 2008; 127:277–288.
- Jalali-Moghadan, N. (2015). *Childhood Bilingualism and Reading Difficulties: Insights from Cognition and Pedagogy*. Örebro, SE: Örebro University Repo 12/15.
- Jenkins, J.R., Fuchs, L.S., van den Broek, P., Espin, C., & Deno, S.L. (2003). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research and Practice*, 18, 237–245.
- Joanisse, M., Manis, F., Keating, P., & Seidenberg, M.S. (2000). Language deficits in dyslexic children: speech perception, phonology and morphology. *Journal of Experimental Child Psychology*, 77, 30–60.
- Johnston, R.S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes: The nature of the relationship. *Reading and Writing*, 8, 217-234.
- Jorm, A. (1983). Specific reading retardation and working memory: A review. *British Journal of Psychology*, 74, 311-342.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149.
- Justi, C.N.G., & Roazzi, A. (2012). A contribuição de variáveis cognitivas para a leitura e a escrita no Português Brasileiro. *Psicologia: Reflexão e Crítica*, 25(3), 605-614.
- Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics*, 20, 303-314.
- Kamhi, A. G., & Catts, H. W. (1986). Toward an understanding of developmental language and reading disorders. *Journal of Speech and Hearing Disorders*, 51, 337–347.
- Kaplan, E., Goddglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia: Lea & Febiger.

- Karpava, S. (2017). Object clitic production by simultaneous Russian–Cypriot Greek bilinguals. *Selected papers on theoretical and applied linguistics from ISTAL 22*, 230–250.
- Katzir, T., Kim, Y.-K., Wolf, M., O'Brien, B., Kennedy, B., Lovett, M., & Morris, R. (2006). Reading fluency: The whole is more than the parts. *Annals of Dyslexia*, 56: 51–82.
- Kauermann, G., & Carroll, R.J. (2001). A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*, 96: 1387-1396.
- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and application* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kenward, M.G., & Roger, J.H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53: 983-997.
- Khodadoust, M., Mohamadi, R., Janani, L., Javadi, Z., & Sadeghi, A. (2019). The effect of phonological awareness on rapid automatized naming. *Medical Journal of The Islamic Republic of Iran*, 33 (1), 1-9.
- Kim, Y.-S., Wagner, R. K., & Foster, E. (2011). Relations Among Oral Reading Fluency, Silent Reading Fluency, and Reading Comprehension: A Latent Variable Study of First-Grade Readers. *Scientific Studies of Reading*, 15 (4), 338-362.
- Kroll, J.F., & Gollan, T.H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In: Goldrick, M., Ferreira, V., & Miozzo, M. (Eds.). *The Oxford Handbook of Language Production* (pp. 165-181). New York, USA: Oxford.
- Kroll, J.F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.

- Kuhl, P. K. (2011). Early Language Learning and Literacy: Neuroscience Implications for Education. *Mind, brain and education: the official journal of the International Mind, Brain, and Education Society*, 5 (3), 128–142.
- Kuhl, P. K., Stevenson, J., Corrigan, N. M., Van Den Bosch, J. J. F., Can, D. D., & Richards, T. (2016). Neuroimaging of the bilingual brain: structural brain correlates of listening and speaking in a second language. *Brain and Language*, 162, 1–9.
- Kuo, L. J., Uchikoshi, Y., Kim, T. J., & Yang, X. (2016). Bilingualism and Phonological Awareness: Re-examining Theories of Cross-Language Transfer and Structural Sensitivity. *Contemporary educational psychology*, 46, 1–9.
- LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Landerl, K., Freudenthaler, H.H., Heene, M., De Jong, P.F., Desrochers, A., Manolitsis, G., Parrila, R., & Georgiou, G.K. (2019). Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency. *Scientific Studies of Reading*, 23 (3), 220-234.
- Layes, S., Lalonde, R., & Rebai, M. (2015). Effectiveness of a Phonological Awareness Training for Arabic Disabled Reading Children: Insights on Metalinguistic Benefits. *Bellaterra Journal of Teaching & Learning Language & Literature*, 8 (4), 24-42.
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32 (1), 69-92.
- Lehtonen, M, Soveri, A, Laine, A, Järvenpää, J, de Bruin, A, & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin Apr.*, 144 (4), 394-425.
- Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

- Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42, 1-22.
- Levelt, W.J.M. (1993). Timing in speech production with special reference to word form encoding. In P. Tallal, A. M. Galaburda, R. R. Llinas, & C. von Euler (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia* (*Annals of the New York Academy of Sciences*, 682 (pp. 283-295). New York: New York Academy of Sciences.
- Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.
- Liao, C.H., Deng, C., Hamilton, J., Lee, C.S., Wei, W., & Georgiou, G.K. (2015). The role of rapid naming in reading development and dyslexia in Chinese. *Journal of Experimental Child Psychology*, 130, 106-122.
- Lieberman, I.Y. (1971). Basic research in speech and lateralization of language: some implications for reading disability. *Annals of Dyslexia*, 21, 71–87.
- Lieberman, I.Y., Shankweiler, D., Liberman, A.M., Fowler, C., & Fischer, F.W. (1977). Phonetic segmentation and recoding in the beginning reader. In A.S. Reber & D. Scarborough (Eds.), *Toward a psychology of reading: The proceedings of the CUNY Conference* (pp. 207–225). Hillsdale: NJ: Lawrence Erlbaum Associates, Inc.
- Locke, A., & Ginsborg, J. (2003). Spoken language in the early years: The cognitive and linguistic development of three- to five-year-old children from socio-economically deprived backgrounds. *Educational and Child Psychology*, 20 (4), 68–79.
- Locke, A., Ginsborg, J., & Peers, I. (2002). Development and Disadvantage: Implications for the early years. *International Journal of Communication Disorders*, (27) 1.
- Lonigan, C.J., Burgess, S.R., Anthony, J.L., & Barker, T.A. (1998). Development of phonological sensitivity in 2- to 5-year-old children. *Journal of Educational Psychology*, 90 (2), 294–311.

- Lukasik, K.M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., & Laine, M. (2018) Bilingualism and working memory performance: Evidence from a large-scale online study. *PLoS ONE* 13 (11), e0205916.
- Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 263-285.
- Lyons, B., & Peters, D. (2000). Applying Skovgaard's modified directed likelihood statistic to mixed linear models. *Journal of Statistical Computation and Simulation*, 65: 225-242.
- MacDonald, M.C., Just, M.A., Carpenter, P.A. 1992. Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology* 24, 56-98.
- Manis, F.R., Doi, L.M., & Bhadha, B. (2000). Naming speed, phonological awareness, and ortographic knowledge in second graders. *Journal of Learning Disabilities*, 33, 325-333.
- Mann, V.A. & Singson, M. (2003). Linking morphological knowledge to English decoding ability: large effects of little suffixes. In Egbert M.H. Assink & Dominiek Sandra (Eds.), *Reading complex words: Cross-language studies*, 1-25. Amsterdam: Kluwer Academic.
- Marcolini, S., Traficante, D., Zoccolotti, P., & Burani, C. (2011). Word frequency modulates morpheme-based reading in poor and skilled Italian readers. *Applied Psycholinguistics*, 32 (3), 513-532.
- Marinelli, C.V., Zoccolotti, P., Romani, C. (2020). The ability to learn new written words is modulated by language orthographic consistency. *PLoS ONE* 15 (2): e0228129.
- Marotta, L., Ronchetti, C., Trasciani, M., & Vicari, S. (2008). *CMF - Valutazione delle competenze metafonologiche*. Trento, IT: Erickson.

- Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In G.E. MacKinnon & T.G. Waller (Eds.), *Reading Research: Advances in theory and practice*, 3. New York: Academic Press.
- Martin, J., Colé, P., Leuwers, C., Casalis, S., Zorman, M., & Sprenger-Charolles, L. (2010). Reading in French-speaking adults with dyslexia. *Annals of Dyslexia*, 60 (2), 238-264.
- Martini, A. (1995). *Le difficoltà di apprendimento della lingua scritta. Criteri di diagnosi e indirizzi di trattamento*. Pisa, IT: Del Cerro.
- Martini, A., Bello, A., & Pecini, C. (2003). Comparsa e sviluppo di abilità metafonologiche in bambini di scuola materna. *Psicologia Clinica dello Sviluppo*, 3, 385-402.
- McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the “specifics” of specific reading disability and specific language impairment. *Journal of Child Psychology Psychiatry*, 41 (7), 869–874.
- McBride-Chang, C., Shu, H., Zhou, A., Wat, C. P., & Wagner, R. K. (2003). Morphological awareness uniquely predicts young children’s Chinese character recognition. *Journal of Educational Psychology*, 95, 743–751.
- McBride-Chang, C., Cho, J. R., Liu, H. Y., Wagner, R. K., Shu, H., Zhou, A., & Muse, A. (2005). Changing Models across Cultures: Associations of Phonological Awareness and Morphological Structure Awareness with Vocabulary and Word Recognition in Second Graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology*, 92, 140-160.
- McBride-Chang, C., Liu, P.D., Wong, T., Wong, A., & Shu, H. (2012). Specific Reading Difficulties in Chinese, English, or Both: Longitudinal Markers of Phonological Awareness, Morphological Awareness, and RAN in Hong Kong Chinese Children. *Journal of Learning Disabilities*, 45 (6), 503-514.
- McCulloch, C.E. (2003). Generalized Linear Mixed Models. *NSF-CBMS Regional Conference Series in Probability and Statistics*, 7: i-viii+1-84.

- McCutchen, D. (1988). "Functional automaticity" in children's writing: A problem of metacognitive control. *Written Communication*, 5, 306–324.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8, 299–325.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35, 13–23.
- Mei, S., Pagni, I., & Vegini, S. (2013). *Alla scoperta delle parole. Attività metafonologiche, semantico-lessicali e narrative per l'avviamento alla letto-scrittura*. Trento, IT: Erickson.
- Meir, N., & Novogrodsky, R. (2019). Prerequisites of Third-Person Pronoun Use in Monolingual and Bilingual Children With Autism and Typical Language Development. *Frontiers in Psychology*, 10, 2289.
- Michael, E.B., & Gollan, T.H. (2005). Being and becoming bilingual. Individual differences and consequences for language production. In J.F. Kroll & A.M.B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389-407). Oxford, UK: Oxford University Press.
- Miller, A.C., Keenan, J.M., Betjemann, R.S., Willcut, E., Pennington, Bruce F., & Olson, Richard, K. (2014). Reading Comprehension in Children with ADHD: Cognitive Underpinnings of the Centrality Deficit. *Journal of Abnormal Child Psychology* volume 41, 473–483 (2013).
- Mitchell, A.M., & Brady, S.A. (2013). The effect of vocabulary knowledge on novel word identification. *Annals of Dyslexia*, 63 (3-4), 201-216.
- Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing. *Scientific Studies of Reading*, 13 (1), 1-25.
- Morais, J. (1991). Phonological Awareness: A Bridge Between Language and Literacy. In D.J. Sawyer & B.J. Fox (Eds.), *Phonological Awareness in Reading. Springer Series in Language and Communication*, 28. New York, NY: Springer.

- Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive*, 7 (5), 415-438.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323-331.
- Morvay, G. (2012). The relationship between syntactic knowledge and reading comprehension in EFL learners. *Studies in Second Language Learning and Teaching* 2 (3), 415–438.
- Müller, B., Richter, T., Karageorgos, P., Krawietz, S., & Ennemoser, M. (2017). Effects of a Syllable-Based Reading Intervention in Poor-Reading Fourth Graders. *Frontiers in Psychology*, 8, 1635.
- Müller, B., Richter, T., & Karageorgos, P. (2020). Syllable-based reading improvement: Effects on word reading and reading comprehension in Grade 2. *Learning and Instruction*, 66, 101304.
- Mumtaz, S., & Humphreys, G. (2001). The effects of bilingualism on learning to read English: evidence from the contrast between Urdu-English bilingual and English monolingual children. *Journal of Research in Reading*, 24, 113-134.
- Muth, C., Bale, S.K.L., Hinde, K., Maninger, N., Mendoza, S.P., & Ferrer, E. (2016). Alternative Models for Small Samples in Psychological Research: Applying Linear Mixed Effects Models and Generalized Estimating Equations to Repeated Measures Data. *Educational and Psychological Measurement*, 76 (1): 64-87.
- Närhi, V., Ahonen, T., Aro, M., Leppäsaari, T., Korhonen, T., Tolvanen, A., & Lyytinen, H. (2005). Rapid serial naming: Relations between different stimuli and neuropsychological factors. *Brain and Language*, 92, 45-57.

- Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27, 342–356.
- Neuhaus, G., Foorman, B.R., Francis, D.J., & Carlson, C.D. (2001). Measures of information processing in rapid automatized naming (RAN) and their relation to reading. *Journal of Experimental Child Psychology*, 78, 359–373.
- Norton, E.S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, 63, 427-452.
- Nouwens, S., Groen, M.A., & Verhoeven, L. (2017). How working memory relates to children's reading comprehension: the importance of domain-specificity in storage and processing. *Read & Write* 30, 105–120.
- Oldfield, R.C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17, 273–281.
- Orsolini, M., Fanari, R., Serra, G., Cioce, R., Rotondi, A., Dassisti, A., & Maronato, C. (2003). Primi progressi nell'apprendimento della lettura: una riconsiderazione del ruolo della consapevolezza fonologica. *Psicologia clinica dello sviluppo*, 3, 403-436.
- Ouellette G, & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, 23 (2), 189–208.
- Paap, K.R., Johnson, H.A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69: 265-78.
- Pape-Neumann, J., Ermingen-Marbach, M.v., Grande, M., Willmes, K, & Heim, S. (2015). The role of phonological awareness in treatments of dyslexic primary school children. *Acta Neurobiologiae Experimentalis*, 75 (1), 80-106.

- Patel, T. K., Snowling, M. J., & De Jong, P. F. (2004). Learning to read in Dutch and English: A crosslinguistic comparison. *Journal of Educational Psychology*, 96 (4), 785-797.
- Patterson, J. (2011). Verbal Fluency. In Kreutzer, J.S., DeLuca, J., & Caplan, B. (Eds) *Encyclopedia of Clinical Neuropsychology*. New York, NY: Springer.
- Pecini, C., Spoglianti, S., Bonetti, S., et al. (2019). Training RAN or reading? A telerehabilitation study on developmental dyslexia. *Dyslexia*, 25, 318–331.
- Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, 144 (1), 48–76.
- Perrotta, E., & Brignola, M. (2000). *Giocare con le parole. Training fonologico per parlare meglio e prepararsi a scrivere*. Trento, IT: Erickson.
- Perrotta, E., & Brignola, M. (2010). *Giocare con le parole. Le rime e i giochi di parole vol.2*. Trento, IT: Erickson.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested modeling and strong inference resting in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 27, 301–333.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, 61, 106–151.
- Perry, C., Ziegler, J.C., & Zorzi, M. (2013). A computational and empirical investigation of graphemes in reading. *Cognitive Science*, 37 (5): 800-828.
- Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science*, 30, 386–395.
- Peterson, C., & Roberts, C. (2003). Like Mother, Like Daughter: Similarities in narrative style. *Developmental Psychology*, 39, 551-562.

- Peterson, R. L., Pennington, B. F., & Olson, R. K. (2012). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition*, 126, 20–38.
- Pierce, L., Chen, J.K., Delcenserie, A., Genesee, F., & Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nature Communications*, 6, 10073.
- Pinker, S. (1997). Foreword. In D. McGuinness (Ed.), *Why Our Children Can't Read and What We Can Do About It: A Scientific Revolution in Reading* (pp. ix–x). New York, NY: Free Press.
- Pinto, M.A. & Candilera, G. (2000). La valutazione del primo sviluppo metalinguistico. Milano, IT: Franco Angeli.
- Plaut, D., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56-115.
- Poarch, G.J., & van Hell, J.G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, 111 (3), 419-438.
- Poulin-Dubois, E., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17 (1), 57-70.
- Raman, N., Mona P.G., Ta, S., Singh, N.C., & Gaonkar, K. (2020). Role of Language in Multilingual Children with Suspected Dyslexia: A Pilot Study. *Journal of Neurology & Neurophysiology* (11) 1, 001-005.
- Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, 136 (2), 630–645.

- Rastle, K., Davis, M.H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, 11 (6), 1090-1098.
- Ricketts, J., Nation, K., & Bishop, D.V.M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies Of Reading*, 11 (3), 235-257.
- Roberts, D., & Walter S.L. (2016). Writing morphophonology, reading lexical tone: Linguistic and experimental evidence in favour of morphographic spelling in Kabiye (Togo). *Writing Systems Research*, 8 (2), 167-186.
- Rondal, J.A., & Comblain, A. (1996). Language in adults with Down syndrome, *Down Syndrome Research and Practice*, 4 (1), 3–14.
- Sage, R. (2005). Communicating with Students Who Have Learning and Behaviour Difficulties: A continuing professional development programme. *Emotional and Behavioural Difficulties* 10 (4), 281-297.
- Santos, F.H., Bueno, O.F.A., & Gathercole, S.E. (2006). Errors in nonword repetition: bridging short- and long-term memory. *Brazilian Journal of Medical and Biological Research*, 39 (3), 371-385.
- Sartori, G., Job, R., & Tressoldi, P. (2007). *DDE-2. Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva-2*. Firenze, IT: Giunti Psychometrics.
- Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tiersley, L. (2005). Relationships among rapid digit naming, phonological processing, motor automaticity and speech perception in poor, average, and good readers and spellers. *Journal of Learning Disabilities*, 38, 12-28.
- Savage, R., Pillay, V., & Melidona, S. (2008). Rapid serial naming is a unique predictor of spelling in children. *Journal of Learning Disabilities*, 41.

- Schaeffer, J. (2000). *The acquisition of direct object scrambling and clitic placement: syntax and pragmatics*. John Benjamins Publishing Company.
- Schuele, C. M., & Boudreau, D. (2008). Phonological awareness intervention: Beyond the basics. *Language, Speech, and Hearing Services in Schools*, 39 (1), 3–20.
- Seidenberg, M.S. (2007). Connectionist models of word reading. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics* (pp. 235–250). New York, NY: Oxford University Press.
- Service, E., Simola, M., Metsänheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *Europe's Journal of Psychology*, 14, 383-408.
- Sevald, C.A., Dell G.S., & Cole, J.S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory & Language*, 34, 807-820.
- Seymour, P.H.K., & MacGregor, C.J. (1984). Developmental dyslexia: A cognitive experimental analysis of phonological, morphemic and visual impairments. *Cognitive Neuropsychology*, 1 (1), 43 - 83.
- Seymour, P.H.K., Aro, M., & Erskine, J.M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94: 143-174.
- Shao, Z., Janse, E., Visser, K., & Meyer, A.S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology* (5), 772.
- Simpson, I.C., Moreno-Pérez, F.J., Rodríguez-Ortiz, I.R., Valdés-Coronel, M., & Saldaña, D. (2020). The effects of morphological and syntactic knowledge on reading comprehension in Spanish speaking children. *Reading and Writing: An Interdisciplinary Journal*, 33, 329–348.
- Siu, C.T.-S., & Ho, C.S.-H. (2020). A longitudinal investigation of syntactic awareness and reading comprehension in Chinese-English bilingual children. *Learning and Instruction*, 67, 101327.

- Snow, P.C., & Powell, M.B. (2008). Oral Language Competence, Social Skills and High-risk Boys: What are Juvenile Offenders Trying to Tell us?. *Children & Society*, 22, 16-28.
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, England: Blackwell
- Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, 7, 37–46.
- Snowling, M.J., Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: creating a virtuous circle. *British Journal of Educational Psychology*, 81 (1): 1-23.
- Sorianello, P. (2012). *Linguaggio e sindrome di Down*. Milano, IT: Franco Angeli.
- Spector, J. (1995). Phonemic awareness training: Application of principles of direct instruction. *Reading and Writing Quarterly*, 11, 37–51.
- Spencer, S., Clegg, J., Stackhouse, J., & Rush, R. (2017). Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *International Journal of Language and Communication Disorders*, 52 (2), 184-196.
- Spinelli, D., Angelelli, P., De Luca, M., Di Pace, E., Judica, A., & Zoccolotti, P. (1997). Developmental surface dyslexia is not associated with deficits in the transient visual system. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 8 (8), 1807–1812.
- Spinnler, H., & Tognoni, G. (1987). Standardizzazione e taratura italiana di test neuropsicologici. Milano, IT: Masson Italia Periodici.
- Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72–110.
- Stainthorp, R., Powell, D., & Stuart, M. (2013). The relationship between rapid naming and word spelling in English. *Journal of Research in Reading*, 36, 371-388.
- Stanovich, K. (1982). Individual differences in the cognitive processes of reading: 1. Word decoding. *Journal of Learning Disabilities*, 15, 485-493.

- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32–71.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable difference model. *Journal of Educational Psychology*, 86 (1), 24-53.
- Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, 20 (4): 147-152.
- Stella, G. (2013). *Pappagallo Lallo. Identificazione precoce di difficoltà fonologiche e laboratori per lo sviluppo linguistico da 3 a 5 anni*. Firenze, IT: Giunti.
- Stella, G., & Apolito, A. (2004). Lo screening precoce nella scuola elementare. Può una prova di sedici parole prevedere i disturbi specifici di apprendimento?. *Dislessia*, 1 (1), 111-120.
- Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, 30 (2): 139-181.
- Stuart-Smith, J., & Martin, D. (1997). Investigating literacy and pre-literacy skills in Panjabi/English schoolchildren. *Educational Review*, 49, 181-197.
- Sullivan, M. D., Poarch, G. J., & Bialystok, E. (2018). Why is Lexical Retrieval Slower for Bilinguals? Evidence from Picture Naming. *Bilingualism (Cambridge, England)*, 21 (3), 479–488.
- Summers, C., Bohman, T.M., Gillam, R.B., Peña, E.D., & Bedore, L.M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders*, 45, 480-493.
- Swanson, H. L., & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of Experimental Child Psychology*, 63 (2), 358 - 385.

- Swanson, H.L. Trainin, G., Necochea, D.M., & Hammill, D.D. (2003). Rapid naming, phonological awareness, and reading: a meta-analysis of the correlation evidence. *Review of Educational Research*, 73, 407–440.
- Talli, I., & Emmanouil, P. (2020). Reading and Non-word Repetition Skills in Bilingual Developmental Dyslexia: The Case of a Greek-Italian Bilingual Dyslexic Adult. *International Journal of Education*, 12 (2), 30-45.
- Talli, I., Sprenger-Charolles, L., & Stavrakaki, S. (2015). Is there an overlap between Specific Language Impairment and Developmental Dyslexia? New insights from French. *Language Acquisition and Language Disorders*, 58, 57-88.
- Terreni, A., Tretti, M. L., Corcella, P. R., Cornoldi, C., & Tressoldi, P. E. (2002). *IPDA - Questionario Osservativo per l'Identificazione Precoce delle Difficoltà di Apprendimento*. Trento, IT: Erickson.
- Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology*, 17 (2): 97-114.
- Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*, 46 (1): 1-16.
- Thordardottir, E., Rioux, E.J. (2019) Does Efficacy Equal Lasting Impact? A Study of Intervention Short Term Gains, Impact on Diagnostic Status, and Association with Background Variables. *Folia Phoniatica et Logopaedica*, 71 (2-3): 71-82.
- Thordardottir, E., Cloutier, G., Ménard, S., Pelland-Blais, E., & Rvachew, S. (2015). Monolingual or bilingual intervention for primary language impairment? A randomized control trial. *Journal of Speech, Language, and Hearing Research*, 58 (2): 287-300.
- Toppelberg, C.O., & Shapiro, T. (2000). Language Disorders: A 10-year research update review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 143-152.

- Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 15* (1), 55–64.
- Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Burgess, S., & Hecht, S. (1997). The contributions of phonological awareness and rapid naming ability to the growth of word reading skills in second-to-fifth-grade children. *Scientific Studies of Reading, 1*, 161–185.
- Traficante, D., Marcolini, S., Luci, A., Zoccolotti, P., & Burani, C. (2011). How do roots and suffixes influence reading of pseudowords: A study of young Italian readers with and without dyslexia. *Language and Cognitive Processes 26* (4), 777–793.
- Traficante, D., Friedmann, N., Rizzi, G., & Luzzatti, C. (2017). Multiple patterns of dyslexia in a shallow orthography: Principles for a diagnostic screening of developmental dyslexias. Poster presented at the *Academy of Aphasia 55th Annual Meeting*, November 5-7, 2017. Baltimore, MD.
- Treiman, R., & Kessler, B. (1995). In defense of an onset-rime syllable structure for English. *Language and Speech, Apr-Jun 38* (2): 127-142.
- Tressoldi, P., Cornoldi, G., & Re, A.M. (2013). BVSCO - 2. Batteria per la Valutazione della Scrittura e della Competenza Ortografica – 2. Firenze, IT: Giunti Psychometrics.
- Troia, G. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34*, 28–52.
- Tuller, L., & Jakubowicz, C. (2004). Développement de la morphosyntaxe du français chez des enfants sourds moyens, *Le langage et l'homme: logopédie, psychologie, audiologie, 14* (2), 191–207.
- Tuller, L., Delage, H., Monjauze, C., Piller, A. G., & Barthez, M. A. (2011). Clitic pronoun production as a measure of atypical language development in French. *Lingua, 121* (3), 423-441.

- Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, 16, 122–140.
- Turkeltaub, P.E., Gareau, L., Flowers, D.L., Zeffiro, T.A., & Eden, G.F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, 6: 767–773.
- Tzouriadou, M. (2020). Concepts and Ambiguities in the Field of Learning Disabilities. In S. Misciagna (Ed.), *Learning Disabilities Neurological Bases, Clinical Features and Strategies of Intervention* (pp. 1-14). London, UK: IntechOpen Limited.
- Ukyo, Y., Noma, H., Maruo, K., & Gosho, M. (2019). Improved Small Sample Inference Methods for a Mixed-Effects Model for Repeated Measures Approach in Incomplete Longitudinal Data Analysis. *Stats*, 2: 174-188.
- Vadasy, P.F., Sanders, E.A., & Herrera, B.L. (2015). Efficacy of Rich Vocabulary Instruction in Fourth- and Fifth-Grade Classrooms. *Journal of Research on Educational Effectiveness*, 8 (3), 325-365.
- Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology*, 105, 213–231.
- Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology*, 102 (4), 827–842.
- Vaessen, A., Gerretsen, P., & Blomert L. (2009). Naming problems do not reflect a second, independent core deficit in dyslexia: double deficits explored. *Journal of Experimental Child Psychology*, 103, 202-221.
- van den Bos, K.P., Zijlstra, J.H., & Spelberg, H.C.I. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading*, 6 (1), 25-49.

- van der Stappen, C., & Van Reybroeck, M. (2018). Phonological Awareness and Rapid Automatized Naming Are Independent Phonological Competencies With Specific Impacts on Word Reading and Spelling: An Intervention Study. *Frontiers in Psychology*, 9, 320.
- van der mosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences, USA*, 107, 10389–10394.
- Varlokosta, S., Belletti, A., Costa, J., Friedmann, N., Gavarró, A., Grohmann, K.K., Guasti, M.T., Laurice Tuller, Lobo, M., Anđelković, D., Argemí, N., Avram, L., Berends, S., Brunetto, V., Delage, H., Ezeizabarrena, M.J., Fattal, I., Haman, E., van Hout, A., Jensen de López, K., Katsos, N., Kologranic, L., Krstić, N., Kuvac Kraljevic, J., Miękisz, A., Nerantzini, M., Queraltó, C., Radic, Z., Ruiz, S., Sauerland, U., Sevcenco, A., Smoczyńska, M., Theodorou, E., van der Lely, H., Veenstra, A., Weston, J., Yachini, M., & Yatsushiro, K. (2016). A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition*, 23 (1), 1-26.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology & Psychiatry*, 45 (1), 2–40.
- Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. New York, NY: Guilford.
- Verhoeven, L., & Perfetti, C. (2003). Introduction to this special issue: The role of morphology in learning to read. *Scientific Studies of Reading*, 7 (3), 209–217.
- Verhoeven, L., & Perfetti, C. A. (2011). Morphological processing in reading acquisition: A crosslinguistic perspective. *Applied Psycholinguistics*, 32 (3), 457–466.

- Vidyasagar, T.R., & Pammer, K., (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing, *Trends in Cognitive Sciences*, 14 (2), 57-63.
- von Bastian, C.C., Souza, A.S., Gade, M. No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*. 2016 Feb; 145(2): 246.
- Vulchanova, M., & Farukh, A. (2018). Can phonological awareness predict concurrent reading outcomes in a deep orthography? *Vigo International Journal of Applied Linguistics*, 15, 145-166.
- Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1999). *Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: Pro-Ed.
- Weikum, W.M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker J.F. (2007). Visual Language Discrimination in Infancy. *Science*, 316, 1159.
- Wells, C. G. (1985). *Language Development in the Pre-School Years*. Cambridge, UK: Cambridge University Press.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23–79.
- Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 11–61). Mahwah, NJ: Erlbaum.
- Wiig, E.H., Zureich, P., & Chan, N.H. (2000). A clinical rationale for assessing rapid automatized naming in children with language disorders. *Journal of Learning Disabilities*, 33, 359-372.
- Willig, A.C. (1985). A Meta-Analysis of Selected Studies on the Effectiveness of Bilingual Education. *Review of Educational Research*, 55 (3), 269-317.

- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, 14, 1–33.
- Wimmer, H., & Hummer, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics*, 11 (4), 349–368.
- Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, 94 (2), 272–277.
- Wolf, M., & Bowers, P.G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91, 415–438.
- Wolf, M., & Denckla, M. B. (2005). The rapid automatized naming and rapid alternating stimulus tests. *Examiner's Manual*. Austin: Pro-Ed.
- Wolf, M., & Obregon, M. (1992). Early naming deficits, developmental dyslexia and a specific deficit hypothesis. *Brain and Language*, 42, 219–247.
- Wolf, M. (2001). *Dyslexia, fluency, and the brain*. York, UK: York Press Inc.
- Wolf, M., Barzillai, M., Gottwald, S., Miller, L., Spencer, K., Norton, E., Lovett, M., & Morris, R. (2009). The RAVE-O intervention: connecting neuroscience to the classroom. *Mind, Brain, and Education*, 3: 84–93.
- Wolf, M., Bowers, P.G., & Biddle, K. (2000). Naming-speed processes, timing, and reading. *Journal of Learning Disabilities*, 33, 387–407.
- Wolf, M., Denckla, M.B. (2005). *RAN/RAS: Rapid Automatized Naming and Rapid Alternating Stimulus Tests*. Austin, TX: Pro-Ed.
- Wolf, M., Goldberg O'Rourke, A., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: an investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing*, 15, 43–72.

- Wolff, U. (2014). RAN as a predictor of reading skills, and vice versa: results from a randomised reading intervention. *Ann. Dyslexia* 64, 151–165.
- Wu, D.H., Martin, R.C., & Damian, M.F. (2002). A Third Route for Reading? Implications from a Case of Phonological Dyslexia. *Neurocase*, 8 (4), 274-295,
- Zachou, A., Partesana, E., Tenca, E., & Guasti, M.T. (2013). Production and comprehension of direct object clitics and definite articles by Italian children with developmental dyslexia. In S. Stavrakaki, M. Lalioti, & P. Konstantinopoulou (Eds.). *Advances in Language Acquisition* (pp. 464-471). Newcastle, UK: Cambridge Scholars Press.
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., & Faísca, L. (2010). Orthographic depth and its impact on universal predictors of reading: a crosslanguage investigation. *Psychological Science*, 21, 551–559.
- Ziegler, J.C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: language-specific or universal? *Journal of Experimental Child Psychology*, 86: 169–93.
- Ziegler, J.C., Perry, C., & Zorzi, M. (2020). Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science*, 29 (3): 293-300.
- Zoccali, A., & Mauro, V. (2006). *Giochi fonologici. Attività di discriminazione uditiva e impostazione articolatoria*. Trento, IT: Erickson.
- Zoccolotti, P., Angelelli, P., Judica, A., & Luzzatti, C. (2005). *I disturbi evolutivi di lettura e scrittura. Manuale per la valutazione*. Roma, IT: Carocci.
- Zoccolotti, P., De Luca, M., Di Pace, E., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme–phoneme correspondence. *Applied Psycholinguistics*, 20 (2), 191–216.
- Zuccheromaglio, C. (1985). Il cammino del bambino verso l'appropriazione della lingua scritta [Child's steps toward writing and reading]. *Età Evolutiva*, 21, 95–104.

Appendixes

Appendix A

Study 1

Complete tables of effects

Dependent variable	Effect	F^a	$df1$	$df2$	p	Predictor	Est.	$StdErr$	t	p	CI	
											Lower	Upper
Non-word repetition	Bilingualism	3.853	1	49	0.055	MONO	0.050	0.025	1.963	0.055	- 0.001	0.101
	Adm. Time	16.246	2	68	0.000	KG-1	- 0.118	0.034	- 3.520	0.001	- 0.186	- 0.050
						KG-2	- 0.060	0.012	- 4.881	0.000	- 0.084	- 0.036
Receptive vocabulary	Bilingualism	22.358	1	48	0.000	MONO	0.301	0.064	4.728	0.000	0.173	0.429
	Adm. Time	16.662	1	50	0.000	KG-1	- 0.276	0.068	- 4.082	0.000	- 0.412	- 0.140
Syntactic comprehension	Bilingualism	14.889	1	52	0.000	MONO	3.565	0.924	3.859	0.000	1.711	5.419
	Adm. Time	14.779	1	52	0.000	KG1	- 3.534	0.919	- 3.844	0.000	- 5.378	- 1.689
Syntactic production (DOcls)	Bilingualism	10.369	1	18	0.005	MONO	0.319	0.099	3.220	0.005	0.110	0.527

^a Satterthwaite approximation

Table A.1. Longitudinal effects of bilingualism and administration time on non-word repetition, receptive vocabulary, syntactic comprehension, and syntactic production

Dependent variable	Effect	F	$df1$	$df2$	p	Predictor	Est.	$StdErr$	t	p	CI	
											Lower	Upper
Productive vocabulary	Bilingu	40.521	1	74	0.000	BI	- 11.453	1.799	- 6.366	0.000	- 15.038	- 7.868
	alism											
	Grade	19.313	1	74	0.000	1-1	- 8.106	1.845	- 4.395	0.000	- 11.782	- 4.431
Syntactic production of DOcl	Bilingu	11.804	1	63	0.001	MONO	0.500	0.145	3.436	0.001	0.209	0.790
	alism											

Table A.2. Significant effects of bilingualism and administration time on productive vocabulary and syntactic production of direct object clitic pronouns in 1st and 2nd grade cross-sectional tests

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
											Lower	Upper
Syllable blending	Bilingualism	4.556	1	78	0.036	MONO	0.072	0.034	2.135	0.036	0.005	0.140
Initial syll recognition	Bilingualism	8.060	1	90	0.006	MONO	0.208	0.073	2.839	0.006	0.062	0.354
Word minimal pair recognition	Bilingualism	29.937	1	87	0.000	MONO	0.188	0.034	5.471	0.000	0.120	0.256
Non-word minimal pair recognition	Bilingualism	17.235	1	83	0.000	MONO	0.179	0.043	4.152	0.000	0.093	0.264
	Adm. Time	4.810	1	29	0.037	KG-1	- 0.149	0.068	- 2.193	0.037	- 0.288	- 0.010
Rhyme recognition	Bilingualism	26.878	1	105	0.000	MONO	0.190	0.037	5.184	0.000	0.118	0.263
	Adm. time	16.038	4	60	0.000	KG-1	- 0.509	0.105	- 4.855	0.000	- 0.724	- 0.293
						KG-2	- 0.315	0.052	- 6.058	0.000	- 0.419	- 0.212
						1-1	- 0.258	0.050	- 5.145	0.000	- 0.357	- 0.158

^a Satterthwaite approximation

Table A.3. Longitudinal effects of bilingualism and administration time on syllable blending, recognition of the first syllable, of word minimal pairs, non-word minimal pairs, and rhymes.

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>P</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
											Lower	Upper
Phoneme blending	Adm. Time	46.321	2	120	0.000	1-1	-	0.062	- 9.625	0.000	- 0.718	- 0.472
							0.595					
						1-2	- 0.086	0.025	- 3.406	0.001	- 0.136	- 0.036
Phoneme segmentation	Adm. Time	23.039	2	75	0.000	1-1	- 0.384	0.057	- 6.788	0.000	- 0.496	- 0.271
						MONO*	3.849	1.040	3.701	0.000	1.772	5.926
						1-1						
Initial syllable deletion	Adm. Time	23.443	2	78	0.000	1-1	- 0.401	0.059	- 6.842	0.000	- 0.518	- 0.284
						1-2	- 0.127	0.053	- 2.373	0.024	- 0.235	- 0.018
						MONO*	3.376	1.117	3.022	0.004	1.115	5.637
						1-2						
Final syllable deletion	Adm. Time	35.341	2	83	0.000	1-1	- 0.493	0.059	- 8.283	0.000	- 0.611	- 0.374
						1-2	- 0.225	0.056	- 4.042	0.000	- 0.338	- 0.112
Phonemic verbal fluency	Bilingualism	9.397	1	26	0.005	MONO	0.338	0.110	3.066	0.005	0.111	0.564
	Adm. time	19.401	2	123	0.000	1-1	- 0.382	0.063	- 6.060	0.000	- 0.507	- 0.257

^a Satterthwaite approximation

Table A.4. Significant longitudinal effects of bilingualism and administration time on phoneme blending and segmentation, initial and final syllable deletion, and phonemic verbal fluency.

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	Lower	Upper
Picture RAN speed	Bilingualism	18.712	1	207	0.000	MONO	0.145	0.033	4.326	0.000	0.079	0.211
	Adm. Time	63.870	4	63	0.000	KG-1	- 0.690	0.067	- 10.284	0.000	- 0.825	- 0.555
						KG-2	- 0.672	0.053	- 12.633	0.000	- 0.778	- 0.566
						1-1	- 0.192	0.048	- 3.956	0.000	- 0.289	- 0.095
Picture RAN errors	Bilingualism	16.444	1	137	0.000	MONO	- 0.401	0.099	- 4.055	0.000	- 0.596	- 0.205
	Adm. time	17.709	4	68	0.000	KG-1	1.191	0.187	6.357	0.000	0.810	1.573
						KG-2	0.876	0.161	5.440	0.000	0.556	1.196
Color RAN speed	Adm. time	9.455	2	88	0.000	1-1	- 0.190	0.044	- 4.330	0.000	- 0.277	- 0.102
						1-2	- 0.094	0.034	- 2.752	0.007	- 0.161	- 0.026
Color RAN errors	Adm. time	4.301	2	49	0.019	1-1	0.452	0.167	2.708	0.010	0.116	0.789
Digit RAN speed	Bilingualism	5.276	1	30	0.029	MONO	0.163	0.071	2.297	0.029	0.018	0.307
	Adm. time	54.593	2	122	0.000	1-1	- 0.578	0.055	- 10.436	0.000	- 0.688	- 0.469
						1-2	- 0.212	0.042	- 4.992	0.000	- 0.296	- 0.128
Digit RAN errors	Bilingualism	5.625	1	64	0.021	MONO	- 0.379	0.160	- 2.372	0.021	- 0.698	- 0.060
	Adm. time	14.142	2	55	0.000	1-1	0.873	0.180	4.850	0.000	0.513	1.232

^a Satterthwaite approximation

Table A.5. Longitudinal effects of bilingualism and administration time on RAN speed and errors for pictures, colors, and digits.

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	Lower	Upper
Digit span forward	Adm. Time	34.065	2	93	0.000	KG-1	- 0.186	0.028	- 6.571	0.000	- 0.242	- 0.129
						1-1	- 0.085	0.031	- 2.717	0.009	- 0.147	- 0.022
Digit span backward	Bilingualism	10.364	1	80	0.002	MONO	0.172	0.053	3.219	0.002	0.066	0.278
	Adm. time	18.671	2	93	0.000	KG-1	- 0.333	0.057	- 5.287	0.000	- 0.447	- 0.220
						1-1	- 0.149	0.059	- 2.526	0.014	- 0.267	- 0.031
Word span	Bilingualism	9.465	1	87	0.003	MONO	0.107	0.035	3.077	0.003	0.038	0.176
	Adm. time	18.330	2	106	0.000	KG-1	- 0.205	0.035	- 5.809	0.000	- 0.276	- 0.135
						1-1	- 0.122	0.041	- 2.971	0.004	- 0.205	- 0.040

^a Satterthwaite approximation

Table A.6. Significant longitudinal effects of bilingualism and administration time on digit span forward, digit span backward and word span.

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
											Lower	Upper
Visual memory errors (one-circle reproduction)	Adm. time	6.460	2	13	0.011	KG-1	0.760	0.255	2.981	0.016	0.181	1.339
Visual memory errors (two-circle reproduction)	Adm. time	11.122	2	70	0.000	KG-1	0.416	0.108	3.861	0.000	0.201	0.630
						KG-2	0.342	0.094	3.650	0.001	0.155	0.529

^a Satterthwaite approximation

Table A.7. Significant longitudinal effects of bilingualism and administration time on digit span forward, digit span backward and word span.

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
											Lower	Upper
Sublexical decoding speed (syll/sec)	Adm. time	79.609	1	92	0.000	1-1	- 0.820	0.092	- 8.922	0.000	- 1.002	- 0.637
Sublexical decoding errors	Bilingualism	6.001	1	76	0.017	MONO	- 0.394	0.161	- 2.450	0.017	- 0.715	- 0.074
	Adm. time	23.284	1	73	0.000	1-1	0.766	0.159	4.825	0.000	0.450	1.083

^a Satterthwaite approximation

Table A.8. Significant longitudinal effects of bilingualism and administration time on sublexical decoding speed and errors.

Dependent Variable	<i>F</i>	<i>Df1</i>	<i>Df2</i>	<i>p</i>	Predictor	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
										Lower	Upper
Word decoding errors	5.783	1	24	0.024	BI	0.048	0.020	2.405	0.024	0.007	0.090
Stress errors	27.132	1	25	0.000	BI	0.976	0.187	5.209	0.000	0.590	1.362

Table A.9. Significant effects of bilingualism on word decoding errors

Dependent variable	Predictor	Coeff.	StdErr	t	p	CI	
						Lower	Upper
Word decoding errors	BI	0.048	0.020	2.405	0.024	0.007	0.090
Stress errors	BI	0.976	0.187	5.209	0.000	0.590	1.362

Table A.10. Significant predictive effects of bilingualism on cross-sectional error scores of word decoding.

Dependent variable	Effect	F ^a	df1	df2	p	Predictor	Est.	StdErr	t	p	CI	
											Lower	Upper
Text decoding speed (syll / sec)	Bilingualism	7.382	1	45	0.009	MONO	0.267	0.098	2.717	0.009	0.069	0.466
	Adm. time	19.333	1	43	0.000	1-2	- 0.481	0.109	- 4.397	0.000	- 0.702	- 0.260
Text decoding errors	Bilingualism	6.896	1	51	0.011	MONO	- 0.493	0.188	- 2.626	0.011	- 0.869	- 0.116
	Adm. time	18.010	1	52	0.000	1-2	0.786	0.185	4.244	0.000	0.415	1.158

^a Satterthwaite approximation

Table A.11. Significant longitudinal effects of bilingualism and administration time on text decoding.

Dependent variable	Effect	F	df1	df2	p	Predictor	Est.	StdErr	t	p	CI	
											Lower	Upper
Reading compr. total score (MT_2)	Bilingualism	19.557 ^a	1	110	0.000	MONO	0.251	0.057	4.422	0.000	0.139	0.364
Reading compr. total score (MT-3-clinical)	Bilingualism	4.694 ^a	1	50	0.035	MONO	0.105	0.048	2.167	0.035	0.008	0.202
Central reading compr. (MT-3-clinical score)	Bilingualism	9829.478 ^b	1	50	0.000	MONO	0.114	0.001	99.144	0.000	0.111	0.116
	Bilingualism* time	4.610	1	50	0.037	MONO* 1-2	- 0.221	0.103	49.742	0.037	- 0.429	- 0.014

^a Satterthwaite approximation

^b Robust model

Table A.12. Significant longitudinal effects of bilingualism and administration time on general and central scores of reading comprehension

Dependent variable	Effect	<i>F</i> ^a	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Est.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
											Lower	Upper
Word writing errors	Bilingualism	17.813	1	96	0.000	MONO	- 0.287	0.068	- 4.221	0.000	- 0.422	- 0.152
	Adm. time	81.919	2	50	0.000	1-1	1.563	0.164	9.359	0.000	1.229	1.897
						1-2	0.789	0.184	4.284	0.000	0.418	1.159

^a Satterthwaite approximation

Table A.13. Significant longitudinal effects of bilingualism and administration time on word writing errors

Dependent variable	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	Predictor	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>p</i>	<i>CI</i>	
										Lower	Upper
Total error score	6.661	1	57	0.012	BI	0.376	0.146	2.581	0.012	0.084	0.668
Phonological errors	6.736	1	53	0.012	BI	0.423	0.163	2.595	0.012	0.096	0.751
Non-phonological errors	5.313	1	55	0.025	BI	0.333	0.145	2.305	0.025	0.044	0.623
Minor errors	5.954	1	50	0.018	BI	0.323	0.132	2.440	0.018	0.057	0.589

Table A.14. Significant effects of bilingualism on text writing errors

Appendix B

Study 2

	<i>Coeff.</i>	Sublexical dec. speed (syll/sec)	Sublexical dec. err.	Word dictation errors	Word dictation syll. err.	Word dictation graph. err.
Digit span backward	<i>Spearman ρ</i>	0.410**	- 0.336**	- 0.517**	- 0.576**	- 0.586**
	<i>p</i>	0,001	0,006	0.000	0.000	0.000
	<i>N</i>	66	66	67	67	67
Word span	<i>Spearman ρ</i>	0.274*	- 0.290*	- 0.154	- 0.062	0.001
	<i>p</i>	0,023	0,016	0,2	0,607	0,994
	<i>N</i>	69	69	71	71	71
Non-word repetition (corr. syll.)	<i>Spearman ρ</i>	0.281*	- 0.402**	- 0.371**	- 0.324**	- 0.296*
	<i>p</i>	0,019	0,001	0,001	0,006	0,012
	<i>N</i>	69	69	71	71	71
Receptive vocabulary	<i>Spearman ρ</i>	0.333	- 0.036	- 0.483*	- 0.430*	- 0.392*
	<i>p</i>	0,089	0,859	0,011	0,025	0,043
	<i>N</i>	27	27	27	27	27
Syntactic comprehension	<i>Spearman ρ</i>	0.409*	- 0.265	- 0.533**	- 0.550**	- 0.494**
	<i>p</i>	0,034	0,181	0,003	0,002	0,008
	<i>N</i>	27	27	28	28	28
Syntactic production of DOcl (Arosio et al., 2014)	<i>Spearman ρ</i>	0.091	- 0.079	- 0.511**	- 0.501*	- 0.544**
	<i>p</i>	0,674	0,712	0,009	0,011	0,005
	<i>N</i>	24	24	25	25	25
Syntactic production of DOcl (Casani & Cardinaletti, 2019)	<i>Spearman ρ</i>	0.069	- 0.164	- 0.373	- 0.334	- 0.437*
	<i>p</i>	0,731	0,413	0,056	0,089	0,023
	<i>N</i>	27	27	27	27	27
Syllable blending	<i>Spearman ρ</i>	0.623**	- 0.569**	- 0.282*	- 0.238*	- 0.165
	<i>p</i>	0	0	0,019	0,049	0,175
	<i>N</i>	69	69	69	69	69
Syllable segmentation	<i>Spearman ρ</i>	0.423**	- 0.396**	- 0.234	- 0.188	- 0.134
	<i>p</i>	0	0,001	0,053	0,123	0,274
	<i>N</i>	69	69	69	69	69
Word minimal pairs	<i>Spearman ρ</i>	0.441**	- 0.464**	- 0.295*	- 0.297*	- 0.240*
	<i>p</i>	0	0	0,014	0,013	0,047
	<i>N</i>	69	69	69	69	69
Non-word minimal pairs	<i>Spearman ρ</i>	0.432**	- 0.435**	- 0.447**	- 0.422**	- 0.349**
	<i>p</i>	0	0	0	0	0,003
	<i>N</i>	69	69	69	69	69
Rhyme recognition	<i>Spearman ρ</i>	0.400**	- 0.367**	- 0.322**	- 0.299*	- 0.242*
	<i>p</i>	0,001	0,002	0,006	0,013	0,042

	<i>N</i>	69	69	71	69	71
Initial syllable recognition	<i>Spearman ρ</i>	0.565**	- 0.488**	- 0.442**	- 0.434**	- 0.390**
	<i>p</i>	0	0	0	0	0,001
	<i>N</i>	69	69	69	69	69
Graphic reproduction of 1 circle (err.)	<i>Spearman ρ</i>	- 0.394*	0.27	0.343	0.384*	0.406*
	<i>p</i>	0,042	0,173	0,08	0,048	0,035
	<i>N</i>	27	27	27	27	27
Graphic reproduction of 2 circles (err.)	<i>Spearman ρ</i>	- 0.331	0.266	0.350	0.286	0.217
	<i>p</i>	0,092	0,179	0,074	0,149	0,278
	<i>N</i>	27	27	27	27	27
Visual recognition of letters (err.)	<i>Spearman ρ</i>	- 0.602**	0.348	0.529**	0.458*	0.297
	<i>p</i>	0,001	0,076	0,005	0,016	0,133
	<i>N</i>	27	27	27	27	27
Picture RAN speed (pict/sec)	<i>Spearman ρ</i>	0.392**	- 0.357**	- 0.536**	- 0.553**	- 0.531**
	<i>p</i>	0,001	0,003	0	0	0
	<i>N</i>	67	67	68	68	68
Picture RAN errors	<i>Spearman ρ</i>	- 0.112	0.281*	0.320**	0.342**	0.351**
	<i>p</i>	0,366	0,021	0,008	0,004	0,003
	<i>N</i>	67	67	68	68	68

* Correlation is significant at 0.05

** Correlation is significant at 0.01

Table B.1. Bivariate relations of short-term memory, language, metaphonology, and RAN at the end of KG with early literacy skills in the middle of the 1st grade of primary school.

		Sublex. dec. speed (syll/sec)	Sublex dec. err.	Word dictation err.	Word dictation syll. err.	Word dictation graph. err.
Digit span forward	<i>Spearman ρ</i>	0.219	- 0.200	- 0.199	- 0.156	- 0.096
	<i>p</i>	0.070	0.099	0.091	0.195	0.418
	<i>N</i>	69	69	73	71	73
Digit span backward	<i>Spearman ρ</i>	0.297*	- 0.308**	- 0.330**	- 0.264*	- 0.213
	<i>p</i>	0.013	0.010	0.004	0.026	0.071
	<i>N</i>	69	69	73	71	73
Word span	<i>Spearman ρ</i>	0.261*	- 0.220	- 0.131	- 0.038	0.006
	<i>p</i>	0.030	0.069	0.269	0.756	0.957
	<i>N</i>	69	69	73	71	73
Non-word repetition	<i>Spearman ρ</i>	0.261*	- 0.393**	- 0.345**	- 0.324**	- 0.309**
	<i>p</i>	0.030	0.001	0.003	0.006	0.009
	<i>N</i>	69	69	71	71	71
Receptive vocabulary	<i>Spearman ρ</i>	0.282	- 0.320*	- 0.491**	- 0.457**	- 0.402**

	<i>p</i>	0.070	0.039	0.000	0.001	0.005
	N	42	42	48	46	48
Productive vocabulary	<i>Spearman</i>					
	<i>ρ</i>	0.363*	- 0.404**	- 0.540**	- 0.537**	- 0.518**
	<i>p</i>	0.018	0.008	0.000	0.000	0.000
	N	42	42	46	46	46
Phoneme blending	<i>Spearman</i>					
	<i>ρ</i>	0.461**	- 0.675**	- 0.587**	- 0.516**	- 0.454**
	<i>p</i>	0.000	0.000	0.000	0.000	0.000
	N	69	69	69	69	69
Initial syll. deletion	<i>Spearman</i>					
	<i>ρ</i>	0.555**	- 0.669**	- 0.572**	- 0.501**	- 0.371**
	<i>p</i>	0.000	0.000	0.000	0.000	0.002
	N	69	69	69	69	69
Final syllable deletion	<i>Spearman</i>					
	<i>ρ</i>	0.523**	- 0.577**	- 0.480**	- 0.377**	- 0.273*
	<i>p</i>	0.000	0.000	0.000	0.001	0.023
	N	69	69	69	69	69
Rhyme recognition	<i>Spearman</i>					
	<i>ρ</i>	0.389**	- 0.382**	- 0.478**	- 0.458**	- 0.424**
	<i>p</i>	0.001	0.001	0.000	0.000	0.000
	N	69	69	69	69	69
Phoneme segmentation	<i>Spearman</i>					
	<i>ρ</i>	0.394**	- 0.590**	- 0.606**	- 0.548**	- 0.469**
	<i>p</i>	0.001	0.000	0.000	0.000	0.000
	N	68	68	68	68	68
Phonemic verbal fluency	<i>Spearman</i>					
	<i>ρ</i>	0.503**	- 0.382**	- 0.351**	- 0.368**	- 0.338**
	<i>p</i>	0.000	0.001	0.003	0.002	0.004
	N	69	69	69	69	69
Visual recognition of 1 circle	<i>Spearman</i>					
	<i>ρ</i>	0.000	0.091	- 0.175	- 0.300	- 0.291
	<i>p</i>	10.000	0.651	0.383	0.128	0.141
	N	27	27	27	27	27
Visual recognition of 2 circles	<i>Spearman</i>					
	<i>ρ</i>	- 0.461*	0.419*	0.350	0.262	0.304
	<i>p</i>	0.016	0.030	0.074	0.187	0.124
	N	27	27	27	27	27
Visual recognition of 3 circles	<i>Spearman</i>					
	<i>ρ</i>	- 0.438*	0.396*	0.352	0.293	0.204
	<i>p</i>	0.022	0.041	0.072	0.138	0.309
	N	27	27	27	27	27
Visual recognition of letters	<i>Spearman</i>					
	<i>ρ</i>	- 0.177	- 0.070	0.377	0.349	0.325
	<i>p</i>	0.377	0.728	0.052	0.074	0.098
	N	27	27	27	27	27
Color RAN speed (col/sec)	<i>Spearman</i>					
	<i>ρ</i>	0.617**	- 0.483**	- 0.532**	- 0.564**	- 0.497**

	<i>p</i>	0.000	0.000	0.000	0.000	0.000
	N	69	69	70	70	70
Picture RAN speed (pict/sec)	<i>Spearman</i>					
	<i>ρ</i>	0.648**	- 0.477**	- 0.659**	- 0.683**	- 0.614**
	<i>p</i>	0.000	0.000	0.000	0.000	0.000
	N	69	69	70	70	70
Digit RAN speed (dig/sec)	<i>Spearman</i>					
	<i>ρ</i>	0.729**	- 0.599**	- 0.590**	- 0.624**	- 0.583**
	<i>p</i>	0.000	0.000	0.000	0.000	0.000
	N	68	68	69	69	69
Color RAN errors	<i>Spearman</i>					
	<i>ρ</i>	- 0.131	0.240*	0.112	0.108	0.097
	<i>p</i>	0.285	0.047	0.354	0.373	0.427
	N	69	69	70	70	70
Picture RAN errors	<i>Spearman</i>					
	<i>ρ</i>	- 0.189	0.360**	0.272*	0.263*	0.227
	<i>p</i>	0.121	0.002	0.023	0.028	0.059
	N	69	69	70	70	70
Digit RAN errors	<i>Spearman</i>					
	<i>ρ</i>	- 0.288*	0.349**	0.254*	0.243*	0.214
	<i>p</i>	0.018	0.004	0.038	0.047	0.082
	N	67	67	67	67	67

* Correlation is significant at 0.05

** Correlation is significant at 0.01

Table B.2. Bivariate relations of short-term memory, language, metaphonology, and RAN with early literacy skills in the middle of the 1st grade of primary school.

Variable 1	Variable 2		Coeff.	<i>p</i>	N
Text decoding speed (syll/sec)	Word span	<i>ρ</i> ^a	0.417*	0.025	29
	Receptive vocabulary	<i>ρ</i>	0.408*	0.039	26
	Productive vocabulary	<i>ρ</i>	0.405*	0.040	26
	Syntactic compr	<i>ρ</i>	0.414*	0.036	26
	Phoneme blending	<i>ρ</i>	0.478**	0.009	29
	Phonemic fluency	<i>ρ</i>	0.490**	0.007	29
	Spoonerism	<i>ρ</i>	0.574**	0.002	26
	Col RAN speed (col/sec)	<i>r</i> ^b	0.404*	0.003	29
	Picture RAN speed (pict/sec)	<i>ρ</i>	0.410*	0.027	29
	Color RAN err	<i>ρ</i>	- 0.559**	0.002	29
	Picture RAN err	<i>ρ</i>	- 0.374*	0.046	29
Text decoding errors	Word span	<i>ρ</i>	- 0.546**	0.002	29
	Phonemic fluency	<i>ρ</i>	- 0.445*	0.016	29
	Spoonerism	<i>ρ</i>	- 0.409*	0.038	26
	Color RAN err	<i>ρ</i>	0.402*	0.030	29
Word decoding speed	Word span	<i>ρ</i>	0.460*	0.012	29
	Phoneme blending	<i>ρ</i>	0.411*	0.027	29

	Rhyme recognition	ϱ	0.388*	0.038	29
	Phonemic fluency	ϱ	0.458*	0.013	29
	Spoonerism	ϱ	0.545**	0.004	26
	Col RAN speed (col/sec)	r	0.421*	0.023	29
	Picture RAN speed (pict/sec)	ϱ	0.367*	0.050	29
	Color RAN err	ϱ	- 0.543**	0.002	29
Word decoding errors	Word span	ϱ	- 0.422*	0.022	29
	Syntactic compr	ϱ	- 0.472*	0.015	26
	Phonemic fluency	ϱ	- 0.579**	0.001	29
	Spoonerism	ϱ	- 0.512**	0.007	26
	Col RAN speed (col/sec)	r	- 0.404*	0.03	29
	Picture RAN err	ϱ	0.392*	0.035	29
Word decoding stress errors	Word span	ϱ	- 0.549**	0.002	29
	Receptive vocabulary	ϱ	- 0.459*	0.018	26
	Syntactic compr	ϱ	- 0.557**	0.003	26
	Phonemic fluency	ϱ	- 0.405*	0.029	29
	Picture RAN err	ϱ	0.465*	0.011	29
Non-word decoding speed (syll/sec)	Phonemic fluency	ϱ	0.510**	0.005	29
	Spoonerism	ϱ	0.554**	0.003	26
	Col RAN speed (col/sec)	r	0.455*	0.013	29
	Picture RAN speed (pict/sec)	ϱ	0.403*	0.03	29
	Digit RAN speed (dig/sec)	ϱ	0.395*	0.034	29
Non-word decoding errors	Word span	ϱ	- 0.485**	0.008	29
	Phoneme blending	ϱ	- 0.463*	0.011	29
	Phonemic fluency	ϱ	- 0.384*	0.040	29
	Spoonerism	ϱ	- 0.554**	0.003	26
Reading compr. total score	Word span	ϱ	0.582**	0.002	26
	Receptive vocabulary	ϱ	0.757**	0.000	26
	Productive vocabulary	ϱ	0.634**	0.000	26
	Syntactic compr	ϱ	0.730**	0.000	26
	Rhyme recognition	ϱ	0.391*	0.048	26
	Phonemic fluency	ϱ	0.433*	0.027	26
	Col RAN speed (col/sec)	r	0.465*	0.017	26
	Picture RAN speed (pict/sec)	ϱ	0.450*	0.021	26
	Digit RAN speed (dig/sec)	ϱ	0.412*	0.037	26
	Picture RAN err	ϱ	- 0.503**	0.009	26
Reading compr. central score	Word span	ϱ	0.601**	0.001	26
	Receptive vocabulary	ϱ	0.685**	0.000	26
	Productive vocabulary	ϱ	0.692**	0.000	26
	Syntactic compr	ϱ	0.769**	0.000	26
	Phonemic fluency	ϱ	0.549**	0.004	26
	Spoonerism	ϱ	0.446*	0.023	26
	Col RAN speed (col/sec)	r	0.458*	0.019	26
	Picture RAN speed (pict/sec)	ϱ	0.419*	0.033	26
	Color RAN err	ϱ	- 0.391*	0.048	26
	Picture RAN err	ϱ	- 0.521**	0.006	26
Reading compr. specific score	Word span	ϱ	0.400*	0.043	26

Receptive vocabulary	ρ	0.574**	0.002	26
Productive vocabulary	ρ	0.450*	0.021	26
Syntactic compr	ρ	0.610**	0.001	26
Rhyme recognition	ρ	0.460*	0.018	26
Col RAN speed (col/sec)	r	0.392*	0.048	26
Picture RAN speed (pict/sec)	ρ	0.409*	0.038	26
Digit RAN speed (dig/sec)	ρ	0.439*	0.025	26
Picture RAN err	ρ	- 0.383	0.054	26

^a Spearman's ρ

^b Pearson's r

* Correlation is significant at 0.05

** Correlation is significant at 0.01

Table B.3. Bivariate relations between pre-literacy skills and reading skills in the middle of the 2nd grade

Variable 1	Variable 2		Coeff.	p	N
Word dict err	Word span	ρ^a	- 0.376*	0.044	29
	Syntactic compr	ρ	- 0.389*	0.049	26
	Phoneme blending	ρ	- 0.609**	0.000	29
	Rhyme recognition	ρ	- 0.489**	0.007	29
	Phonemic fluency	ρ	- 0.514**	0.004	29
	Spoonerism	ρ	- 0.635**	0.000	26
Non-word dict err	Phoneme blending	ρ	- 0.364	0.052	29
	Rhyme recognition	ρ	- 0.413*	0.026	29
	Spoonerism	ρ	- 0.387	0.050	26
Text dict tot err	Syntactic compr	ρ	- 0.442*	0.024	26
	Phoneme blending	ρ	- 0.608**	0.001	26
	Phoneme segmentation	ρ	- 0.432*	0.028	26
	Final syll deletion	ρ	- 0.449*	0.021	26
	Spoonerism	ρ	- 0.472*	0.015	26
	Col RAN speed (col/sec)	ρ	- 0.394*	0.046	26
	Color RAN err	ρ	0.453*	0.020	26
Text dict phonol err	Receptive vocabulary	ρ	- 0.383	0.053	26
	Productive vocabulary	ρ	- 0.448*	0.022	26
	Syntactic compr	ρ	- 0.503**	0.009	26
	Phoneme blending	ρ	- 0.557**	0.003	26
	Rhyme recognition	ρ	- 0.440*	0.025	26
	Phoneme segmentation	ρ	- 0.525**	0.006	26
	Final syll deletion	ρ	- 0.525**	0.006	26
	Spoonerism	ρ	- 0.425*	0.030	26
	Col RAN speed (col/sec)	ρ	- 0.562**	0.003	26
	Picture RAN speed (pict/sec)	ρ	- 0.420*	0.033	26

	Color RAN err	ρ	0.454*	0.020	26
Text dict non-phonol err	Dig span backward	ρ	- 0.409*	0.034	27
	Phoneme blending	ρ	- 0.556**	0.003	26
	Phoneme				
	segmentation	ρ	- 0.390*	0.049	26
	Spoonerism	ρ	- 0.407*	0.039	26
	Color RAN err	ρ	0.387	0.051	26
Text dict stress/geminate err.	Phoneme blending	ρ	- 0.476*	0.014	26
Text dictation omissions	Receptive vocabulary	ρ	- 0.466*	0.016	26
	Syntactic compr	ρ	- 0.432*	0.027	26
Text dict minor err	Receptive vocabulary	ρ	- 0.436*	0.026	26
	Syntactic compr	ρ	- 0.407*	0.039	26
	Picture RAN err	ρ	0.409*	0.038	26
Text dict inappropriate words	Receptive vocabulary	ρ	- 0.569**	0.002	26
	Syntactic compr	ρ	- 0.405*	0.040	26

^a Spearman's ρ

Table B.4. Bivariate relations between pre-literacy skills and writing skills in the middle of the 2nd grade

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>StdErr.</i>	<i>t</i>	<i>CI</i>	
									Lower	Upper
Sublexical decoding speed (syll / sec)	Syllable segmentation	12.650	1	53	0.001	0.056	0.016	3.557	0.025	0.088
	Word minimal pair recognition	6.105	1	53	0.017	0.083	0.033	2.741	0.016	0.150
Sublexical decoding errors	Syllable blending	4.416	1	60	0.040		0.053	- 2.101	- 0.216	- 0.005
Word dictation errors (word errors)	Digit span backward	7.273	1	51	0.009	- 0.203	0.075	- 2.697	- 0.355	- 0.052
Word dictation errors (syllable errors)	Digit span backward	8.218	1	52	0.006	- 0.267	0.093	- 2.867	- 0.453	- 0.080
Word dictation errors (grapheme errors)	Digit span backward	13.043	1	53	0.001	- 0.379	0.105	- 3.612	- 0.589	- 0.168

Table B.5 Significant effects of short-term memory, metaphonology, and RAN skills at the end of KG on sublexical decoding and word writing in the middle of the 1st grade of primary school.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>Std. err.</i>	<i>t</i>	<i>CI</i>	
									Lower	Upper
Word dictation errors (word errors)	Digit span backward	7.151	1	66	0.009	- 0.352	0.132	- 2.674	- 0.615	- 0.089
	Non-word repetition	4.788	1	66	0.032	- 0.049	0.022	- 2.188	- 0.094	- 0.004
	Phoneme blending	25.326	1	115	0.000	- 0.150	0.030	- 5.024	- 0.209	- 0.091
	Digit RAN speed	20.019	1	111	0.000	- 1.259	0.281	- 4.474	- 1.817	- 0.701

Table B.6. Significant longitudinal effects of short-term memory measures, analytical metaphonological skills, and RAN on word dictation between the middle of the 1st grade and the middle of the 2nd grade of primary school.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Est.</i>	<i>Std. err.</i>	<i>t</i>	<i>CI</i>	
									Lower	Upper
Text decoding speed (syll/sec)	Phoneme blending	10.531	1	42	0.002	0.089	0.027	3.245	0.034	0.144
	Final syllable deletion	10.147	1	42	0.003	0.069	0.022	3.185	0.025	0.113
	Digit RAN speed	7.208	1	42	0.010	0.581	0.216	2.685	0.144	1.018
Text decoding errors	Final syllable deletion	9.413	1	42	0.004	- 0.140	0.046	- 3.068	- 0.232	- 0.048
Reading comprehension total score	Syntactic comprehension	6.985	1	42	0.012	0.020	0.008	2.643	0.005	0.035
Reading comprehension central score	Syntactic comprehension	4.395	1	44	0.018	0.009	2.096	0.042	0.001	0.035
Reading comprehension specific score	Syntactic comprehension	8.242	1	44	0.006	0.023	0.008	2.871	0.007	0.039
Reading comprehension specific score	Initial syllable deletion	5.297	1	44	0.026	0.030	0.013	2.301	0.004	0.056
Reading comprehension specific score	Digit RAN speed	6.087	1	44	0.018	0.250	0.101	2.467	0.046	0.453
Word dictation errors	Phoneme blending	6.416	1	42	0.015	- 0.155	0.061	- 2.533	- 0.279	- 0.032

Table B.7. Significant longitudinal effects of memory measures, analytical metaphonological skills, and RAN on text decoding, text comprehension, and word dictation between the end of the 1st grade and the middle of the 2nd grade of primary school.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>Std. err.</i>	<i>t</i>	<i>CI</i>	
									Lower	Upper
Text decoding speed (syll/sec)	Phonemic verbal fluency	5.834	1	60	0.019	0.010	0.004	2.415	0.002	0.019
Text decoding errors	Final syllable deletion	4.361	1	75	0.040	- 0.031	0.015	- 2.088	- 0.061	- 0.001
Text decoding errors	Color RAN speed	6.868	1	75	0.011	- 0.931	0.355	- 2.621	- 1.638	- 0.223
Text decoding errors	Color RAN errors	4.558	1	75	0.036	6.900	3.232	2.135	0.461	13.339
Word decoding speed	Phonemic verbal fluency	14.172	1	62	0.000	0.014	0.004	3.765	0.007	0.022
Word decoding speed	Graphic reproduction (2 circles) errors	4.739	1	62	0.033	- 0.034	0.016	- 2.177	- 0.065	- 0.003
Word decoding errors	Phonemic verbal fluency	4.160	1	76	0.045	- 0.014	0.007	- 2.040	- 0.027	- 0.000

Word decoding errors	Picture RAN errors	5.537	1	76	0.021	6.086	2.586	2.353	0.935	11.236
Non-word decoding speed	Phonemic verbal fluency	6.369	1	75	0.014	0.007	0.003	2.524	0.002	0.013
Non-word decoding speed	Digit RAN speed	3.993	1	75	0.049	0.181	0.091	1.998	0.001	0.362
Non-word decoding errors	Graphic reproduction (2 circles) errors	4.160	1	62	0.046	0.060	0.029	2.040	0.001	0.118
Word decoding stress errors	Picture RAN errors	4.499	1	63	0.038	7.429	3.503	2.121	0.430	14.429
Word decoding attentional errors	Color RAN errors	8.648	1	63	0.005	8.829	3.002	2.941	2.829	14.829
Reading comprehension total score	Syntactic production of DOcls	6.245	1	61	0.015	0.262	0.105	2.499	0.052	0.471
Reading comprehension total score	Phonemic verbal fluency	4.473	1	61	0.039	0.010	0.005	2.115	0.001	0.019
Reading comprehension central score	Syntactic production of DOcls	4.259	1	60	0.043	0.338	0.164	2.064	0.010	0.666
Reading comprehension central score	Final syllable deletion	4.834	1	60	0.032	0.034	0.015	2.198	0.003	0.064
Reading comprehension central score	Phonemic verbal fluency	7.428	1	60	0.008	0.018	0.006	2.725	0.005	0.031
Reading comprehension specific score	Graphic reproduction (2 circles) errors	5.987	1	64	0.017	- 0.038	0.015	- 2,447	- 0.068	- 0.007
Non-word dictation errors	Initial syllable deletion	5.331	1	81	0.045	- 0.039	0.019	- 2.037	- 0.076	- 0.001
Text-dictation total error score	Picture RAN errors	4.846	1	61	0.032	10.869	4.938	2.201	0.996	20.742
Text-dictation phonological errors	Initial syllable deletion	8.368	1	65	0.005	- 0.065	0.022	- 2.893	- 0.109	- 0.020
Text-dictation non-phonological errors	Phoneme blending	10.419	1	72	0.002	- 0.053	0.016	- 3.228	- 0.085	- 0.020
Text-dictation stress/geminate errors	Non-word repetition	4.786	1	52	0.033	- 0.032	0.014	- 2.188	- 0.061	- 0.003
Text-dictation omissions	Syntactic production of DOcls	7.404	1	32	0.010	- 0.703	0.258	- 2.721	- 1.229	- 0.177
Text-dictation omissions	Digit span backward	10.718	1	32	0.003	- 0.468	0.143	- 3,274	- 0.759	- 0.177
Text-dictation minor errors	Non-word repetition	5.175	1	52	0.027	- 0.021	0.009	- 2.275	- 0.040	- 0.003
Text-dictation minor errors	Graphic reproduction (2 circles) errors	7.137	1	52	0.010	0.073	0.027	2,671	0.018	0.128

Table B.8. Significant effects of longitudinal scores of short-term memory, metaphonology, and RAN obtained in the three administrations of 1st and 2nd grade, on literacy scores obtained in the middle of the 2nd grade.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Est.</i>	<i>Std. err.</i>	<i>t</i>	<i>CI</i>	
									Lower	Upper
Word decoding speed (syll/sec)	Spoonerism	4.733	1	22	0.041	0.010	0.005	2.176	0.000	0.020
Word decoding stress errors	Syntactic comprehension	5.046	1	23	0.035	- 0.071	0.032	- 2.246	- 0.136	- 0.006
Non-word decoding speed	Spoonerism	7.596	1	22	0.012	0.012	0.004	2.756	0.003	0.021
Non-word decoding errors	Spoonerism	4.596	1	22	0.043	- 0.003	0.001	- 2.144	- 0.006	- 0.000
Word-dictation errors	Spoonerism	4.671	1	21	0.042	- 0.035	0.016	- 2.161	- 0.069	- 0.001
Text decoding errors	Syntactic comprehension	4.492	1	24	0.045	- 0.074	0.035	- 2.119	- 0.146	- 0.002
Reading comprehension total score	Syntactic comprehension	11.499	1	21	0.003	0.033	0.010	3.391	0.013	0.054
Reading comprehension central score	Syntactic comprehension	6.948	1	20	0.016	0.033	0.012	2.636	0.007	0.059
Reading comprehension specific score	Syntactic comprehension	5.863	1	21	0.025	0.030	0.012	2.421	0.004	0.055
Text dictation phonological errors	Productive vocabulary	4.532	1	22	0.045	- 0.037	0.017	- 2.129	- 0.073	- 0.001
Text dictation phonological errors	Spoonerism	6.057	1	22	0.022	- 0.031	0.012	- 2.461	- 0.057	- 0.005

Table B.9. Significant effects of language skills on literacy skills in the middle of the 2nd grade.

Appendix C

Study 3

C.5.3. Results

This section describes the results of the additional analyses of the Standardized Mean Differences (SMDs) conducted on the separate subgroups by using Hedge's g value of the effect size with small-sample correction for the independent and correlated samples.

These analyses constitute an alternative option to those described in section 5.3.

The following paragraphs summarize the main findings and are followed by full statistical tables.

Tables of independent samples include also the *Mean* and *Standard Deviation* of each measure.

Additional rank tests were performed on the independent (Kruskal-Wallis) and correlated (Wilcoxon) samples. Their outcomes are not reported here for a matter of space but their relevant results are discussed in section 5.4.

C.5.3.1. Effects of the training on pre-literacy skills

After excluding children with developmental disorders (see 5.2.1), there was a significant effect of group on the graphic reproduction of 2 semicircles among monolingual children ($H(2) = 8.432, p = 0.015$). The post hoc comparison revealed that RAN monolinguals ($M = 7.000, SD = 1.414$) made significantly more errors ($p = 0.015$) than GRAPH monolinguals ($M = 2.000, SD = 1.414$). The effect size was very large ($g = 3.143, SE = 1.194, CI [0.802, 12.217]$).

This effect was not present in bilingual groups.

Global metaphonological skills

An analysis of Hedge's value (see 5.2.4) after administration KG-2 confirmed

- a large effect size for the difference between the mono-META group and the mono-CG in the recognition of word minimal pairs ($g = 0.915$);
- a large effect size for the difference between the META group and the CG in rhyme recognition among both monolinguals ($g = 0.922$) and bilinguals ($g = 1.155$);
- a very large effect size for the difference between the bi-META and the bi-CG in the initial-syllable recognition ($g = 1.365$).

The within-sample analysis confirmed

- a very large improvement in rhyme recognition in the mono-META ($g_{av} = - 1.163$) and an even larger improvement in the bi-META group ($g_{av} = - 1.694$);
- a very large improvement in the recognition of the initial syllable by both the mono-META ($g_{av} = - 1.363$) and an even larger improvement in the bi-META group ($g_{av} = - 2.056$);
- a very large improvement in syllable blending by the mono-RAN group ($g_{av} = - 1.642$).

There was also a very large improvement of the mono-META group in syllable segmentation ($g_{av} = - 1.121$).

All these effects were significant at non-parametric tests (Kruskal Wallis / Wilcoxon) ($p < 0.05$).

Analytical metaphonological skills

After administration 1-1, the analysis of SMDs showed (see 5.2.4)

- a large advantage of the mono-GRAPH group compared to the mono-CG in phoneme blending ($g = 1.130$) and final syllable deletion ($g = 1.690$). The former was not significant, whereas the latter was significant at the Kruskal-Wallis test ($p = 0.002$).
- large advantages of the bi-META group compared to the bi-CG in phoneme blending ($g = 1.032$); phoneme segmentation ($g = 1.128$); and phonemic verbal fluency ($g = 1.285$);
- a large advantage of the bi-META group compared to the bi-RAN group in phonemic segmentation ($g = 1.789$).

None of these effects was significant at the Kruskal-Wallis test.

Between administrations 1-1 and 1-2

- a very large improvement of syllable blending was confirmed by the mono-META group ($g_{av} = - 1.385$), the mono-RAN group ($g_{av} = - 1.878$), and the bi-RAN group ($g_{av} = - 1.713$);
- initial syllable deletion largely improved in the mono-META group ($g_{av} = - 1.059$);
- a tremendous improvement of phonemic verbal fluency took place among bi-META children ($g_{av} = - 6.598$).

As for rhyme recognition in primary school (for rhyme recognition in kindergarten, see the previous section), small improvement was confirmed between KG-2 and 1-1 in the mono-CG ($g_{av} = - 0.417$) and bi-CG ($g_{av} = - 0.396$). The global (mono + bi) META group confirmed a strong improvement between administrations 1-1 and 1.2 ($g_{av} = - 0.857$), whereas the global RAN group showed a moderate improvement between administrations 1-2 and 2-1 ($g_{av} = - 0.549$).

RAN

A large effect size was confirmed in administration 1-1 of digit RAN for the speed difference between the global GRAPH group and the global CG ($g = 1.197$).

Large improvement was confirmed for picture RAN speed between administrations KG-1 and KG-2 by the global META group ($g_{av} = -1.467$), the global RAN group ($g_{av} = -2.527$), and the global GRAPH group ($g_{av} = -2.176$). Large improvement of the same ability was also present between administrations KG-2 and 1-1 in both the META groups (mono $g_{av} = -1.566$; bi $g_{av} = -1.131$), the RAN groups (mono $g_{av} = -2.012$; bi $g_{av} = -2.490$), in the mono GRAPH group ($g_{av} = -0.420$), and in the CGs (mono $g_{av} = -1.449$; bi $g_{av} = -1.624$).

Visual skills

Moderate improvement was confirmed between administrations KG-1 and KG-2 of the graphic reproduction of 2 semicircles by the global META group ($g = 0.686$), whereas the effects in the mono-RAN group ($g = 1.044$) and bi RAN group ($g = 1.111$) were larger.

Non-word repetition

A moderate effect size was confirmed for the difference between the global META group and the global CG in administration KG-2 of the non-word repetition test ($g = 0.747$). Larger improvement was confirmed between KG-2 and 1-1 by the global RAN group ($g_{av} = -1.346$) and the bi-RAN group ($g_{av} = -1.177$).

Syntactic skills

The global GRAPH group confirmed a large advantage compared to the global CG in syntactic production of direct object clitic pronouns in administration 2-1 ($g = 1.145$). A large effect was present in both monolingual ($g = 1.209$) and bilingual ($g = 4.190$) GRAPH children, whereas a moderate advantage in the same administration was also present in the global META group ($g = 0.750$).

A moderate to large improvement took place between administration KG-2 and 2-1 only among META children.

No effect of syntactic comprehension emerged.

Receptive and productive vocabulary

No effect was confirmed for both receptive and productive vocabulary.

Digit and word span

A large improvement of digit span forward and backward took place between administrations KG-1 and 1-1 in the global META group (forward: $g_{av} = -1.136$; backward: $g_{av} = -1.126$); in the global RAN group (forward: $g_{av} = -0.910$; backward: $g_{av} = -1.064$); and of digit span forward in the mono META group ($g_{av} = -1.214$) and in the global GRAPH group ($g_{av} = -1.162$).

There was also a moderate improvement of digit span forward ($g_{av} = -0.604$) and a quite larger improvement in digit span backward ($g_{av} = -0.834$) by the bi-CG.

In administration 1-1 there was a large advantage of the mono-GRAPH group compared to the mono-CG in digit span backward ($g = 1.163$).

In administration KG-2, word span showed a large advantage of the global META group compared to the GRAPH group ($g = 0.999$) and a moderate too high effect compared to the CG ($g = 0.793$).

In administration 1-1, there was a moderate difference between the global META group and the global CG ($g = 0.750$).

5.3.2.A Effects of the training on literacy skills

Word and non-word decoding

The global META group was confirmed to be faster than the RAN group in word decoding with a very large effect size ($g = 1.345$), which was significant at the Kruskal-Wallis test (Adapt. $p = 0.040$).

No effect was confirmed for word decoding errors.

As for non-word decoding speed, the global META group was confirmed to be much largely faster than the global CG ($g = 2.560$) and largely faster than the RAN group ($g = 1.247$). The difference with the CG was only significant at the Kruskal-Wallis test (Adapt. $p = 0.044$).

As for non-word decoding errors, a large effect for the difference between the global META and the global GRAPH group was found ($g = 1.744$), with the former making more errors than the latter. The difference was not significant at the Kruskal-Wallis test.

Sublexical decoding

In administration 1-1, the global and the monolingual GRAPH group confirmed higher speed than the respective CGs, with very large effects (global sample $g = 1.223$; mono-sample $g = 1.989$).

The global RAN group confirmed to make more errors than the global GRAPH group, with a large effect size ($g = 1.239$). The difference was significant at the Kruskal-Wallis test ($p = 0.037$).

For administration 2-1, see above the results of the non-word decoding test.

Text decoding

In the global META group, there was a large increase in text decoding speed from the former to the latter administration ($g_{av} = - 1.734$); in the global RAN group, the increase was even larger ($g_{av} = - 2.126$); in the GRAPH group, it was moderate ($g_{av} = 0.588$).

A large increase of speed was present in the mono-META group ($g_{av} = - 1.834$) but not in the bi-META group; in the mono-RAN ($g_{av} = - 1.747$), and in the bi-RAN ($g_{av} = - 1.814$) group.

In administration 1-2, the global RAN group made more errors than the global GRAPH group, with a large effect ($g = 1.376$). The mono-RAN group showed the same difference, with a larger effect size ($g = 3.594$). In addition, the mono-META group made more errors than the mono-GRAPH group ($g = 2.995$), and the difference was significant at the Kruskal-Wallis test (ad. $p = 0.025$).

The within-subject analysis showed a large decrease of word decoding errors between the first and second administration in the global META group ($g_{av} = 1.286$) and in the RAN group ($g_{av} = 1.414$) but not in the GRAPH group. In particular, the error-decrease effect was very large in the mono-META group ($g_{av} = 2.084$) and in the bi-RAN group ($g_{av} = 2.866$).

Reading comprehension

In administration 1-2, the MT-2 collective reading comprehension tests showed a large difference in favor of the global META group compared to the global RAN group ($g = 1.057$). As for the MT-3 tests, there was a very large difference in favor of the global GRAPH group compared to the CG in specific comprehension ($g = 2.683$). These differences were not significant. In administration 2-1, the global META group had a higher score than the RAN group in reading central comprehension ($g = 1.055$). These differences were not significant at the Kruskal-Wallis test.

The within-subject analysis showed an improvement between the two administrations only for MT-3 scores of the META group. The effect for comprehension total score was moderate ($g_{av} = -0.621$), whereas for the central comprehension it was larger ($g_{av} = -0.893$).

Similar improvement emerged in the mono-META group, with slightly larger effects (total score: $g_{av} = -0.717$; central score: $g_{av} = -1.050$) but not in the BI groups.

Word and non-word writing

In administration 1-1, the global RAN group made more word errors than the global GRAPH group ($g = 1.447$). The difference was not significant at the Kruskal-Wallis test.

There were large improvements between administration 1-1 and 1-2 for every group except the GRAPH groups.

No effect was confirmed for non-word writing.

Text writing

There were large effect sizes for the difference of stress geminate errors between the META and the GRAPH group (global sample ($g = 1.239$); monolinguals ($g = 1.830$) in stress/geminate errors, omissions (in monolinguals ($g = 1.421$)), which were more numerous in the META than in the GRAPH group.

Moreover, total errors ($g = 1.606$), phonological errors ($g = 1.369$), stress/geminate errors ($g = 1.586$), and minor errors ($g = 1.566$) were more numerous in the global RAN group than in the global GRAPH group. The difference in minor errors was also present in monolingual RAN/GRAPH children, with a very large effect ($g = 3.561$), but not in bilinguals.

The differences between RAN and GRAPH children in total errors (ad. $p = 0.039$) and phonological errors (ad. $p = 0.049$) were only significant at the Kruskal-Wallis test.

Statistics tables

Test	Contrast	M_1	SD_1	M_2	SD_2	SMD		Conf.Int.	
						g	Std.Err.	Lower	Upper
KG-2_WORD_SPAN	METAvsGRAPH	3,333	0,492	2,857	0,378	0,999	0,509	0,001	1,996
KG-2_WORD_SPAN	METAvsCG	3,333	0,492	2,854	0,618	0,793	0,331	0,143	1,278
KG-2_SYLL_SYNTH	METAvsCG	14,083	1,311	12,255	2,793	0,701	0,330	0,053	1,138
KG-2_W_MIN_PAIRS	METAvsCG	14,417	0,669	12,106	2,906	0,868	0,334	0,213	1,407
KG-2_RHYME_REC	METAvsCG	13,500	2,316	8,809	4,821	1,041	0,338	0,377	1,746
KG-2_INITIAL_SYLL_REC	METAvsCG	13,667	1,303	9,511	4,601	0,983	0,337	0,323	1,626
KG-2_SYLL_SEGM	METAvsCG	13,500	3,989	10,553	4,466	0,664	0,330	0,018	1,087
KG-2_NW_REP_CORR_SYLL	METAvsCG	53,250	4,181	48,083	7,304	0,747	0,330	0,100	1,206
KG-2_W_MIN_PAIRS	METAvsRAN	14,417	0,669	12,778	1,856	1,203	0,487	0,250	2,402
KG-2_RHYME_REC	METAvsRAN	13,500	2,316	9,556	4,640	1,085	0,479	0,148	2,116
KG-2_INITIAL_SYLL_REC	METAvsRAN	13,667	1,303	10,222	4,086	1,168	0,484	0,219	2,314
KG-2_NW_REP_CORR_SYLL	METAvsRAN	53,250	4,181	48,000	3,742	1,259	0,491	0,297	2,548
KG-2_SYLL_BLEND	RANvsCG	14,222	1,302	12,255	2,793	0,738	0,371	0,011	1,272
KG-2_RHYME_REC	GRAPHvsCG	12,714	2,430	8,809	4,821	0,835	0,414	0,024	1,508
KG-2_INITIAL_SYLL_REC	GRAPHvsCG	13,286	2,360	9,511	4,601	0,845	0,414	0,034	1,526

Table C.1. Means, Standard Deviations, and Standardized Mean Differences between global (mono + bi) groups in administration KG-2 (Post-intervention 1)

META			SMD		Conf.Int.	
		Mean difference	g_{av}	SE	Lower	Upper
Pair	KG-1_W_MIN_PAIRS - KG-2_W_MIN_PAIRS	-1,583	-0,903	0,609	-2,923	-0,243
Pair	KG-1_RHYME_REC - KG-2_RHYME_REC	-4,417	-1,429	0,892	-6,379	-2,454
Pair	KG-1_NW_MIN_PAIRS - KG-2_NW_MIN_PAIRS	-1,917	-0,670	0,857	-3,803	-0,031
Pair	KG-1_INITIAL_SYLL_REC - KG-2_INITIAL_SYLL_REC	-3,500	-1,704	0,435	-4,458	-2,542
Pair	KG-1_SYLL_SEGM - KG-2_SYLL_SEGM	-4,167	-0,887	1,546	-7,570	-0,763
Pair	KG-1_VIS_CIR_ERR_TOT - KG-2_VIS_CIR2	2,000	0,686	0,896	0,028	3,972
Pair	KG-1_NW_REP_CORR_SYLL - KG-2_NW_REP_CORR_SYLL	-5,250	-0,889	1,887	-9,404	-1,096

RAN	Contrast	Mean difference	SMD		Conf.Int.	
			g_{av}	SE	Lower	Upper
Pair	KG-1_NW_MIN_PAIRS - KG-2_NW_MIN_PAIRS	-2,778	-0,668	1,011	-5,109	-0,447
Pair	KG-1_VIS_CIR1_ERR - KG-2_VIS_CIR1	1,111	0,661	0,389	0,214	2,008
Pair	KG-1_VIS_CIR_ERR_TOT - KG-2_VIS_CIR_tot_ERR	4,222	1,138	0,683	2,648	5,796
GRAPH KG			SMD		Conf.Int.	
		Mean difference	g_{av}	SE	Lower	Upper
Pair	KG-1_INITIAL_SYLL_REC - KG-2_INITIAL_SYLL_REC	-2,143	-0,561	0,800	-4,100	-0,186
Pair	KG-1_NW_REP_CORR_SYLL - KG-2_NW_REP_CORR_SYLL	-3,143	-0,563	1,262	-6,230	-0,056

Table C.2. Within-subject SMD (g_{av}) between Pre-intervention 1 and Post-intervention 1

						SMD		Conf.Int.	
	Contrast	M_1	SD_1	M_2	SD_2	g	Std.Err.	Lower	Upper
MONO									
KG-2_W_MIN_PAIRS	METAvsCG	14,714	0,488	13,417	1,530	0,915	0,447	0,039	1,715
KG-2_RHYME_REC	METAvsCG	14,143	1,864	10,625	4,063	0,922	0,447	0,045	1,727
BI									
KG-2_RHYME_REC	METAvsCG	12,600	2,793	7,227	4,770	1,155	0,524	0,128	2,360
KG-2_INITIAL_SYLL_REC	METAvsCG	13,600	1,342	7,955	4,337	1,365	0,535	0,317	2,911

Table C.3. Means, Standard Deviations, and Standardized Mean Differences between mono- and bi- groups in administration KG-2 (Post-intervention 1)

MONO - META*Conf. Int.*

	Mean difference	g_{av}	SE	Lower	Upper
KG-1_RHYME_REC - KG-2_RHYME_REC	-3,714	-1,163	1,267	-6,815	-0,614
KG-1_INITIAL_SYLL_REC - KG-2_INITIAL_SYLL_REC	-3,714	-1,363	0,680	-5,379	-2,050
KG-1_SYLL_SEGM - KG-2_SYLL_SEGM	-6,143	-1,121	2,304	-11,779	-0,506

BI META

	Mean difference	g_{av}	SE	Lower	Upper
KG-1_RHYME_REC - KG-2_RHYME_REC	-5,400	-1,694	1,208305	-8,754791	-2,045209
KG-1_INITIAL_SYLL_REC - KG-2_INITIAL_SYLL_REC	-3,200	-2,056	0,489898	-4,560175	-1,839825

MONO RAN

	Mean difference	g_{av}	SE	Lower	Upper
KG-1_SYLL_BLEND - KG-2_SYLL_BLEND	-2,000	-1,642	0,632456	-3,755978	-0,244022
KG-1_VIS_CIR2_ERR - KG-2_VIS_CIR2	2,800	1,044	0,860233	0,411612	5,188388
KG-1_VIS_CIR_ERR_TOT - KG-2_VIS_CIR_tot_ERR	3,800	0,912	0,663325	1,958315	5,641685

BI RAN

	Mean difference	g_{av}	SE	Lower	Upper
KG-1_NW_MIN_PAIRS - KG-2_NW_MIN_PAIRS	-4,250	-0,736	1,030776	-7,530391	-0,969609
KG-1_VIS_CIR2_ERR - KG-2_VIS_CIR2	3,500	1,111	0,866025	0,743921	6,256079
KG-1_VIS_CIR_ERR_TOT - KG-2_VIS_CIR_tot_ERR	4,750	0,914	1,376893	0,368113	9,131887

BI GRAPH

	Mean difference	g_{av}	SE	Lower	Upper
KG-1_RAN_SPEED_OBJSEC - KG-2_RAN_SPEED_OBJSEC	-0,051	-0,343	0,009849	-0,093376	-0,008624

Table C.4. Within-subject SMD (g_{av}) between Pre-intervention 1 and Post-intervention 1

	Contrast	M_1	SD_1	M_2	SD_2	SMD		Conf.Int	
						g	Std.Err.	Lower	Upper
1-1_PHON_FLUENCY	METAvsCG	16,833	8,494	11,522	6,352	0,768	0,332	0,116	1,241
1-1_RHYME_REC	METAvsRAN	13,167	2,368	9,444	4,640	1,018	0,474	0,089	1,967
1-1_SUBLEX_READING_ERR_PERC	RANvsGRAPH	0,242	0,176	0,057	0,043	1,239	0,589	0,084	2,690
1-1_DICT_16_WORD_ERR_PERC	RANvsGRAPH	0,785	0,160	0,480	0,249	1,447	0,610	0,251	3,290
1-1_PHON_SYNTH	GRAPHvsCG	11,000	2,280	6,196	4,097	1,197	0,451	0,313	2,317
1-1_INITIAL_SYLL_DEL	GRAPHvsCG	12,833	2,229	8,152	5,258	0,915	0,444	0,045	1,708
1-1_FIN_SYLL_DEL	GRAPHvsCG	12,500	3,987	7,739	4,519	1,049	0,447	0,173	1,977
1-119_RAN_SPEED_NUMSEC	GRAPHvsCG	1,333	0,379	0,945	0,399	0,962	0,444	0,091	1,796
1-1_SUBLEX_READING_SPEED_SYLLSEC	GRAPHvsCG	0,912	0,491	0,476	0,332	1,223	0,453	0,336	2,382

Table C.5. Means, Standard Deviations, and Standardized Mean Differences between global (mono + bi groups) in administration 1-1 (Follow-up 1)

META KG-2 - 1-1			SMD		Conf. Int.	
		M difference	g_{av}	SE	Lower	Upper
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,286	-1,468	0,055	-0,407	-0,165
Pair	KG-1_DIG_SPAN_FW - 1st_gr_DIG_SPAN_FW	-0,600	-1,136	0,163	-0,969	-0,231
Pair	KG-1_DIG_SPAN_BW - 1st_gr_DIG_SPAN_BW	-0,800	-1,126	0,291	-1,457	-0,143
Pair	KG-2_CLIT_PERC_Casani_Cardinaletti – 2-1_CLIT_PERC_Casani_et_al	-0,279	-0,798	0,078	-0,455	-0,104

RAN KG2 - 1-1			SMD		Conf. Int.	
		M difference	g_{av}	SE	Lower	Upper
Pair	KG-2_VIS_CIR1 - 1-1_VIS_CIR1_REPR_ERR	0,889	1,077	0,309	0,176	1,602
Pair	KG-2_VIS_LETT_ERR - 1-1_VIS_LETT_REC_ERR	2,000	1,143	0,687	0,415	3,585
Pair	KG-2_NW_REP_CORR_SYLL - 1-1_NW_REP	-5,556	-1,346	1,324	-8,609	-2,502
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,450	-2,527	0,048	-0,561	-0,339
Pair	KG-1_DIG_SPAN_FW – 1-1_DIG_SPAN_FW	-0,444	-0,910	0,176	-0,850	-0,039
Pair	KG-1_DIG_SPAN_BW – 1-1_DIG_SPAN_BW	-1,111	-1,064	0,389	-2,008	-0,214

GRAPH KG2 - 1-1

		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RAN_SPEED_PICT/SEC - 1-1_RAN_SPEED_PICTSEC	-0,391	-2,176	0,069	-0,568	-0,214
Pair	KG-1_DIG_SPAN_FW – 1-1_DIG_SPAN_FW	-0,667	-1,162	0,211	-1,209	-0,125

Table C.6. Within-subject SMD (*g_{av}*) between Post-intervention 1 and Follow-up 1

						<i>SMD</i>		<i>Conf.Int</i>	<i>CI</i>
MONO	Contrast	<i>M₁</i>	<i>SD₁</i>	<i>M₂</i>	<i>SD₂</i>	<i>g</i>	<i>Std.Err.</i>	<i>Lower</i>	<i>Upper</i>
1-1_PHON_SYNTH	GRAPHvsCG	12,000	1,414	7,458	4,118	1,130	0,564	0,024	2,382
1-1_FIN_SYLL_DEL	GRAPHvsCG	14,750	0,500	8,000	4,118	1,690	0,592	0,529	4,019
1-1_SUBLEX_READING_SPEED_SYLLSEC	GRAPHvsCG	1,100	0,432	0,492	0,274	1,989	0,611	0,790	5,154
1-1_DIG_SPAN_BW	GRAPHvsCG	3,250	0,500	2,417	0,717	1,163	0,565	0,055	2,461
1-1_CLIT_PERC_Casani_et_al	GRAPHvsCG	0,970	0,020	0,794	0,150	1,209	0,567	0,096	2,573
BILINGUALS									
1-1_PHON_SYNTH	METAvsCG	8,800	4,025	4,818	3,686	1,032	0,518	0,016	2,080
1-1_PHON_SEGM	METAvsCG	11,000	3,674	5,429	4,976	1,128	0,526	0,097	2,303
1-1_PHON_FLUENCY	METAvsCG	15,400	3,782	8,818	5,161	1,285	0,530	0,246	2,691
1-1_PHON_SEGM	METAvsRAN	11,000	3,674	4,250	2,872	1,789	0,875	0,073	4,916
1-1_CLIT_PERC_Casani_et_al	GRAPHvsCG	0,96	0,056569	0,38	0,351392	1,630	0,782	0,097	4,190

Table C.7. Means, Standard Deviations, and Standardized Mean Differences between mono- and bi- groups in administration 1-1 (Follow-up 1)

MONO META			SMD		Conf.Int.	
	Contrast	M difference	g _{av}	SE	Lower	Upper
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,332	-1,566	0,087	-0,544	-0,120
Pair	KG-1_DIG_SPAN_FW - 1st_gr_DIG_SPAN_FW	-0,714	-1,214	0,184	-1,166	-0,263

BI META

		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,221	-1,131	0,048	-0,355	-0,088
MONO RAN						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,461	-2,012	0,082	-0,689	-0,234
BI RAN						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_NW_REP_CORR_SYLL - 1-1_NW_REP	-5,750	-1,177	1,652	-11,007	-0,493
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,436	-2,490	0,050	-0,597	-0,276
MONO GRAPH						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,420	-2,627	0,092	-0,712	-0,128
Pair	KG-2_RAN_OBJ_ERR_PERC - 1-1_RAN_FIG_ERR_PERC	0,021	0,498	0,003	0,010	0,032
MONO CG						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RHYME_REC - 1-1_RHYME_REC	-1,625	-0,417	0,531	-2,723	-0,527
Pair	KG-2_NW_REP_CORR_SYLL - 1-1_NW_REP	-1,583	-0,265	0,747	-3,128	-0,039
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,289	-1,449	0,034	-0,360	-0,218
BI CG						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	KG-2_RHYME_REC - 1-1_RHYME_REC	-1,857	-0,396	0,718	-3,355	-0,359
Pair	KG-2_NW_REP_CORR_SYLL - 1-1_NW_REP	-3,652	-0,477	1,218	-6,179	-1,125
Pair	KG-2_RAN_SPEED_OBJSEC - 1-1_RAN_SPEED_PICT/SEC	-0,316	-1,624	0,041	-0,401	-0,231
Pair	KG-2_RAN_OBJ_ERR_PERC - 1-1_RAN_FIG_ERR_PERC	0,081	0,938	0,023	0,032	0,129
Pair	KG-1_DIG_SPAN_FW - 1st_gr_DIG_SPAN_FW	-0,350	-0,604	0,109	-0,579	-0,121
Pair	KG-1_DIG_SPAN_BW - 1st_gr_DIG_SPAN_BW	-0,700	-0,834	0,164	-1,043	-0,357

Table C.8. Within-subject SMD (*g_{av}*) between Post-intervention 1 and Follow-up 1

	Contrast	M_1	SD_1	M_2	SD_2	SMD		Conf. Int.	
						g	Std.Err.	Lower	Upper
1st_gr_WORD_SPAN	METAvsCG	3,556	0,527	3,021	0,729	0,750	0,370	0,024	1,289
ALL_JAN20_CLIT_PERC_Casani_et_al	METAvsCG	0,839	0,203	0,596	0,336	0,750	0,372	0,021	1,292
1-2_1st_SYLL_DEL	METAvsRAN	14,000	1,118	8,625	5,878	1,245	0,544	0,180	2,615
1-2_RHYME_REC	METAvsRAN	14,444	0,726	10,375	4,340	1,282	0,547	0,210	2,716
1-2_PHON_SEGM	METAvsRAN	13,333	2,121	10,375	2,326	1,265	0,545	0,196	2,670
1-2_FIN_SYLL_DEL	METAvsRAN	12,222	2,167	8,250	4,590	1,073	0,529	0,036	2,190
1-2_MT2_TEXT_COMPR_CORR_PERC	METAvsRAN	0,822	0,156	0,625	0,198	1,057	0,528	0,022	2,152
1-2_MT3_DEC_ERR_PERC	RANvsGRAPH	0,125	0,063	0,039	0,052	1,376	0,621	0,159	3,111
1-2_CLIT_PERC_Casani_et_al	GRAPHvsCG	0,967	0,030	0,596	0,336	1,145	0,450	0,264	2,193
1-2_MT3_COMPR_SPEC	GRAPHvsCG	0,750	0,105	0,467	0,058	2,683	1,101	0,526	9,357

Table C.9. Means, Standard Deviations, and Standardized Mean Differences between global groups in administration 1-2 (Post-intervention 2)

META 1-1 - 1.2			SMD		Conf.Int.	
		Mean difference	g_{av}	SE	Lower	Upper
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-4,556	-1,244	1,203	-7,330	-1,781
Pair	1-1_1st_SYLL_DEL - 1-2_1st_SYLL_DEL	-3,333	-0,893	1,258	-6,235	-0,432
Pair	1-1_RHYME_REC - 1-2_RHYME_REC	-1,778	-0,857	0,722	-3,443	-0,112
Pair	1-1_PHON_FLUENCY - 1-2_PHON_FLUENCY	-5,222	-0,441	1,362	-8,363	-2,082
Pair	1-1_FIN_SYLL_DEL - 1-2_FIN_SYLL_DEL	-3,222	-1,023	1,341	-6,316	-0,129
Pair	1-119_RAN_SPEED_NUMSEC - 1-2_RAN_NUMSEC	-0,232	-0,631	0,078	-0,412	-0,052
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,493	1,908	0,092	0,280	0,706
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,464	1,976	0,093	0,249	0,679

Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,375	1,989	0,075	0,203	0,547
RAN 1-1 - 1-2		SMD				
		<i>Mean difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-5,750	-1,898	0,773	-7,579	-3,921
Pair	1-1_1st_SYLL_DEL - 1-2_1st_SYLL_DEL	-3,875	-0,671	1,342	-7,049	-0,701
Pair	1-1_FIN_SYLL_DEL - 1-2_FIN_SYLL_DEL	-2,375	-0,415	0,925	-4,561	-0,189
Pair	1-119_RAN_SPEED_COLSEC - 1st_gr_RAN_JUNE_COLSEC	-0,150	-0,593	0,046	-0,259	-0,041
Pair	1-119_RAN_SPEED_NUMSEC - 1-2_RAN_NUMSEC	-0,353	-1,339	0,060	-0,499	-0,206
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,438	2,221	0,049	0,322	0,553
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,425	1,972	0,050	0,307	0,543
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,368	1,664	0,049	0,252	0,485
GRAPH 1-1 - 1-2						
		<i>Mean difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-2,333	-0,923	0,211	-2,875	-1,791
Pair	1-1_PHON_FLUENCY - 1-2_PHON_FLUENCY	-3,333	-1,565	1,054	-6,043	-0,624
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,281	6,837	0,083	0,067	0,496
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,229	0,986	0,080	0,024	0,434
CG 1-1 - 1-2		SMD				
		<i>Mean difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
					Inferiore	Superiore
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,379	1,453	0,042	0,294	0,465
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,395	1,560	0,039	0,315	0,475
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,393	0,039	0,314	0,473	1,632

Table C.10. Within-subject SMD (*g_{av}*) between Follow-up 1 and Post-intervention 2

						<i>SMD</i>		<i>Conf.Int</i>	
MONO 1-2	Contrast	M_1	SD_1	M_2	SD_2	g	<i>Std.Err.</i>	<i>Lower</i>	<i>Upper</i>
1-2_MT3_baco_DEC_ERR_PERC	META vs GRAPH	0,079	0,027	0,007	0,008	2,995	1,076	0,887	11,078
1-2_MT3_baco_DEC_ERR_PERC	RANvsGRAPH	0,074	0,021	0,007	0,008	3,594	1,446	0,760	15,750
1-2_DIG_SPAN_BW	GRAPHvsCG	3,250	0,500	2,417	0,717	1,163	0,565	0,055	2,461

Table C.11. Means, Standard Deviations, and Standardized Mean Differences between mono. and bi- groups in administration 1-2 (Post-intervention 2)

MONO META		SMD			Conf.Int.	
		M difference	g _{av}	Std.Err.	Lower	Upper
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-5,667	-1,385	1,606	-9,794	-1,539
Pair	1-1_1st_SYLL_DEL - 1-2_1st_SYLL_DEL	-4,500	-1,059	1,668	-8,789	-0,211
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,510	1,512	0,140	0,150	0,871
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,442	1,432	0,140	0,081	0,803
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,322	1,381	0,106	0,050	0,594

BI META						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_PHON_FLUENCY - 1-2_PHON_FLUENCY	-6,667	-6,598	0,667	-9,535	-3,798
Pair	1-1_RAN_COL_ERR_PERC - 1-2_RAN_COL_ERR_PERC	0,027	1,866	0,003	0,012	0,041
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,458	2,137	0,055	0,221	0,695
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,508	3,265	0,065	0,228	0,788
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,480	4,341	0,051	0,261	0,698

MONO RAN						
		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-5,750	-1,878	1,436	-10,320	-1,180
Pair	1-119_RAN_SPEED_COLSEC - 1st_gr_RAN_JUNE_COLSEC	-0,205	-0,531	0,058	-0,391	-0,018

Pair	1-119_RAN_SPEED_NUMSEC - 1-2_RAN_NUMSEC	-0,441	-1,250	0,032	-0,544	-0,338
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,469	2,195	0,065	0,261	0,676
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,481	2,713	0,075	0,242	0,721
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,392	2,003	0,073	0,159	0,624

BI RAN

		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_PHON_SYNTH - 1-2_PHON_SYNTH	-5,750	-1,713	0,854	-8,468	-3,032
Pair	1-1_RAN_COL_ERR_PERC - 1-2_RAN_COL_ERR_PERC	0,033	0,743	0,008	0,009	0,056
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,407	1,691	0,079	0,156	0,657
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,369	1,129	0,062	0,173	0,565
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,345	1,008	0,075	0,107	0,583

MONO CG

		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,356	1,534	0,055	0,238	0,473
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,342	1,597	0,053	0,229	0,456
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,309	1,546	0,052	0,199	0,419

BI CG

		<i>M difference</i>	<i>g_{av}</i>	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-1_DICT16_WORD_ERR_PERC - 1-2_DICT16_WORD_ERR	0,406	1,714	0,065	0,266	0,547
Pair	1-1_DICT16_SYLL_ERR_PERC - 1-2_DICT16_SYLL_ERR	0,455	1,939	0,055	0,338	0,573
Pair	1-1_DICT16_GRAPH_ERR_PERC - 1-2_DICT16_GRAPH_ERR	0,489	2,170	0,049	0,384	0,594

Table C.12. Within-subject SMD (*g_{av}*) between Follow-up 1 and Post-intervention 2

	Contrast	M_1	SD_1	M_2	SD_2	SMD		Conf.Int	
						g	Std.Err.	Lower	Upper
2-1_BVSCO_PHONETIC_PERC	METAvsGRAPH	0,035	0,021	0,011	0,015	1,239	0,589	0,084	2,689
2-1_JAN_TILTAN_NW_SPEED_SYLL/SEC	METAvsCG	1,385	0,170	0,945	0,105	2,560	0,922	0,752	8,360
2-1_JAN_SPOONERISM	METAvsCG	21,667	7,036	8,333	10,116	1,588	0,775	0,069	4,041
2-1_JAN_MT3CIOP_SYLLSEC	METAvsRAN	2,786	0,632	1,959	0,474	1,393	0,557	0,301	3,033
2-1_JAN_TILTAN_WL_SPEED_SYLL/SEC	METAvsRAN	1,895	0,496	1,336	0,230	1,345	0,553	0,262	2,891
2-1_JAN_TILTAN_NW_SPEED_SYLL/SEC	METAvsRAN	1,385	0,170	1,153	0,184	1,247	0,544	0,181	2,621
2-1_JAN_PHON_SYNTH	METAvsRAN	13,667	1,871	11,625	1,847	1,042	0,527	0,009	2,119
2-1_JAN_SPOONERISM	METAvsRAN	21,667	7,036	8,375	10,378	1,441	0,562	0,340	3,177
2-1_JAN_MT3_COMPR_CENTR	METAvsRAN	0,789	0,183	0,538	0,267	1,055	0,528	0,020	2,147
2-1_JAN_DICT_BVSCO_TOT_ERR_PERC	RANvsGRAPH	0,183	0,089	0,062	0,030	1,606	0,648	0,336	3,848
2-1_BVSCO_PHONOL_PERC	RANvsGRAPH	0,096	0,064	0,024	0,015	1,369	0,620	0,153	3,091
2-1_BVSCO_PHONETIC_PERC	RANvsGRAPH	0,037	0,016	0,011	0,015	1,586	0,645	0,321	3,780
2-1_BVSCO_MINOR_ERRORS_PERC	RANvsGRAPH	0,072	0,020	0,026	0,035	1,566	0,643	0,306	3,714

Table C.13. Means, Standard Deviations, and Standardized Mean Differences between global groups in administration 2-1 (Follow-up 2)

META		SMD			Conf.Int.	
		<i>M difference</i>	g_{av}	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-2_RAN_PICT/SEC - 2-1_JAN19_RAN_SPEED_PICT/SEC	-0,138	-0,519	0,052	-0,257	-0,019
Pair	1-2_FIN_SYLL_DEL - 2-1_JAN_FIN_SYLL_DEL	-2,222	-1,242	0,813	-4,096	-0,348
Pair	1-2_MT3_baco_DEC_SPEED_SYLLSEC - 2-1_JAN_MT3CIOP_SYLLSEC	-1,268	-1,734	0,191	-1,709	-0,826
Pair	1-2_MT3_baco_DEC_ERR_PERC - 2-1_JAN_MT3CIOP_ERR_PERC	0,033	1,286	0,010	0,011	0,056
Pair	1-2_MT3_COMPR_TOT_PERC - 2-1_JAN_MT3_COMPR_TOT_PERC	-0,128	-0,621	0,053	-0,249	-0,006
Pair	1-2_MT3_COMPR_CENTR - 2-1_JAN_MT3_COMPR_CENTR	-0,178	-0,893	0,062	-0,320	-0,035
RAN						
		<i>M difference</i>	g_{av}	<i>SE</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-2_INITIAL_SYLL_DEL - 2-1_JAN_INITIAL_SYLL_DEL	-5,375	-1,118	1,880	-9,820	-0,930
Pair	1-2_RHYME_REC - 2-1_JAN_RHYME_REC	-2,375	-0,549	0,944	-4,607	-0,143
Pair	1-2_PHON_SEGM - 2-1_JAN_PHON_SEGM	-2,750	-1,310	1,146	-5,459	-0,041

Pair	1-2_FIN_SYLL_DEL - 2-1_JAN_FIN_SYLL_DEL	-5,250	-1,318	1,590	-9,009	-1,491
Pair	1-2_DICT16_WORD_ERR - 2-1_JAN_DICT_DDE_W_ERR_PERC	0,222	1,223	0,053	0,097	0,346
Pair	1-2_MT3_baco_DEC_SPEED_SYLLSEC - 2-1_JAN_MT3CIOP_SYLLSEC	-0,942	-2,126	0,130	-1,249	-0,635
Pair	1-2_MT3_baco_DEC_ERR_PERC - 2-1_JAN_MT3CIOP_ERR_PERC	0,081	1,414	0,023	0,026	0,136
GRAPH						
		M difference	g_{av}	SE	Lower	Upper
Pair	1-2_RAN_PICT/SEC - 2-1_JAN19_RAN_SPEED_FIG/SEC	-0,163	-0,711	0,060	-0,316	-0,010
Pair	1-2_RAN_NUMSEC - 2-1_JAN19_RAN_SPEED_NUM/SEC	-0,302	-0,777	0,037	-0,397	-0,206
Pair	1-2_DICT16_WORD_ERR - 2-1_JAN_DICT_DDE_W_ERR_PERC	0,132	1,101	0,050	0,004	0,260
Pair	1-2_MT3_baco_DEC_SPEED_SYLL/SEC - 2-1_JAN_MT3CIOP_SYLL/SEC	-0,775	-0,588	0,249	-1,416	-0,135
CG						
		M difference	g_{av}	SE	Lower	Upper
Pair	1-2_MT3_COMPR_SPEC - 2-1_JAN_MT3_COMPR_SPEC_CORR_PERC	-0,300	-2,969	0,058	-0,548	-0,052

Table C.14. Within-subject SMD (g_{av}) between Post-intervention 2 and Follow-up 2

MONO 2-1	Contrast	M_1	SD_1	M_2	SD_2	SMD	Std.Err.	Conf.Int	
						g		Lower	Upper
2-1_JAN_MT3CIOP_ERR_PERC	META vs GRAPH	0,029	0,010	0,013	0,005	1,736	0,816	0,137	4,614
2-1_JAN_TILTAN_NW_ERR_PERC	META vs GRAPH	0,121	0,043	0,038	0,043	1,744	0,817	0,143	4,644
2-1_BVSCO_PHONETIC_PERC	META vs GRAPH	0,041	0,020	0,007	0,008	1,830	0,832	0,198	4,981
2-1_BVSCO_OMISSIONS_PERC	META vs CG	0,024	0,023	0,004	0,009	1,421	0,522	0,398	3,042
2-1_BVSCO_MINOR_ERRORS_PERC	RAN vs GRAPH	0,064	0,015	0,007	0,013	3,561	1,436	0,747	15,494
BILINGUALS 2-1									
2-1_JAN_SPOONERISM	META vs RAN	26,000	3,000	5,000	8,042	2,716	1,337	0,095	9,997

Table C.15. Means, Standard Deviations, and Standardized Mean Differences between global groups in administration 2-1 (Follow-up 2)

MONO META		SMD			Conf.Int.	
		<i>M difference</i>	<i>g_{av}</i>	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-2_RAN_PICT/SEC - 2-1_JAN19_RAN_SPEED_PICT/SEC	-0,183	-0,895	0,053	-0,320	-0,046
Pair	1-2_MT3_DEC_SPEED_SYLLSEC - 2-1_JAN_MT3CIOP_SYLLSEC	-1,548	-1,834	0,155	-1,947	-1,149
Pair	1-2_MT3_DEC_ERR_PERC - 2-1_JAN_MT3CIOP_ERR_PERC	0,050	2,084	0,007	0,031	0,069
Pair	1-2_MT3_COMPR_TOT_PERC - 2-1_JAN_MT3_COMPR_TOT_PERC	-0,183	-0,717	0,049	-0,310	-0,056
Pair	1-2_MT3_COMPR_CENTR - 2-1_JAN_MT3_COMPR_CENTR	-0,250	-1,050	0,056	-0,395	-0,105
MONO RAN		SMD			Conf.Int.	
		<i>M difference</i>	<i>g_{av}</i>	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair	1-2_DICT16_WORD_ERR - 2-1_JAN_DICT_DDE_W_ERR_PERC	0,167	0,942	0,045	0,024	0,310
Pair	1-2_MT3_baco_DEC_SPEED_SYLLSEC - 2-1_JAN_MT3CIOP_SYLLSEC	-0,946	-1,747	0,170	-1,488	-0,405
BI		SMD			Conf.Int.	
		<i>M difference</i>	<i>g_{av}</i>	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair 5	1-2_RAN_COL_ERR_PERC - 2-1_JAN19_RAN_COL_ERR_PERC	-0,028	-0,507	0,008	-0,055	-0,001
Pair 11	1-2_1st_SYLL_DEL - 2-1_JAN_1st_SYLL_DEL	-8,000	-1,680	2,041	-14,496	-1,504
Pair 15	1-2_FIN_SYLL_DEL - 2-1_JAN_FIN_SYLL_DEL	-7,250	-1,953	1,887	-13,257	-1,243
Pair 18	1-2_MT3_baco_DEC_SPEED_SYLLSEC - 2-1_JAN_MT3CIOP_SYLLSEC	-0,937	-1,814	0,223	-1,646	-0,228
Pair 19	1-2_MT3_baco_DEC_ERR_PERC - 2-1_JAN_MT3CIOP_ERR_PERC	0,130	2,866	0,027	0,043	0,217
#DIV/0!						
MONO GRAPH		SMD			Conf.Int.	
		<i>M difference</i>	<i>g_{av}</i>	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair 4	1-2_RAN_JUNE_COLSEC - 2-1_JAN19_RAN_SPEED_COLSEC	-0,202	-0,998	0,061	-0,395	-0,008
Pair 8	1-2_RAN_NUMSEC - 2-1_JAN19_RAN_SPEED_NUMSEC	-0,297	-0,553	0,056	-0,474	-0,120
Pair 14	1-2_PHON_FLUENCY - 2-1_JAN_PHON_FLUENCY	-5,500	-0,499	1,443	-10,093	-0,907
BI GRAPH		SMD			Conf.Int.	
		<i>M difference</i>	<i>g_{av}</i>	<i>StdErr</i>	<i>Lower</i>	<i>Upper</i>
Pair 9	1-2_RAN_NUM_ERR_PERC - 2-1_JAN19_RAN_NUM_ERR_PERC	0,021	0,000	0,001	0,014	0,027

Table C.16. Within-subject SMD (g_{av}) in mono- and bi- children between Post-intervention 2 and Follow-up 2

Appendix D

Study 4

Deficit	Associated variable	<i>Fisher's coeff.</i>	<i>df</i>	<i>p</i>	<i>Cramer's V (p)</i>
Decoding speed	ADHD	4.042	1	0.044	0.287 (0.042)
	Social deprivation	11.416	1	0.003	0.483 (0.003)
	Familiarity with LD	4.636	1	0.031	0.308 (0.030)
Decoding accuracy	Social deprivation	10.967	1	0.002	0.473 (0.002)
	Familiarity with LD	5.940	1	0.022	0.348 (0.022)
	Family education	7.242	1	0.021	0.384 (0.024)
Reading global comprehension	ADHD	3.868	1	0.049	0.281 (0.047)
	Familiarity with LD	5.872	1	0.027	0.346 (0.027)
Reading central comprehension	Social deprivation	6.635	1	0.015	0.372 (0.015)
	Familiarity with LD	8.926	1	0.007	0.431 (0.007)
	Family education	4.579	1	0.037	0.341 (0.058)
Reading specific comprehension	Social deprivation	3.424	1	0.055	0.267 (0.062)
	Familiarity with LD	8.926	1	0.007	0.431 (0.007)
	Bilingualism	3.995	1	0.053	0.298 (0.053)
Word writing	LI	6.000	1	0.033	0.354 (0.033)
	ID	4.174	1	0.041	0.295 (0.039)
	ADHD	7.273	1	0.007	0.389 (0.011)
	Social deprivation	17.308	1	0.000	0.600 (0.000)
	Familiarity with LD	9.344	1	0.003	0.441 (0.003)
	Family education	8.895	1	0.009	0.434 (0.009)
Non-word writing	LI	6.646	1	0.025	0.372 (0.025)
	ID	12.209	1	0.007	0.504 (0.007)
	ADHD	5.264	1	0.051	0.331 (0.051)
	Social deprivation	18.052	1	0.000	0.613 (0.000)
	Familiarity with LD	3.841	1	0.050	0.283 (0.048)
Text writing	Social deprivation	13.000	1	0.000	0.520 (0.000)
	Familiarity with LD	8.100	1	0.006	0.411 (0.006)

Table D.1. Significant associations between impaired abilities and children's individual characteristics.

Post hoc tests showed that children are significantly more likely

- to have poor decoding speed when they have ADHD, conditions of social deprivation, or familiarity with LD;
- to have poor decoding accuracy when they live a condition of social deprivation, familiarity with LD, or their parents have a lower education level;

- to have poor reading global comprehension when they have ADHD and familiarity with LD; in particular, poor central comprehension is more associated with social deprivation, familiarity with LD, or parents' lower education level, whereas poor specific comprehension is associated with social deprivation, familiarity with LD, or bilingualism;
- to have poor word writing when they have LI, mild ID, ADHD, social deprivation, familiarity with LD, or their parents have a lower education level;
- to have poor non-word writing when they have LI, a mild ID, ADHD, social deprivation, or familiarity with LD;
- to have poor text writing when they have social deprivation and familiarity with LD.

Table D.2 shows significant results of the post hoc test.

Impairment	Variable	Likelihood		Z
		With deficit	Without deficit	
Decoding speed	ADHD	33.3%	6.8%	2.0
	Social deprivation	83.3%	18.2%	3.4
	Familiarity with LD	50.0%	13.6%	2.2
Decoding accuracy	Social deprivation	56.3%	11.8%	3.3
	Familiarity with LD	37.5%	8.8%	2.5
	Family education	Primary = 56.3% University = 0%	Primary = 23.5% University = 23.5%	Primary = 2.3 University = - 2.1
Reading global comprehension	ADHD	25.0%	5.3%	2.0
	Familiarity with LD	41.7%	10.5%	2.4
Reading central comprehension	Social deprivation	53.8%	16.7%	2.6
	Familiarity with LD	46.2%	8.3%	3.0
	Family education	Primary = 61.5%	Primary = 25%	Primary = 2.4
Reading specific comprehension	Social deprivation	46.2%	19.4%	1.9
	Familiarity with LD	46.2%	8.3%	3.0
Reading specific comprehension	Bilingualism	15.4%	47.2%	- 2.0
Word writing	LI	28.6%	3.6%	2.5
	ID	14.4%	0.0%	2.1
	ADHD	23.8%	0.0%	2.7
	Social deprivation	57.1%	3.6%	4.2
	Familiarity with LD	38.1%	3.6%	3.1
	Family education	Primary = 57.1%	Primary = 17.9%	2.9
Non-word writing	LI	40.0%	7.7%	2.6
	ID	30%	0.0%	3.5
	ADHD	30%	5.1%	2.3
	Social deprivation	80%	12.8%	4.3
	Familiarity with LD	40%	12.8%	2
Text writing	Social deprivation	46.4%	0.0%	3.6
	Familiarity with LD	32.1%	0.0%	2.9

Table D.2. Post hoc significant differences between proportions of different individual characteristics found in children with/without different impairments (Bonferroni, $p \leq 0.05$)

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>Conf. Interval</i>	
									Lower	Upper
Decoding speed deficit	Social deprivation	6.343	1	48	0.015	2.975	1.181	2.518	0.600	5.351
Decoding accuracy deficit	Social deprivation	5.513	1	47	0.023	1.846	0.786	2.348	0.264	3.428
Word writing deficit	LI	4.257	1	48	0.045	2.581	1.251	2.063	0.063	5.099
	Social deprivation	9.554	1	46	0.003	3.568	1.154	3.091	1.244	5.892
Non-word writing deficit	Social deprivation	11.053	1	47	0.002	3.134	0.943	3.325	1.238	5.031

Table D.3. Significant effects of individual characteristics on literacy impairments in 2nd and 3rd grade.

Dependent variable	Predictor	Predicted proportions		
		Global accuracy (Akaike)	Without impairment	With impairment
Decoding speed deficit	Digit span forward + backward	88.1%	97.5%	7.1%
Decoding accuracy deficit	Digit span forward + word span	76.9%	91.3%	45.2%
	Letter visual discrimination	73.1%	92.8%	34.3%
Reading comprehension global deficit	Digit span backward	85.1%	90.1%	69.7%
	Graphic reproduction errors (2 semicircles)	86.5%	89.5%	78.6%
Reading comprehension central deficit	Digit span backward	79.1%	91.8%	44.4%
	Non-word repetition	86.0%	97.8%	33.3%
	Graphic reproduction errors (2 semicircles)	77.9%	86.5%	56.7%
Reading comprehension specific deficit	Digit span backward	88.8%	94.9%	72.2%
Word writing deficit	Digit span forward + word span	77.6%	89.6%	61.4%
	Non-word repetition	72.8%	93.2%	36.0%

	Graphic reproduction of 2 semicircles	77.9%	87.7%	66.0%
Non-word writing deficit	Digit span forward	84.3%	94.5%	40.0%
	Graphic reproduction of 1 circle + letter visual discrimination	85.6%	97.5%	43.5%
Text writing deficit	Digit span backward	72.4%	53.6%	85.9%

Table D.4. Global predictive accuracy and proportion of cases predicted by each significant covariate.

Impairment	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>Conf. Interval</i>	
									Lower	Upper
Decoding speed	Digit span forward	3.838	1	131	0.052	0.731	0.373	1.959	- 0.007	1.740
	Digit span backward	12.798	1	131	0.000	2.957	0.827	3.577	1.322	4.592
Decoding accuracy	Digit span forward	4.429	1	116	0.037	0.734	0.349	2.105	0.043	1.426
	Word span	6.858	1	108	0.010	1.004	0.383	2.619	0.244	1.764
	Letter visual discrimination errors	8.436	1	102	0.005	- 0.367	0.126	- 2.904	- 0.617	- 0.116
Reading global comprehension (central + c)	Digit span backward	6.139	1	132	0.014	0.920	0.371	2.478	0.185	1.564
	Graphic reproduction errors (2 semicircles)	5.826	1	102	0.018	- 0.332	0.138	- 2.414	- 0.605	- 0.059
Reading comprehension central	Digit span backward	12.374	1	132	0.001	1.296	0.363	3.568	0.577	2.014
	Non-word repetition	26.962	1	112	0.000	0.004	0.001	5.193	0.003	0.006
	Graphic reproduction errors (2 semicircles)	4.900	1	102	0.029	- 0.267	0.121	- 2.214	- 0.507	- 0.028
Reading specific comprehension	Digit span backward	8.137	1	111	0.005	0.985	0.345	2.852	0.301	1.669
Word writing	Digit span forward	9.913	1	62	0.003	1.298	0.412	3.148	0.474	2.122
	Word span	4.664	1	74	0.034	0.867	0.401	2.160	0.067	1.666
Word writing	Non-word repetition	5.965	1	112	0.016	0.062	0.025	2.442	0.012	0.112
	Graphic reproduction errors (2 semicircles)	9.232	1	86	0.003	- 0.350	0.115	- 3.038	- 0.579	- 0.121
Non-word writing deficit	Digit span forward	11.354	1	97	0.001	1.279	0.380	3.370	0.526	2.032
	Graphic reproduction errors (1 semicircle)	3.819	1	101	0.053	- 0.357	0.183	- 1.954	- 0.720	0.005

	Letter visual discrimination errors	6.987	1	101	0.010	- 0.414	0.157	- 2.643	- 0.725	- 0.103
Text writing	Digit span backward	10.028	1	132	0.002	0.793	0.250	3.167	0.297	1.288

Table D.5. Longitudinal predictive effects of memory span, phonological, and visual memory skills from KG to the 3rd grade on diverse impairments in 2nd and 3rd grade.

Deficit	Predictor	Coeff	StdErr	t	p	Conf. Int. (95%)	
						Lower	Upper
Decoding speed	Digit span backward * KG-2	3.019	1.169	2.584	0.011	0.707	5.331
	Digit span backward * 1-1	2.604	0.652	3.993	0.000	1.314	3.895
	Digit span backward * 2-1	2.118	0.586	3.617	0.000	0.959	3.277
	Digit span backward * 3-1	2.400	0.531	4.523	0.000	1.350	3.450
Decoding accuracy	Digit span forward * KG-2	1.557	0.426	3.653	0.000	0.714	2.400
	Digit span forward * 1-1	1.426	0.363	3.926	0.000	0.707	2.144
	Digit span forward * 2-1	1.236	0.328	3.763	0.000	0.586	1.886
	Digit span forward * 3-1	1.297	0.355	3.653	0.000	0.595	1.999
	Word span * KG-2	1.503	0.403	3.729	0.000	0.706	2.300
	Word span * 1-1	1.555	0.373	4.172	0.000	0.818	2.292
	Word span * 2-1	1.304	0.345	3.774	0.000	0.620	1.987
	Word span * 3-1	1.744	0.477	3.658	0.000	0.801	2.687
Reading global comprehension	Digit span backward * KG-2	2.191	0.734	2.985	0.004	0.734	3.649
	Digit span backward * 1-1	1.286	0.434	2.966	0.004	0.426	2.146
	Digit span backward * 2-1	1.071	0.371	2.885	0.005	0.336	1.806
	Digit span backward * 3-1	0.882	0.425	2.077	0.045	0.020	1.745
Reading central comprehension	Digit span backward * KG-2	3.074	0.735	4.183	0.000	1.620	4.528
	Digit span backward * 1-1	1.811	0.430	4.214	0.000	0.961	2.661
	Digit span backward * 2-1	1.552	0.379	4.095	0.000	0.802	2.301
	Digit span backward * 3-1	1.308	0.406	3.223	0.002	0.505	2.110
Reading specific comprehension	Digit span backward * KG-2	1.629	0.565	2.880	0.005	0.504	2.754
	Digit span backward * 1-1	1.148	0.393	2.923	0.004	0.370	1.927
	Digit span backward * 2-1	0.945	0.339	2.793	0.006	0.275	1.616
Word writing	Digit span forward * KG-2	2.115	0.481	4.393	0.000	1.160	3.071
	Digit span forward * 1-1	1.885	0.399	4.720	0.000	1.092	2.677
	Digit span forward * 2-1	1.679	0.375	4.478	0.000	0.934	2.423
	Digit span forward * 3-1	1.778	0.406	4.383	0.000	0.964	2.592
	Graphic reproduction errors (2 semicircles) * KG-2	-0.349	0.130	-2.687	0.009	-0.608	-0.090

	Graphic reproduction errors (2 semicircles) * 1-1	-0.509	0.170	-3.002	0.004	-0.847	-0.171
Non-word writing	Digit span forward * KG-2	1.506	0.492	3.062	0.003	0.532	2.480
	Digit span forward * 1-1	1.386	0.412	3.361	0.001	0.569	2.203
	Digit span forward * 2-1	1.178	0.375	3.144	0.002	0.436	1.921
	Digit span forward * 3-1	1.372	0.396	3.464	0.001	0.581	2.163
Text writing	Digit span backward * KG-2	1.458	0.428	3.405	0.001	0.611	2.305
	Digit span backward * 1-1	1.017	0.304	3.343	0.001	0.415	1.619
	Digit span backward * 2-1	0.948	0.283	3.349	0.001	0.388	1.508

Table D.6. Significant combined effects of each predictor and administration time on different deficits

Dependent variable	Fixed effect	Predictor	Est	StdErr	df	p	95% Conf. Interv.	
							Limite inferiore	Limite superiore
Decoding accuracy	Letter discrimination errors	BI	-1.00E-16	0.000	102	0.001	-2.00E-16	-1.00E-16
Reading global comprehension	Graphic reproduction errors (2 semicircles)	BI	-3.000E-16	1.000E-16	102.000	0.000	-5.000E-16	-2.000E-16
		MONO	3.000E-16	2.000E-16	102.000	0.034	0.000	7.000E-16
Text writing	Digit span backward	BI * 2nd grade	-4.000E-16	1.000E-16	132.000	3.085E-5	-6.000E-16	-2.000E-16
		MONO * 2nd grade	5.000E-16	1.000E-16	132.000	7.458E-6	3.000E-16	7.000E-16

Table D.7. Significant random effects emerged on literacy impairments, with short-term memory and visual discrimination skills as main effects.

Dependent variable	Predictor	Predicted proportions		
		Global accuracy (Akaike)	Without impairment	With impairment
Decoding speed	Productive vocabulary	84.8%	95.0%	16.7%
	Syntactic production of DOcls	91.3%	97.5%	50.0%
Decoding accuracy	Productive vocabulary	71.7%	86.7%	43.8%
	Syntactic production of DOcls	73.9%	86.7%	50.0%
Reading comprehension central	Productive vocabulary	71.7%	87.9%	30.8%
	Syntactic comprehension	80.0%	90.6%	53.8%

Reading comprehension specific	Productive vocabulary	87.0%	90.0%	76.9%
Word writing	Productive vocabulary	80.4%	80.0%	81.0%
	Syntactic comprehension	75.6%	79.2%	71.4%
Non-word writing	Productive vocabulary	87.0%	97.2%	50.0%
	Syntactic comprehension	82.2%	91.4%	50.0%
Text writing	Productive vocabulary	82.6%	83.3%	82.1%

Table D.8. Predictive global accuracy and proportions of cases without/with impairment predicted by each significant covariate.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>Conf. Interval</i>	
									Lower	Upper
Decoding speed	Productive vocabulary	4.982	1	44	0.031	0.172	0.077	2.232	0.017	0.327
	Syntactic production of DOcls	6.327	1	44	0.016	4.300	1.700	2.515	0.855	7.745
Decoding accuracy	Productive vocabulary	6.942	1	44	0.012	0.149	0.056	2.635	0.035	0.262
	Syntactic production of DOcls	6.264	1	44	0.016	2.976	1.189	2.503	0.580	5.373
Reading central comprehension	Productive vocabulary	4.654	1	44	0.036	0.123	0.057	2.157	0.008	0.238
	Syntactic comprehension	4.164	1	44	0.047	0.238	0.117	2.041	0.003	0.474
Reading specific comprehension	Productive vocabulary	5.398	1	44	0.025	0.232	0.100	2.323	0.031	0.433
Word dictation	Productive vocabulary	6.547	1	44	0.014	0.139	0.054	2.559	0.030	0.249
	Syntactic comprehension	8.216	1	43	0.006	0.373	0.130	2.866	0.111	0.635
Non-word dictation	Productive vocabulary	8.740	1	44	0.005	0.270	0.091	2.956	0.086	0.455
	Syntactic comprehension	6.038	1	43	0.018	0.331	0.135	2.457	0.059	0.603
Text dictation	Productive vocabulary	5.990	1	44	0.018	0.147	0.060	2.447	0.026	0.268

Table D.9. Significant effects of vocabulary and syntactic skills on diverse literacy impairments in 2nd and 3rd grade.

Dependent variable	Predictor	Global accuracy (Akaike)	Predicted proportions	
			Without impairment	With impairment
Decoding accuracy	Initial syllable recognition	77.8%	91.2%	45.8%
Reading global comprehension	Initial syllable recognition	94.8%	98.5%	70.0%
Reading central comprehension	Initial syllable recognition	85.7%	96.8%	35.7%
Reading specific comprehension	Initial syllable recognition	90.9%	98.5%	50.0%
Word writing	Initial syllable recognition	81.9%	93.9%	64.7%
Non-word writing	Initial syllable recognition	86.7%	95.5%	50.0%
Text writing	Initial syllable recognition	72.3%	80.5%	64.3%

Table D.10. Predictive accuracy of recognition of the initial syllable on each impairment. Predictive global accuracy and proportions of cases without/with impairment predicted by each significant covariate.

Dependent variable	Predictor	<i>F</i>	<i>df1</i>	<i>df2</i>	<i>p</i>	<i>Coeff.</i>	<i>StdErr</i>	<i>t</i>	<i>Conf. Interval</i>	
									Lower	Upper
Decoding accuracy	Initial syllable recognition	12.485	1	79	0.001	0.264	0.075	3.533	0.115	0.412
Global reading comprehension	Initial syllable recognition	8.667	1	75	0.004	0.426	0.145	2.944	0.138	0.714
Central reading comprehension	Initial syllable recognition	9.290	1	75	0.003	0.250	0.082	3.048	0.086	0.413
Specific reading comprehension	Initial syllable recognition	12.118	1	61	0.001	0.280	0.080	3.481	0.119	0.440
Word writing	Initial syllable recognition	22.183	1	81	0.000	0.459	0.097	4.710	0.265	0.653
Non-word writing	Initial syllable recognition	4.256	1	81	0.042	0.166	0.081	2.063	0.006	0.327
Text writing	Initial syllable recognition	11.184	1	75	0.001	0.231	0.069	3.344	0.093	0.368

Table D.11. Longitudinal effects of recognition of the initial syllable in the two KG administrations on different literacy impairments found in 2nd and 3rd grade.

Dependent variable	Predictor	Est	StdErr	df	p	95% Conf. Interv.	
						Lower	Upper
Reading global comprehension	BI	-3.19E-14	8.70E-15	75	0.000	-4.94E-14	-1.45E-14
	MONO	3.19E-14	1.46E-14	75	0.032	2.80E-15	6.11E-14
Reading central comprehension	MONO * 2nd grade	3.000E-16	1.000E-16	75	2.227E-5	2.000E-16	5.000E-16
	BI * 3rd grade	-2.000E-16	0.000	75	2.682E-5	-3.000E-16	-1.000E-16
Word writing	BI	-2.000E-15	4.000E-16	81	4.805E-6	-2.900E-15	-1.200E-15
	MONO	2.000E-15	7.000E-16	81	0.006	6.000E-16	3.500E-15
Non-word writing	MONO * 2nd grade	3.640E-14	1.280E-14	81	0.006	1.100E-14	6.190E-14
	BI * 3rd grade	-3.980E-14	6.300E-15	81	1.654E-8	-5.240E-14	-2.710E-14

Table D.12. Significant random effects of bilingualism and grade on different literacy skills, with initial syllable recognition as a longitudinal fixed effect.

Dependent variable	Predictor	Predicted proportions		
		Global accuracy (Akaike)	Without deficit	With deficit
Decoding speed deficit	Initial syllable deletion + Phonemic verbal fluency	89.6%	100.0%	0%
Decoding accuracy deficit	Final syllable deletion + phonemic verbal fluency	76.4%	92.8%	42.6%
Reading comprehension global deficit	Phoneme blending + final syllable deletion + phonemic verbal fluency	75.7%	99.0%	15.0%
Reading comprehension central deficit	Final syllable deletion	70.8%	97.0%	9.3%
Reading comprehension specific deficit	Final syllable deletion	72.9%	96.0%	20.5%
Word writing deficit	Phoneme blending + final syllable deletion + phonemic verbal fluency	75.0%	67.5%	84.4%
Non-word writing deficit	Initial syllable deletion + final syllable deletion + phonemic verbal fluency	86.1%	97.4%	39.3%
Text writing deficit	Phoneme blending	68.8%	30.4%	93.2%

Table D.13. Global and partial accuracy of predictive effects of analytical metaphonological skills on different literacy impairments. .

Deficit	Predictor	Coeff.	Std.Err.	t	p	Conf. Int. (95%)	
						Lower	Upper
Decoding speed	Initial syllable deletion * 1-2	0.426	0.123	3.475	0.001	0.179	0,672
	Initial syllable deletion * 2-1	0.197	0.052	3.769	0.000	0.093	0.300
	Initial syllable deletion * 2-2	0.196	0.083	2.346	0.035	0.016	0.376
	Initial syllable deletion * 3-1	0.186	0.084	2.208	0.051	-0.001	0,374
	Phonemic verbal fluency * 1-1	0.537	0.201	2.674	0.008	0.140	0,935
	Phonemic verbal fluency * 1-2	0.425	0.114	3.718	0.000	0.199	0,651
	Phonemic verbal fluency * 2-1	0.294	0.065	4.529	0.000	0.166	0,422
	Phonemic verbal fluency * 2-2	0.318	0.100	3.183	0.002	0.120	0,516
	Phonemic verbal fluency * 3-1	0.241	0.080	3.001	0.003	0.082	0.400
Decoding accuracy	Final syllable deletion * 1-1	0.739	0.195	3.797	0.000	0.351	1,128
	Final syllable deletion * 1-2	0.392	0.081	4.806	0.000	0.23	0,554
	Final syllable deletion * 2-1	0.255	0.052	4.884	0.000	0.152	0.359
	Final syllable deletion * 2-2	0.191	0.067	2.872	0.007	0.056	0,327
	Final syllable deletion * 3-1	0.201	0.068	2.973	0.007	0.062	0,341
	Phonemic verbal fluency * 1-1	0.329	0.074	4.463	0.000	0.182	0.476
	Phonemic verbal fluency * 1-2	0.281	0.053	5.298	0.000	0.175	0,386
	Phonemic verbal fluency * 2-1	0.205	0.040	5.134	0.000	0.126	0,284
	Phonemic verbal fluency * 2-2	0.179	0.064	2.812	0.011	0.046	0,312
	Phonemic verbal fluency * 3-1	0.159	0.057	2.802	0.013	0.039	0,279
Reading global comprehension	Phoneme blending * 1-1	0.190	0.092	2.070	0.040	0.009	0.372
	Final syllable deletion * 1-1	0.464	0.143	3.251	0.002	0.177	0,751
	Final syllable deletion * 1-2	0.204	0.060	3.367	0.001	0.083	0,324
	Final syllable deletion * 2-1	0.133	0.043	3.075	0.003	0.048	0,219
	Phonemic verbal fluency * 1-1	0.212	0.071	2.982	0.004	0.070	0.353
	Phonemic verbal fluency * 1-2	0.08	0.036	2.196	0.031	0.007	0,152
	Phonemic verbal fluency * 2-1	0.077	0.035	2.22	0.029	0.008	0,146
Reading central comprehension	Final syllable deletion * 1-1	0.563	0.154	3.65	0.001	0.254	0.871
	Final syllable deletion * 1-2	0.229	0.060	3.82	0.000	0.11	0,348
	Final syllable deletion * 2-1	0.156	0.043	3.582	0.000	0.07	0.242
Reading comprehension specific	Final syllable deletion * 1-1	0.721	0.300	2.4	00.02	0.119	1,323
	Final syllable deletion * 1-2	0.346	0.076	4.543	0.000	0.195	0,498
	Final syllable deletion * 2-1	0.254	0.058	4.342	0.000	0.138	0.369
Word writing	Phon_blend*1-1	1.074	0.541	1.985	0.053	-0.015	2,162
	Phon_blend*1-2	0.713	0.147	4.862	0.000	0.422	1,004

	Phon_blend*2-1	0.618	0.129	4.786	0.000	0.361	0,874
	Phon_blend*3-1	0.523	0.128	4.098	0.000	0.268	0,777
	Final syllable deletion * 1-1	0.688	0.192	3.582	0.001	0.304	1,071
	Final syllable deletion * 1-2	0.386	0.081	4.754	0.000	0.224	0,548
	Final syllable deletion * 2-2	0.185	0.077	2.405	0.021	0.03	0,34
	Final syllable deletion * 3-1	0.196	0.074	2.64	0.012	0.046	0,346
	Phonemic verbal fluency * 1-1	0.293	0.064	4.586	0.000	0.166	0.420
	Phonemic verbal fluency * 1-2	0.202	0.049	4.107	0.000	0.105	0,3
	Phonemic verbal fluency * 2-1	0.175	0.040	4.434	0.000	0.097	0,254
	Phonemic verbal fluency * 2-2	0.138	0.058	2.383	0.025	0.019	0,257
	Phonemic verbal fluency * 3-1	0.119	0.056	2.129	0.046	0.002	0,236
Non-word writing	Initial syllable deletion * 1-1	1.694	0.598	2.831	0.008	0.481	2,907
	Initial syllable deletion * 1-2	0.292	0.070	4.198	0.000	0.153	0,432
	Initial syllable deletion * 2-1	0.193	0.043	4.534	0.000	0.108	0,278
	Initial syllable deletion * 2-2	0.175	0.067	2.629	0.019	0.033	0,317
	Initial syllable deletion * 3-1	0.159	0.067	2.381	0.036	0.013	0.306
	Final syllable deletion * 1-1	0.583	0.171	3.413	0.001	0.24	0,925
	Final syllable deletion * 1-2	0.286	0.081	3.547	0.001	0.125	0,446
	Final syllable deletion * 2-1	0.192	0.048	3.99	0.000	0.097	0.288
	Final syllable deletion * 2-2	0.157	0.067	2.35	0.029	0.018	0,296
	Final syllable deletion * 3-1	0.176	0.071	2.492	0.024	0.026	0,327
Text writing	Phoneme blending * 1-1	0,395	0,071	5,531	0	0,253	0,536
	Phoneme blending * 1-2	0,227	0,04	5,616	0	0,146	0,308
	Phoneme blending * 2-1	0,188	0,036	5,213	0	0,116	0,261

Table D.14. Significant combined effects of analytical metaphonological skills and administration time on literacy deficits

Dependent variable	Predictor	Global accuracy (Akaike)	Predicted proportions	
			Without impairment	With impairment
Decoding speed	Picture RAN speed	89.9%	98.5%	5.0%
Decoding accuracy	Picture RAN errors	77.0%	90.2%	45.3%
Reading comprehension global	Rhyme recognition	86.3%	89.7%	76.3%
Reading comprehension central	Rhyme recognition	78.2%	91.1%	44.6%
Reading comprehension specific	Rhyme recognition	89.0%	90.6%	78.1%
Word writing deficit	Picture RAN errors + rhyme recognition	72.4%	74.0%	69.9%
Non-word writing	Picture RAN speed + errors	88.9%	97.2%	48.6%
Text writing	Picture RAN errors	74.7%	90.6%	62.0%

Table D.15. Global and partial accuracy of predictive effects of the RAN of pictures on different literacy deficits.

						Conf. Interv. (95%)	
Deficit	Predictor	Coeff.	Std.Err.	t	p	Lower	Upper
Decoding speed	Picture RAN speed * KG-1	25.848	3.806	6.791	0.000	18.301	33.396
	Picture RAN speed * KG-2	21.282	2.938	7.243	0.000	15.466	27.098
	Picture RAN speed * 1-1	12.498	4.522	2.764	0.008	3.401	21.595
	Picture RAN speed * 1-2	10.867	1.705	6.375	0.000	7.494	14.240
	Picture RAN speed * 2-1	9.172	1.264	7.258	0.000	6.669	11.674
	Picture RAN speed * 2-2	9.578	1.790	5.352	0.000	5.924	13.231
	Picture RAN speed * 3-1	8.129	1.173	6.930	0.000	5.753	10.504
Decoding accuracy	Picture RAN errors * KG-1	- 8.974	2.538	- 3.535	0.001	- 13.977	- 3.97
	Picture RAN errors * KG-2	- 13.122	3.223	- 4.071	0.000	- 19.475	- 6.768
	Picture RAN errors * 1-1	- 19.526	8.344	- 2.340	0.020	- 35.975	- 3.076
	Picture RAN errors * 1-2	- 26.770	8.069	- 3.318	0.001	- 42.677	- 10.863
	Picture RAN errors * 2-1	- 32.520	9.367	- 3.472	0.001	- 50.985	- 14.054
	Picture RAN errors * 2-2	- 18.582	7.553	- 2.460	0.015	- 33.471	- 3.693
Reading global comprehension	Rhyme recognition * KG-1	0.368	0.072	5.084	0.000	0.224	0.512
	Rhyme recognition * KG-2	0.267	0.059	4.512	0.000	0.15	0.385
	Rhyme recognition * 1-1	0.295	0.068	4.309	0.000	0.158	0.432
	Rhyme recognition * 1-2	0.247	0.062	3.971	0.000	0.124	0.37
	Rhyme recognition * 2-1	0.210	0.057	3.663	0.000	0.096	0.323
	Rhyme recognition * 2-2	0.244	0.070	3.482	0.001	0.102	0.386
	Rhyme recognition * 3-1	0.232	0.070	3.316	0.002	0.089	0.375
Reading central comprehension	Rhyme recognition * KG-1	0.335	0.074	4.519	0.000	0.188	0.483
	Rhyme recognition * KG-2	0.262	0.054	4.838	0.000	0.155	0.369
	Rhyme recognition * 1-1	0.270	0.062	4.358	0.000	0.147	0.393
	Rhyme recognition * 1-2	0.243	0.054	4.472	0.000	0.135	0.350
	Rhyme recognition * 2-1	0.207	0.050	4.179	0.000	0.109	0.306
	Rhyme recognition * 2-2	0.206	0.066	3.115	0.004	0.072	0.341
	Rhyme recognition * 3-1	0.229	0.065	3.505	0.002	0.095	0.363
Reading specific comprehension	Rhyme recognition * KG-1	0.378	0.072	5.219	0.000	0.234	0.521

	Rhyme recognition * KG-2	0.286	0.055	5.206	0.000	0.178	0.395
	Rhyme recognition * 1-1	0.264	0.060	4.399	0.000	0.145	0.384
	Rhyme recognition * 1-2	0.253	0.057	4.408	0.000	0.14	0.367
	Rhyme recognition * 2-1	0.216	0.051	4.191	0.000	0.114	0.317
	Rhyme recognition * 2-2	0.259	0.070	3.710	0.001	0.118	0.401
	Rhyme recognition * 3-1	0.233	0.071	3.272	0.003	0.086	0.379
Word writing	Rhyme recognition * KG-1	0.228	0.064	3.564	0.001	0.101	0.355
	Rhyme recognition * KG-2	0.199	0.051	3.914	0.000	0.099	0.300
	Rhyme recognition * 1-1	0.190	0.054	3.551	0.001	0.084	0.297
	Rhyme recognition * 1-2	0.195	0.050	3.936	0.000	0.097	0.293
	Rhyme recognition * 2-1	0.155	0.044	3.509	0.001	0.068	0.242
	Rhyme recognition * 2-2	0.203	0.065	3.136	0.003	0.072	0.334
	Rhyme recognition * 3-1	0.218	0.066	3.293	0.003	0.083	0.354
	Picture RAN errors * KG-1	- 11.330	2.951	- 3.839	0.000	- 17.256	- 5.404
	Picture RAN errors * KG-2	- 14.327	4.011	- 3.572	0.001	- 22.373	- 6.282
	Picture RAN errors * 1-1	- 35.598	9.115	- 3.906	0.000	- 53.815	- 17.382
	Picture RAN errors * 1-2	- 32.840	9.361	- 3.508	0.001	- 51.513	- 14.167
	Picture RAN errors * 2-1	- 35.708	10.649	- 3.353	0.001	- 56.943	- 14.473
	Picture RAN errors * 2-2	- 27.974	12.205	- 2.292	0.038	- 54.219	- 1.729
	Picture RAN speed * KG-1	9.557	4.070	2.348	0.020	1.533	17.581
Non-word writing	Picture RAN speed * KG-2	9.137	4.332	2.109	0.036	0.598	17.677
	Picture RAN speed * 1-1	6.314	1.596	3.956	0.000	3.168	9.461
	Picture RAN speed * 1-2	6.103	1.646	3.708	0.000	2.858	9.348
	Picture RAN speed * 2-1	4.911	1.311	3.747	0.000	2.328	7.495
	Picture RAN speed * 2-2	5.972	1.587	3.763	0.000	2.844	9.100
	Picture RAN speed * 3-1	5.230	1.218	4.292	0.000	2.828	7.632
	Picture RAN errors * KG-1	- 12.982	3.825	- 3.394	0.001	- 20.643	- 5.321
	Picture RAN errors * KG-2	- 14.698	5.328	- 2.759	0.008	- 25.407	- 3.989
	Picture RAN errors * 1-1	- 38.152	8.271	- 4.613	0.000	- 54.648	- 21.657
	Picture RAN errors * 1-2	- 30.261	9.484	- 3.191	0.002	- 49.177	- 11.346
	Picture RAN errors * 2-1	- 42.880	14.153	- 3.030	0.004	- 71.151	- 14.61
	Picture RAN errors * 2-2	- 22.525	5.961	- 3.778	0.001	- 34.953	- 10.096

Text writing	Picture RAN errors * KG-1	- 9.200	2.912	- 3.159	0.003	- 15.066	- 3.334
	Picture RAN errors * KG-2	- 12.415	4.151	- 2.991	0.004	- 20.748	- 4.082
	Picture RAN errors * 1-1	- 29.038	9.472	- 3.066	0.003	- 48.01	- 10.066
	Picture RAN errors * 1-2	- 35.187	10.522	- 3.344	0.001	- 56.16	- 14.214
	Picture RAN errors * 2-1	- 37.139	10.487	- 3.541	0.001	- 58.027	- 16.250

Table D.16. Significant effects of the combination of picture RAN and rhyme recognition, respectively, with testing time on literacy impairments found in 2nd and 3rd grade.

Dependent variable	Predictor	Global accuracy (Akaike)	Predicted proportions	
			Without impairment	With impairment
Decoding accuracy deficit	Color RAN speed	72.5%	88.2%	34.2%
Reading comprehension global deficit	Color + digit RAN speed	86.8%	92.9%	66.7%
Reading comprehension central deficit	Color + digit RAN speed	80.2%	91.8%	47.1%
Word writing deficit	Color RAN speed + errors	76.9%	85.7%	64.2%
Non-word writing deficit	Color RAN speed + digit RAN errors	88.2%	100%	21.1%

Table D.17. Global and partial accuracy of predictive effects of color and digit RAN on different literacy impairments..

Dependent variable	Predictor	Coeff.	Std.Err.	t	p	Conf. Interv. (95%)	
						Lower	Upper
Decoding accuracy	Color RAN speed * 1-1	5.831	1.442	4.043	0.000	2.977	8,685
	Color RAN speed * 1-2	5.685	1.243	4.575	0.000	3.226	8,144
	Color RAN speed * 2-1	4.350	1.051	4.140	0.000	2.270	6,429
	Color RAN speed * 2-2	3.768	1.310	2.876	0.005	1.175	6,361
	Color RAN speed * 3-1	3.647	1.209	3.016	0.003	1.254	6,039
Reading central comprehension	Color RAN speed * 1-1	3.364	1.115	3.017	0.003	1.151	5,577
	Color RAN speed * 1-2	2.841	1.045	2.719	0.008	0.768	4,915
	Color RAN speed * 2-1	1.933	0.859	2.251	0.026	0.231	3,634
	Digit RAN speed * 1-1	2.616	0.987	2.65	0.009	0.662	4,569
	Digit RAN speed * 1-2	1.778	0.759	2.342	0.021	0.275	3,281
	Digit RAN speed * 2-1	1.145	0.594	1.929	0.056	-0.03	2,32
Word writing	Color RAN speed * 1-1	4.428	1.184	3.739	0.000	2.081	6,774
	Color RAN speed * 1-2	4.080	1.082	3.77	0.000	1.934	6,226

	Color RAN speed * 2-1	3.430	0.929	3.693	0.000	1.588	5,272
	Color RAN speed * 2-2	3.111	1.267	2.455	0.019	0.548	5,674
	Color RAN speed * 3-1	3.025	1.139	2.657	0.012	0.707	5,344
	Color RAN errors * 1-2	-25.539	10.375	-2.462	0.017	-46.381	-4,696
	Color RAN errors * 2-1	-43.478	13.914	-3.125	0.003	-71.19	-15,766
	Color RAN errors * 2-2	-42.131	16.547	-2.546	0.024	-77.841	-6,421
Non-word writing	Digit RAN errors * 1-2	-32.050	14.156	-2.264	0.025	-60.075	-4,025
	Digit RAN errors * 2-1	-42.390	15.675	-2.704	0.008	-73.423	-11,358

Table D.18. Significant predictive combined effects of color/digit RAN and administration time on different decoding deficits found in 2nd and 3rd grade.

Acknowledgments

I would like to express my deep gratitude to my Supervisor, Prof. Anna Cardinaletti, for her strenuous support and dedication in every phase of this challenging project, from the genesis to the write-up. Her dedicated advice helped me face every choice in the 3 years of my Ph.D.

My deep gratitude goes to my Co-Supervisor, Prof. Giulia Bencini, for her precious and targeted advice.

I am particularly grateful to the reviewers of this thesis, Prof. Mila Vulchanova and Prof. Daniela Traficante, whose valuable suggestions allowed me to question my methodological choices and to gain insight in discussing my data, thus increasing the added value of this work.

My special thanks to Prof. Daniela Traficante and Prof. Claudio Giuseppe Luzzatti for providing the still unpublished Italian version of the TILTAN test, which allowed me to collect a body of data that certainly deserves further analyses.

Special thanks to Prof. Anna Cardinaletti and Prof. Francesca Volpato, who shared the realization of the clitic-production test that has been here employed for the first time. A valuable contribution to my expertise for building this test was given by discussion with Prof. Laurice Tuller, Prof. Philippe Prévost, and Prof. Hélène Delage. I would like also to acknowledge the help of Sara Cerutti, who generously drew the test pictures.

Many thanks to Prof. Ianthi Tsimpli for her suggestions on testing.

Thanks to Dr. Giorgio Arcara and Dr. Gianluca Lebani for discussing with me some thorny statistical issues during the revision phase of the present thesis.

I would like finally to express my deepest gratitude to all the people who made this project possible.

First of all, the psychologist Letizia Moretti (Association Psicologia Insieme of Rome/Varese) who provided contact with the hosting Institute and generously gave her clinical advice and indispensable practical support during the administration of tests and interventions. The hosting Institute, in the person of the school counselor, the school principal and assistant principal, and the teachers, who patiently tolerated our presence during school hours and vividly discussed their pupils' outcomes.

Above all, I would like to thank the children, who enthusiastically participated in the scheduled activities, albeit demanding, and their families, who confidently gave their consent.

I would like eventually to acknowledge the help of a selected team of collaborators who assisted in the administration of tests and interventions after specific training: MDs in psychology Vanessa Costa, Brigida Fasano, Stefano Pugliese, Gianmarco Quigesi, Vanessa Saralli, and Debora Zacchi; graduate students in psychology Alice Gargiulo and Tania Nicolucci; and students in speech therapy Maria Cristina Bonaventura, Miriana Ciferri, Giorgia Gaudenzi, and Caterina Marchesi.

Errata corrige

The following amendments have been integrated into the present version compared to the digital version uploaded on April 9, 2021:

("wrong version" → "amended version")

- p.3, Abstract, line 20: "and memory" → "and short-term memory"
- p.3, Abstract, lines 25-26: "and experimented proposing" → "with"
- p. 3, lines 35-36: "whereas the other two groups did not increase significantly in the treated abilities" is erased
- p. 19, line 11: "Neither more proficient ((1) and (2)) nor less proficient ((3) and (4))" → "Neither less proficient ((1) and (2)) nor more proficient ((3) and (4))"
- p. 59, line 3: "This test is validated" → "This test is standardized"
- p. 59, line 6: "which is validated" → "which is standardized"
- p. 60, footnote 20: "See 2.1.3" → "See 2.3"
- p. 62, footnote 21: "See 2.1.3" → "See 2.3"
- p. 65, lines 9-10: the sentence "In addition, children were assessed on different tasks to examine their linguistic and non-linguistic abilities" moved to lines 15-16, between "made in class." and "The families"
- p. 66, line 15: "see sections 2.1.2 – 2.1.14" → "see sections 2.2 – 2.14"
- p. 67, last line: "in sections 2.1.2 – 2.1.14" → "in sections 2.2 – 2.14"
- p. 68, column 1, line 2: "forward/backward digit span forward and backward" → "forward/backward digit span"
- p. 68, column 1, line 8: "RAN of colors, pictures, colors, and digits" → "RAN of colors, pictures, and digits"
- p. 70, footnote 25, last line: "see 2.1.3" → "see 2.3"
- p. 71, Table 2 caption: "on non-word repetition." → "on non-word repetition and receptive vocabulary."
- p. 72, end of footnote 26: "(see 2.1.3)" → "(see 2.3)"
- p. 87, end of footnote 27: "see 2.1.9." → "see 2.9."
- p. 91, line 2 after Table 16: "in the middle of the 1st grade" → "at the end of the 1st grade"
- p. 91, end of line 3 after Table 16: "is non-significant." → "is marginally non-significant."
- p. 91, 3rd-to-last line: "(see 2.1.11)" → "(see 2.11)"
- p. 110, line 18: "reduced working memory" → "reduced verbal working memory"
- p. 116, line 8 after title: "(see 2.1.12)" → "(see 2.12)"
- p. 120, lines 3-4: "are" → "were"
- p. 122, line 2: "investigate" → "investigated"
- p.122, line 14: "in sections 2.1.2 – 2.1.14" → "in sections 2.2 – 2.14"
- p. 123, lines 10-11: "see Table B.1; Complete correlation tables are in Appendix B):" → "see Table B.1 in Appendix B)."
- p. 126, line 13: score → scores
- p. 133, line 6: " , as shown in Figure 24" is erased and replaced with " ,"
- p. 137, line 3: "predicts" → "predicted"
- p. 143, line 8: "test" → "text"
- p. 146, line 15: "emerge" → "emerged"
- p. 146, line 19: "2.1.14" → "2.14"
- p. 152, 5th-to-last line: "phonological route of reading" → "phonological route of writing"
- p. 156, footnote 47, end of 2nd-last line: "see footnote 1" → "see footnote 22"
- p. 158, column 1, line 3: "forward/backward digit span forward and backward" → "forward/backward digit span"
- p. 158, column 1, line 9: "RAN of colors, pictures, colors, and digits" → "RAN of colors, pictures, and digits"
- p. 158, column "1-2", subcolumn "CG": "6 (3+3)" is added to data lines 1, 2, 5, 7, and 17 (excluding title lines)

p. 158, column "2-1", subcolumn "CG": "6 (3+3)" is added to data lines 1, 2, 5, 7, 11, 14, 15, 16, and 17 (excluding title lines)

p. 159: title lines are added as in p. 158.

p. 159, column "1-2", subcolumn "CG": "6 (3+3)" is added to data lines 3 and 4

p.159, column "2-1": data line 6 is erased

p. 159, column "2-1", subcolumn CG: "6 (3+3)" is added to data lines 1, 2, 3, 4, and 7.

p. 168, line 1: "tests, intervention, the CG" → "tests, the CG"

p. 172, Figure 31 caption, end of line 2: "for" → "of"

p. 175, Table 27 caption: "on the 3 administrations" → "on the 2 administrations of the syntactic-comprehension test and on the 3 administrations"

p. 176, line 5: "advantage" → "disadvantage"

p. 179, line 4 after Figure 35: "table 7" → "Table 29"

p. 181, line 14: "see 2.1.9" → "see 2.9"

p. 188, Table 33 caption: "marginal" → "marginally"

p. 190, 4th-to-last line: "fewer errors" → "fewer phonological errors"

p. 198, line 7: "see 2.1.3" → "see 2.3"

p. 201, line 15: "**Text decoding and comprehension**" → "**Text decoding**"

p. 207, line 9: "stimulated by" → "through"

p. 209, line 3: "children who" → "children (17 mono + 13 bilinguals) who"

p. 209, line 12: "These children and all" → "These children including all"

p. 211, line 3 after Table 36: "see sections 2.1.2 – 2.1.14." → "see sections 2.2 – 2.14."

p. 213, line 11: "effects." → "effects of the mixed regressions."

p. 217, last line: "Figure 40" → "Figure 41"

p. 222, line 3 after Figure 42: "on the decoding speed" → "on decoding speed"

p. 224, line 2 after Figure 43: "Table 4.6 shows significant effects emerged in respective administrations." is erased

p. 224, line 3 after Figure 43: "(table 4.7)" is erased

p. 229, line 1 after Figure 46: "predictive longitudinal strength" → "longitudinal predictive strength"

p. 245, lines 13 - 14: "who developed reading or writing under the norm during the project." → "who were under the norm in reading or writing in 2nd and 3rd grade."

p. 251, line 9: "impaired short-term memory" → "impaired verbal short-term memory"

p. 252, line 18: "and in syntactic → and of syntactic"

p. 255, line 17: "imspirment" → "difficulty"

p. 262, 2nd-last line: "only picture RAN errors only" → "picture RAN errors only"

p. 267, line 2: "(see 4.3.2, Figure 14)" → "(see 4.4.2, Figure 26)"

p. 278, 3rd-to-last line: "tasks, and in" → "tasks. In"

p. 379, after line 19: the following text is added: "Many thanks to Prof. Ianthi Tsimpli for her suggestions on testing."

p. 380, line 1: "Moretti, who" → Moretti (Association Psicologia Insieme of Rome/Varese), who"