This is the final peer-reviewed author’s accepted manuscript (postprint) of the following publication:

**Published Version:**

**Availability:**
This version is available at: https://hdl.handle.net/11585/895345 since: 2024-05-20

**Published:**
DOI: http://doi.org/10.1080/13670050.2022.2132813

**Terms of use:**
Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher’s website.
Title

SPELLING AND WRITING SKILLS IN MINORITY-LANGUAGE BILINGUAL CHILDREN EXPOSED TO A TRANSPARENT ORTHOGRAPHY: MULTILEVEL PROFILES AND CONCURRENT PREDICTORS.

Authors

Alexandra Affranti¹, Valentina Tobia², Stephanie Bellocchi³, Paola Bonifacci¹

¹Department of Psychology, University of Bologna, Viale Berti Pichat n.5, 40127 Bologna (Italy)
²Faculty of Psychology, University Vita-Salute San Raffaele, Via Olgettina 58, 20132 Milan (Italy)
³Univ. Paul Valéry Montpellier 3, Univ. Montpellier, EPSYLON EA 4556, F34000, Montpellier, (France)

Corresponding author:

Paola Bonifacci
Department of Psychology,
University of Bologna,
Viale Berti Pichat n.5,
40127 Bologna (Italy)
E-mail: paola.bonifacci@unibo.it

Keywords: Bilingualism, Writing, Orthographic Skills, Written production, Morpho-syntactic comprehension, Phonology.

Declaration of interest: none
Spelling and writing skills in minority-language bilingual children exposed to a transparent orthography: multilevel profiles and concurrent predictors.

ABSTRACT

Many studies have explored how Language-Minority Bilingual Children (LMBC) read and comprehend, while the way they learn to spell and write has received less attention. This study aimed to assess a comprehensive profile of spelling and writing skills in LMBC, comparing performances of 4th and 5th grade bilingual (n = 74) and monolingual (n = 131) children in word and nonword reading and multilevel spelling and writing tasks (word, nonword and passage dictation, and written narrative production). Furthermore, we explored the role of linguistic and cognitive predictors (nonverbal Intellectual Quotient, verbal knowledge, morphosyntactic comprehension, nonword repetition) in spelling and writing outcomes. Our findings showed that overall LMBC did not reach a monolingual-like proficiency in spelling by the end of primary school, while they were similar to monolinguals in reading tasks and were able to produce written narratives with adequate macrostructure, syntactic complexity, and lexical variety. Moreover, morphosyntactic comprehension predicted spelling in both groups. Nonverbal intelligence and verbal knowledge predicted spelling skills only for the bilingual group. With regards writing skills, morphosyntactic comprehension emerged as a predictor exclusively in the bilingual group. These results are discussed with reference to educational and clinical implications.
1. Introduction

Writing can be considered as a multifaceted construct that includes different subcomponents such as spelling, and text production and that supports other academic skills (e.g., Graham & Hebert, 2011). For language-minority bilingual children (LMBC), who speak a home language that is different from the language in which they learn at school, spelling and writing tasks in L2 might be more challenging than their monolingual peers, due to lack of exposure and to the different time spans needed to develop specific academic skills. However, these gaps could be filled if given adequate time and educational opportunities. Cummins (1979) proposed a distinction between BICS (Basic Interpersonal Communication Skills), which are acquired earlier and mastered faster and CALP (Cognitive Academic Language Proficiency), which might require up to 5-7 years to reach monolingual-like levels of performance. Therefore, it is very important to understand developmental trajectories in LMBC in specific academic tasks to understand what might be considered an index of disorder or a peculiar characteristic of a bilingual trajectory. While groups' mean scores might be influenced by scores distribution, recent approaches focus on profile analysis, which allows to identify what percentage of a group could be at risk for developing spelling and writing disorders.

The present study aims to evaluate spelling and writing skills in LMBC and monolinguals in the final years of primary school, considering spelling (word, nonword, text) and written narrative production, using macrostructural and microstructural parameters. The investigation focuses on the groups' mean differences and the profile analyses, in order to understand if the groups differ in the probability of being poor spellers. Finally, concurrent higher-order cognitive skills (nonverbal IQ) and linguistic predictors of orthographic skills were analysed.

1.1. Spelling and writing development in bilingual children

Different theoretical models have examined the role of component skills in writing development. The Simple View of Writing (SVW, Berninger et al., 2002) includes two main skills: ideation, or the ability to create ideas through language, and transcription, the ability to put these
ideas into text form, including spelling and handwriting. This model is enriched in the Not-So-Simple View of Writing (Berninger & Winn, 2006) which includes domain-general resources (executive functions, motivation, self-efficacy), and self-regulatory composition processes, supporting planning, reviewing, revising. Further, working memory was hypothesized as a central component, with working memory demands during writing affecting each subcomponent. Finally, in Kim and Schatschneider (2017) language (vocabulary inference, grammar, theory of mind) and cognitive skills (working memory) had both Direct and Indirect Effects on writing (DIEW model), at discourse-level oral language, spelling, and handwriting fluency. Some studies investigated higher-order factors in spelling and writing, notably the role of general intelligence (nonverbal IQ), known to be highly related to working memory (Salthouse & Pink, 2008). Zarić et al., 2021 reported a significant contribution of general intelligence to spelling, although smaller than orthographic knowledge. Rindermann et al. (2011) found that a latent intelligence factor had the strongest direct effect on expressive writing in monolinguals aged 9-14. However, the contribution of general intelligence to these skills in bilingual children is less documented. A possible reciprocal interaction between higher-order cognitive functions and component skills of writing, and amongst writing components, should be noted. Spelling is reported to be the most stable longitudinal predictor of writing achievement from first to seventh grade (e.g., Abbott, et al., 2010). However, interaction paths might differ in monolingual and bilingual children. A recent meta-analysis (Graham & Eslami, 2020) evaluated how the SVW and, particularly, oral language, vocabulary, and transcription, relate to writing in L2 English learners. Transcription had the strongest correlation, followed by vocabulary and oral language. However, transcription also accounted for the highest percentage of variation at 31.22%, which is a possible variation in writing unexplained by the SVW. Considering spelling, Arab-Moghaddam and Sénéchal (2001) found that phonological and orthographic processing skills similarly predicted spelling in English, whereas orthographic skills alone predicted spelling in Persian (more transparent orthography) in English-Persian bilinguals in
2nd and 3rd grade. For Jongejan and colleagues (2007), phonological awareness was a significant predictor of written spelling in bilinguals from 1st to 4th grades with English as L2. Lexical access was the stronger predictor for older bilinguals (3rd – 4th graders); it is possible that L2 vocabulary size (Bialystok et al., 2010) may affect bilinguals’ writing outcomes from the 3rd grade, especially in lexical variety. The study also found higher associations between predictor variables and word spelling skills in older bilinguals (3rd and 4th grade), compared to younger students. Harrison and colleagues (2016), highlighted phonological awareness as a common predictor of word and texting spelling in bilinguals (L2 English) and monolinguals in 3rd grade, whereas syntactic awareness was a common predictor of writing quality. Savage et al. (2017) found, in their longitudinal study (1st to 6th grade), that ability to formulate grammatically correct sentences in L1 (English) is a significant predictor of microstructural indices like the number of words used in written production tests in L2 (French), suggesting that skills used in written production tasks are transferable across languages.

In summary, a wide range of linguistic predictors of spelling or writing have been proposed, with phonology, orthographic processing, and vocabulary as main predictors of spelling skills, and syntactic/grammatical knowledge and awareness as predictors of text writing. In line with this evidence, our study investigated linguistic predictors of spelling and writing skills, with a focus on lexical knowledge, phonological memory, and morpho-syntactic knowledge.

1.2. Bilinguals’ performance in word and nonword spelling

Zhao and colleagues’ meta-analysis (2016) demonstrated that bilinguals outperformed monolinguals in real-word spelling tasks, while monolinguals seem to perform better in nonword spelling tasks. The studies focused on L2-English, in which spelling in monolinguals is acquired later compared to transparent orthographies, limiting the generalizability of their results (Seymour, et al., 2003). Since monolingual English children commit more spelling errors than monolingual children acquiring a transparent language, the bilingual’s gap in writing skills might be more consistent for transparent vs. opaque languages. Studies on spelling skills have reported worse performances in L2 for bilingual children compared to monolingual peers when exposed to a
transparent orthography, albeit with limited evidence. Verhoeven (2000) reported that bilingual children exposed to Dutch in the first two grades of primary school underperformed compared to monolinguals in word spelling. Zhang et al. (2021) analysed spelling errors profiles of 3rd-4th graders with different linguistic backgrounds exposed to German, reporting Polish and Turkish students were more likely to be classified as poor spellers. A study involving LMBC 4th-5th graders with Italian as L2 (Authors, 2017) showed an underperformance in writing tasks compared to monolingual peers, even when controlling for socio-economic status (SES) (Authors, 2020, 2022). Interestingly, these differences did not persist when comparing results gathered in the English tasks, for which groups shared the same starting point and exposure. In summary, a partial explanation of conflicting results in bilinguals’ spelling performance may derive from language system characteristics (opaque vs. transparent); orthographic knowledge in LMBC may be limited due to less exposure to L2 orthography and/or a smaller L2 vocabulary, creating a larger gap between LMBC and monolingual peers in spelling compared to reading.

1.2. Bilinguals’ performance in text writing

Oral narrative skills have already been established as a valuable tool in exploring bilinguals’ linguistic skills, and share many commonalities with written narrative skills (Gagarina et al., 2016). Usually, oral narratives are analysed at the micro- and macrostructural level. The former refers to linguistic and quantitative features (word number, lexical width, and syntactic complexity), whereas the latter refers to the general structure of the story’s content (often referred to as Story Grammar) more specifically the occurrence of main events and characters and the Goal-Attempts-Outcomes (GAO) structure. Some studies found no differences among microstructural indexes between monolinguals and bilinguals (Cleave et al., 2010). Others note a better performance by monolinguals on some microstructural indexes like number of words, type and token and lexical errors (Authors, 2018). In macrostructure, bilinguals were comparable to monolingual peers in identifying key story elements and conveying them in a complex narrative (Boerma et al., 2016).
Some evidence seems to underline differences in various error types across genres and languages. Bahr et al. (2015) analysed writing samples of Spanish-English bilinguals attending middle-school, finding differences among three types of errors: phonological, orthographic and morphological. Phonological errors are related to incorrect one-to-one phoneme-grapheme correspondences, coupled with difficulty analysing the phoneme sequence in the word (e.g., *big* for *bik*). Orthographical errors involve the misapplication of language-specific orthographic knowledge, resulting in a phonologically plausible but incorrect attempt (e.g., *nite* for *night*). Morphological errors involve mistakes in inflectional or derivational morphemes or homonym errors (e.g., *they’re* for *their*). In Bahr’s study, participants wrote two expository and two narrative texts, in English and Spanish (eight writing samples from each participant). Authors found no differences in phonological errors, neither for language nor genre. English samples presented more morphological errors, while Spanish samples showed more orthographic errors, despite Spanish’s transparent orthography. English narratives showed a higher number of morphological errors than expository ones, which, in turn, showed a higher number of orthographical errors. These results highlighted how bilingual writers may show different levels of proficiency depending on language and writing task.

1.3. The connection between reading and spelling

Studies on monolinguals have shown significant links between reading and spelling, with early reading skills predicting later spelling skills in both opaque and transparent languages (Georgiou et al. 2020). Phonological awareness and Rapid Automatized Naming (RAN) are predictors of both reading and spelling skills, but vary with respect to tasks and orthographies. However, phonology might be more relevant for decoding and nonword spelling whereas lexical knowledge has a greater effect on word reading and spelling (Notarnicola et al., 2012). A meta-analysis found a moderate correlation between reading and spelling and RAN, with a stronger association with word spelling over nonword spelling, and opaque orthographies over transparent orthographies (Chen et al.,
Previous research showed LMBC might reach adequate decoding skills within their first two years of schooling (see August and Shanahan, 2006). However, they may still underperform, when assessed only in L2, in words with increased orthographic complexity (Droop & Verhoeven, 2003) or minor frequencies (Authors, 2016). Finally, spelling may be affected by differences in the amount of reading in L2 (Georgiou et al. 2020).

In summary, reading skills seem to consolidate earlier in bilinguals compared to spelling and writing skills, although they might be affected negatively by a limited L2 vocabulary, leading bilinguals to rely more on phonology than semantic knowledge in decoding.

Present study

The main aim of this study was to explore, in the final years of primary school, spelling and writing skills and their linguistic and cognitive concurrent predictors in bilingual language-minority children learning a transparent language (Italian) as L2, and in their monolingual peers. The focus was threefold:

1) To analyse differences in writing performance, considering different typologies of spelling errors in words, nonwords, and text dictation tasks and micro- and macrostructural aspects of written production. LMBC were expected to underperform in word spelling tasks due to their limited consolidation of L2 orthographic knowledge, and in line with previous studies on transparent languages. On the contrary, they may master the phoneme-grapheme correspondence more easily, thus reaching higher scores in nonword writing. In text production, LMBC may obtain lower scores regarding microstructural indexes, as previously found in bilinguals’ oral narratives (Authors, 2018), given their narrower vocabulary in L2 (Bialystok et al., 2010). We also hypothesized that the macrostructural level may be a strength for bilinguals, as the abilities to identify the story grammar and to build a mental scheme of the story are not language-specific.

2) To assess the patterns of weaknesses and strengths in a profile analysis that considers, besides mean values, the percentage children that fall in a range of deficient, borderline, or typical
performance. Such analysis would allow understanding whether the mean differences in performances actually reflect the possibility that bilingual children may be considered poor spellers. As writing is an active skill that may require more time to be acquired, we hypothesized that a higher percentage of LMBC may not reach an adequate level of proficiency in spelling and writing tasks, compared to the reading ones.

3) To assess linguistic and cognitive concurrent predictors of orthographic skills and narrative productions. Measures of phonological memory (nonword repetition), lexical knowledge (verbal knowledge), and morphosyntactic knowledge were included in the regression model. Nonword repetition was chosen as a phonological task because previous evidence highlighted that it requires less lexical knowledge and is reliable in discriminating bilingual children with language/learning impairments (e.g., Boerma et al., 2015). We expected vocabulary and phonology to be significant predictors of spelling skills and, according to Harrison et al., (2016), morphosyntactic knowledge to predict mainly text writing. Concerning higher-order cognitive skills, we expected nonverbal IQ would predict spelling and writing skills in bilinguals, in line with previous studies on monolinguals (Zarić et al., 2021; Rindermann et al. 2011).

Information regarding participants’ reading levels was included in the study, seeing that reading is known to be a powerful predictor of spelling skills and because reading has received greater attention in literature on LMBC, with more consistent results compared to writing skills. However, since the population examined was at the end of primary school, reading was not included as a predictor since, at least in typical developing monolinguals exposed to a transparent orthography, decoding skills reach a ceiling effect and might not elucidate the developmental process linking reading to spelling. Therefore, analysis of reading performance is mainly aimed at assessing consistency with previous studies on reading, and, to increase the generalizability of the results in terms of writing skills.

2. Methods
2.1. Participants

The initial sample involved 240 children that were recruited from ten classes of mainstream primary schools in a suburb area in the north of Italy; they attended Grade 4 or 5 and aged between 9 and 11 years old. Exclusionary criteria were having sensory or neurological impairments, a diagnosis of neurodevelopmental disorders and having less than two years of continuous exposure to Italian in the scholastic setting. Furthermore, we included only bilingual children with both parents speaking a minority language at home. Information was collected through a questionnaire administered by teachers to parents. The sample involved in the study totalled 205 children (45.9% males, mean age = 9.5, SD = .5); of these, one hundred thirty-one were monolinguals (48.1% males, mean age = 9.4, SD = .5), while seventy-four were bilinguals (41.9% males, mean age = 9.5, SD = .5). The bilingual group comprised mostly early bilinguals (73.3%), meaning they were consistently exposed to Italian (entering in the preschool system) before the age of three, and the bilingual group’s mean Age of First Exposure to Italian was 2.4 years (SD = 2.2 years). Their mean socio-economic status (SES), measured according to Hollingshead (1975) was 21.3 (SD = 8.7), which corresponds to a medium-low range. Children spoke a vast heterogeneity of L1 languages: Arabic (31.1%), Romanian (13.5 %), Albanian (13.5%), Tagalog (8.1%), Chinese (6.8%), Spanish (6.8%), Niger-kordofanian Languages (Twi, Ibo, 6.8%), Pidgin-English (5.4%) and other languages (Polish, Syrian, Turkish, Indian, Urdu, 8.1%). Most parents of the monolingual group, who were not directly interviewed by teachers, did not return SES questionnaires. However, L1 and L2 children were recruited children from the same classrooms, attended by children from the same neighbourhoods and this should minimise possible differences in socio-economic status.

2.2. Materials

2.2.1. Verbal Knowledge and Non-verbal intellectual quotient

Kaufman Brief Intelligence Test-2 (K-BIT2; Kaufman & Kaufman, 2004, Authors, 2016). In the verbal knowledge subtest, a word or a sentence is read aloud by the experimenter and the child is asked to choose one picture, among a group of six, that best depicts it. In the matrix subtest, the
child is asked to choose one picture, among a group of six, that best fits in the matrix presented. A score of 1 was given for every correct response and standard scores were calculated from the norms in the test manual. The split-half reliability coefficient was 0.86 for the nonverbal scale and .90 for the verbal scale.

2.2.2. *Morphosyntactic comprehension task*

BVN 5-11 (Bisiacchi et al., 2005), simplified version of the TROG test (Bishop, 2003): the task comprises a list of 18 sentences: for each one, the child is asked to choose one picture, among a group of four, that best depicts the meaning of the sentence. A score of 1 was given for every correct response; analyses were run on raw scores since the norms for our sample’s age range are not available.

2.2.3. *Nonword Repetition task*

BVN 5-11 (Bisiacchi et al., 2005), nonword repetition: Each nonword was read aloud by the experimenter: a score of 1 was given if the child was able to repeat the item after one presentation, 0.5 if he/she was able to repeat the item after two repetitions. Z-scores were derived from norms given by the test manual. The test-retest reliability index reported in the text manual was .23

2.2.4. *Word and nonword reading task*

ALCE Battery (Authors, 2014): The word reading task comprises three lists of 20 words increasing in length and decreasing in frequency, while the nonword reading task comprises two lists of 15 nonwords. For both tasks, the reading speed was measured in syllables per second, and the accuracy was measured by the percentage of items that the child read incorrectly. The analyses were performed on the T-scores (mean = 50; SD = 10) provided by the test manual. The KR-20 reliability index reported in the text manual was .89 for word reading and .96 for nonwords.

2.2.5. *Word and nonword writing task*

Tasks 6 (words) and 7 (nonwords) of the BDDE-2 (Sartori et al., 2007): Forty-eight words and twenty-four nonwords were read aloud by the experimenter. Raw scores consisted in the number of incorrect items; each word spelled incorrectly was coded as 1 error (irrespective of the number of
2.2.6. *Passage dictation task*

BVSCO (Tressoldi et al., 2012): The word count of the text for children in 4th grade was 135 words, the one for 5th graders was of 148 words. A score of 1 was given to every misspelt word and it was assigned to one of the error categories considered:

- *Phonological errors* – regarding phoneme-grapheme conversion, they included inversion, omission, addition or substitution of one or more graphemes.
- *Non-phonological errors* – the phoneme-grapheme conversion is followed, but the child did not apply orthographic rules correctly (i.e., illegal fusion or separation of a word, apostrophes use and incorrect *h* in the verb “avere” [to have], which in the third person requires the *h*, “lui ha” [he has], whereas the preposition “a” [to] is spelled without the *h*);
- *Accents/doubles errors (phonetic errors)* – regarding an omission or addition of accents or double letters (in Italian words may have double letters, as for example in *letto* [bed]).

Z-scores were calculated from the norms in the test manual.

2.2.7. *Written text production*

BVSCO, Tressoldi et al., 2012: Participants were asked to write a story based on a pictorial stimulus made of 5 vignettes as they would tell it to a friend.

The macrostructure was coded through the identification of the main character’s goals, actions and outcomes, according to the MAIN method (Multilingual Assessment Instrument for Narratives; Gagarina, et al., 2012). We assigned different levels of complexity as follows:

- no complexity (0) – there are references only to an attempt (A) or an outcome (O);
- low complexity (1) – there are references to an attempt and an outcome (AO);
- medium complexity (2) – there are references to both a goal and an attempt or outcome (GA -GO);
- high complexity (3) – there are references to a goal, an attempt and an outcome (GAO).

We also calculated the total number of story elements included (meaning the total number of goals, actions and results). The maximum score was seven.

The analysis of microstructural indexes included: Mean Length of Utterances (the number of words divided by the number of phrases), number of words, number of phrases, number of type (the number of different words used), type/word ratio, and syntactic complexity (number of principal, coordinate and subordinate clauses and principal/complex clauses ratio). We also counted the number of errors regarding the same categories used in the dictation, and, in addition, we considered morphological errors (e.g., when the noun and its article have a different number or gender; i.e., the article is plural while the noun is singular, or the article is masculine while the noun is feminine). Z-scores were calculated from the manual for the number of words, phrases and percentages of phonological, non-phonological and accents and double letter errors.

For the remaining variables, z-scores were obtained on the total (monolinguals and bilinguals) sample’s means and standard deviations.

2.3. Procedure

Parents signed the informed consent for participating in the study. The matrix subtest and reading tasks were administered individually in a quiet room at the children’s schools; writing tasks were administered collectively in the classroom in two different sessions. The study obtained the approval of the Ethics Committee of the University of [Blind].

3. Results

3.1. Analysis of group differences

In order to investigate the first aim of the study, t-tests were run on T-scores or z-scores for the following variables: matrix, verbal knowledge, passage dictation errors, and total story elements. For word and nonword reading, word and nonword spelling, text spelling errors, narrative
structure, and errors in the written text production, separate MANOVAs were run, with Group (monolinguals vs. bilinguals) as between-subjects factor. For variables with a skewness not included in the range ± 2 (Curran, et al., 1996), a square root transformation was performed. Descriptive statistics with mean scores and SDs for each task are presented in Tables 1 and 2.

As expected, the results on the matrix task showed no difference between the groups (t (203) = - .87, p = .38, Cohen’s d = -.13). For word and nonword reading tasks, a multivariate significant effect of Group was found, F (4,199) = 5.668, p = .00, η2 = .102, with a significantly better performance observed in monolinguals for both word reading speed, F (1,203) = 14.401, p = .00, η2 = .067, and accuracy, F (1,203) = 4.879, p = .03, η2 = .024. On the contrary, no differences emerged regarding the nonword reading task, F (1,203) = 1.829-2.879, p = .09-.18, η2 = .009-.014.

Considering spelling skills, the MANOVA run on the word and nonword spelling tasks showed a significant multivariate effect of Group, F (2,201) = 11.929, p = .00, η2 = .106, with LMBC making more errors in both word, F (1,203) = 23.890, p = .00, η2 = .106 and nonword, F (1,203) = 6.484, p = .01, η2 = .031, dictations. Also, for passage dictation, more errors were observed for LMBC, t (101,179) = -5.47, p = .00, Cohen’s d = .82. The types of errors in the passage dictation task were analysed with a MANOVA that showed a significant multivariate effect of Group, F (3,194) = 12.723, p = .00, η2 = .164, with LMBC making more phonological, F (1,197) = 22.291, p = .00, η2 = .102, non-phonological, F (1,197) = 25.460, p = .00, η2 = .115, and phonetic, F (1,197) = 24.019, p = .00, η2 = .109, errors.

Turning to the written text production, no difference emerged regarding the macrostructural index that measured the total number of elements of the story (t (202) = -.01, p = .99, Cohen’s d = .00). LMBC made a higher percentage of total errors (t (96.382) = 3.56, p = .01, Cohen’s d = .58).
More specifically, the MANOVA run on the types of errors showed a significant multivariate effect of Group, $F(4, 199) = 7.834, p = .00, \eta^2 = .136$, and LMBC made more phonological, $F(1, 203) = 4.281, p = .04, \eta^2 = .021$, non-phonological, $F(1, 203) = 9.292, p = .00, \eta^2 = .044$, accents and doubles, $F(1, 203) = 18.231, p = .00, \eta^2 = .083$, and morphosyntactic, $F(1, 203) = 12.198, p = .00, \eta^2 = .057$, errors.

Regarding microstructural indexes, the results were not homogeneous. The multivariate effect of Group proved significant, $F(9, 194) = 4.338, p = .00, \eta^2 = .168$, but although LMBC wrote less words, $F(1, 203) = 9.532, p = .00, \eta^2 = .045$, and less phrases, $F(1, 203) = 8.337, p = .00, \eta^2 = .040$, and employed less different words, $F(1, 203) = 18.133, p = .00, \eta^2 = .082$, their type/word ratio, $F(1, 203) = 2.218, p = .14, \eta^2 = .011$, and Mean Length of Utterances, $F(1, 203) = .239, p = .63, \eta^2 = .001$, were nevertheless comparable to the monolinguals’ productions. Lastly, the results showed that the LMBC’s narratives contain less principal clauses, $F(1, 203) = 6.193, p = .01, \eta^2 = .029$, and coordinate clauses, $F(1, 203) = 4.031, p = .05, \eta^2 = .020$, but the same number of subordinate clauses, $F(1, 203) = 1.050, p = .31, \eta^2 = .005$, and a similar principal clauses/complex clauses ratio, $F(1, 203) = .014, p = .91, \eta^2 = .000$, compared to their monolingual peers.

3.2 Profile analysis

To explore the second aim of the study, for each task we divided the sample into three groups, based on their performance’s $z$-scores: typical (> -1 SD), borderline (> -1; < -2 SD) and deficient (< - 2 SD). For macrostructural complexity, the text productions were coded on a four-level scale of complexity (absence, low, medium, high). We ran Chi-Square tests on the percentages of LMBC and monolinguals in each performance range.

Figure 1 shows the distributions across the three ranges of proficiency for different types of errors in the passage dictation task.

Insert Figure 1 here
In the word reading task, there was a significant difference in the distribution of performance ranges for the speed parameter ($\chi^2(2) = 14.03, p = .01$). More bilinguals fall in the borderline range than the deficient range. Results showed no differences regarding the nonword reading task, both for speed ($\chi^2(2) = 5.89, p = .05$) and accuracy ($\chi^2(2) = 3.67, p = .16$), nor in the word reading task for accuracy ($\chi^2(2) = 5.55, p = .06$). In spelling tasks, results showed no differences in nonword spelling ($\chi^2(2) = 4.31, p = .12$), but a higher percentage of the LMBC group struggled in word spelling ($\chi^2(2) = 24.70, p = .00$). LMBC also displayed more severe difficulties in the writing task than in the corresponding reading task. Overall, the passage dictation task was more challenging for LMBC than monolinguals, as they underperformed in proficiency considering all error categories: total ($\chi^2(2) = 21.36, p = .00$), phonological ($\chi^2(2) = 15.81, p = .00$), non-phonological ($\chi^2(2) = 15.97, p = .00$) and accents and doubles ($\chi^2(2) = 19.03, p = .00$). Overall, in the passage dictation task, distributions showed around 25% of LMBC made enough errors to fall in the deficient range, with phonological errors being the most common.

Considering the written text production, Figure 2 shows the distributions for total errors and the three categories of errors considered in the narrative task; in the absence of standardized values in the manual for this type of errors these were not included in profile analysis.

Insert Figure 2 here

In this case, non-phonological ($\chi^2(2) = 9.43, p = .01$), accents and doubles ($\chi^2(2) = 16.74, p = .00$) and total errors ($\chi^2(2) = 10.94, p = .01$) showed differences in the distribution for monolinguals and bilinguals, while no difference was found for phonological errors ($\chi^2(2) = .80, p = .67$).

Regarding macrostructure, bilinguals’ performance was comparable to monolingual peers ($\chi^2(3) = 2.64, p = .45$), as the two groups showed a similar distribution across the four levels.
3.3. Cognitive and Linguistic predictors of writing skills

A multigroup structural equation model (Kline, 2010), including confirmatory factor analysis (CFA) and path analysis, was run with the latent factors Spelling and Written text production as dependent variables and Matrix, Verbal knowledge, Morphosyntactic comprehension and Nonword repetition as potential concurrent predictors. The Spelling factor was obtained via a CFA run on the number of errors in the writing tasks of words, nonwords, and text; for words and text errors, scores were transformed with a square root transformation because their skewness was higher than 2 (Curran et al., 1996). The CFA for creating the Written text production factor was run on the following variables: number of Tokens, number of Types, number of sentences, MLU, and macrostructural complexity. The model (Figure 3), performed with the MPlus software (Muthén & Muthén, 1998–2010), was tested on monolinguals and bilinguals with the multigroup technique. Multiple indices were used to evaluate model fit: the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the standardized root mean squared residual (SRMR). TLI and CFI values equal to or higher than .90 indicate an acceptable model fit; RMSEA and SRMR values close to, respectively, .10 and .08 indicate a reasonable fit (e.g., Hu & Bentler, 1999). The results are displayed in Figures 4 a,b.

Insert Figure 3 here

Insert Figure 4 a,b here

Relevant differences were observed between monolinguals and bilinguals in the pattern of predictors. For monolinguals, spelling (explained variance: 12.4 %) was significantly predicted by morphosyntactic comprehension, whereas the variables considered were not significantly related to monolinguals’ written text production (explained variance: 3.5 %). For LMBC, spelling was

---

1 The matrix with all correlations between variables included in the study has been added as supplementary material (Appendix 1)
predicted, as for monolinguals, by morphosyntactic comprehension, but also by non-verbal reasoning; verbal knowledge showed a predicting role that was marginally significant. The explained variance was 35.6%. Then, written text production was significantly predicted by morphosyntactic comprehension, with an explained variance of 8.5%. The model presented acceptable fit indices: RMSEA = .10 (90% confidence interval = .08-.12); CFI = .92; TLI = .90; SRMR = .08.

4. Discussion

The main purpose of this study was to explore spelling and writing skills in LMBC learning a transparent language as L2, compared to monolingual peers. There were three specific goals: analyse group differences, deepen the understanding of group differences through profile analysis, and investigate linguistic and higher-order cognitive predictors of spelling and writing skills. Because of the strong connection documented between reading and spelling, the study included reading measures.

The analysis of group differences evidenced that dictation spelling tasks were a challenge for bilinguals, as differences emerged for all tasks administered, with incremental effect sizes from nonwords to words and text spelling. This trend is consistent with previous findings on spelling in LMBC exposed to an L2 transparent language (Verhoeven, 2000; Zhang et al., 2021; Authors, 2017, 2020), but contrasting with previous studies on English as L2 (Zhao et al., 2016). Considering reading skills, in line with previous studies (Authors, 2016, 2017), the present sample of LMBC showed adequate nonword decoding, not significantly different from monolinguals. For word reading, LMBC underperformed in both speed and accuracy. Considering written text production, this study replicates and extends previous findings on oral narrative skills (Boerma et al., 2016; Authors, 2018), with adequate LMBC performances at the macrostructural level and weaknesses, although heterogeneous, in microstructure. In particular, when bilingual children wrote stories in L2, they were as capable as their monolingual peers to point out main events, and did not differ
from monolinguals in MLU, type/token ratio, and principal clauses/complex clauses ratio. However, they underperformed in phonological, orthographic, phonetic, and morphosyntactic errors and wrote fewer words and sentences, but a similar number of subordinates.

Profile analysis allowed us to understand fine-grained patterns of LMBC’s performance. Nonword spelling and reading had similar percentages of poor spellers and readers in both groups. Conversely, a higher percentage of bilinguals underperformed in the word spelling task and word reading speed, but not accuracy. Considering text spelling, 60% of LMBC children had a fully adequate performance, against 88% of monolinguals. Finally, in writing production, profile analyses and group differences had a similar pattern of results.

In summary, although bilinguals globally underperformed monolinguals, profile analyses showed they more frequently underperformed only in word spelling, word reading speed, text spelling, and orthographic accuracy in writing production, but not in nonword reading and nonword spelling. This suggests that phoneme-grapheme correspondence in writing might require more effort than grapheme-phoneme matching in reading.

This pattern of results confirmed previous findings suggesting that spelling skills, at the orthographic level, might require more time and additional skills to develop compared to reading skills (Authors, 2017, 2020). Furthermore, LMBC weaknesses in orthographic competence involved all error types and therefore seem to be generalized to the whole process, not specific to the phonologic vs. orthographic domain.

Concerning microstructure, LMBC’s text productions were shorter with a limited lexical variability. However, no differences in type/word ratio emerged, suggesting that LMBC were able to produce adequate written productions in terms of lexical richness although “in few words” (Authors, 2018). We found similar results regarding syntax: LMBC used fewer coordinate and principal clauses, but adequate syntactic complexity (principal/complex clauses ratio). These results highlighted the importance of evaluating microstructural indices to avoid penalizing bilinguals’ written productions, weighing scores on the actual length of production.
Finally, regarding spelling and writing skills predictors, our findings are in line with Harrison and colleagues (2016), as morphosyntactic skills help predict both spelling and written text production skills for the LMBC group, but only spelling skills in monolinguals and the absence of such a relationship deserves more investigation in transparent languages. Further, Harrison et al.’s study involved a syntactic awareness task, whereas this study employed a syntactic knowledge task. Nonverbal IQ was another significant predictor of spelling skills (Zarić et al., 2021) associated with a marginal effect of vocabulary, in line with Jongejan and colleagues (2007). Globally, the explained variance for spelling was higher than in writing, particularly for bilinguals. This suggests that writing skills, due to their complexity, may involve other higher-order factors not part of this study, like planning, translating or revising.

In conclusion, LMBC exposed to a transparent orthography underperformed in word, nonword, and text spelling tasks when assessed only in L2 in the final years of primary school, considering both phonological and non-phonological errors. Their level of performance was similar to monolinguals for word reading accuracy and nonword reading and spelling, but about 25% of the sample for the text dictation task and 30% for the written production task fell in the deficient range. In written text production, LMBC showed adequate macrostructural parameters and weaknesses in microstructural parameters. Finally, linguistic skills were stronger predictors for LMBC. This pattern of results underlines the multicomponent nature of spelling and writing skills and shows a peculiar profile of strengths and weaknesses.

Some considerations can be proposed to interpret the gap in spelling skills. The first concerns the developmental trajectory of L2 acquisition. In fact, the bilingual gap in written tasks may be related to the BICS/CALP distinction (Cummins, 1979). Writing can be considered a CALP component and involves more than phonological-based correspondence, possibly taking longer to develop, even if most children were born in the host country and have an acceptable level of spoken language skills.
In addition, participants were exposed to a highly transparent language which may be responsible for the gap with monolinguals, who develop writing skills faster when exposed to transparent languages compared to opaque languages (Treiman & Kessler, 2005). Furthermore, spelling in Italian seems to be strongly determined by lexical skills (Notarnicola et al., 2012), which may widen the gap, since bilinguals might rely more on phonological skills due to a limited L2 vocabulary (Bialystok et al., 2010).

Alternatively, the macrostructural aspects of writing production could be a relative strength for LMBC children, pointing to the utility of evaluating written narrative macrostructure in linguistic assessment of bilingual children. In clinical contexts, macrostructural indices might help distinguish between Developmental Language Disorder (DLD) and insufficient L2 exposure, (Boerma et al., 2016). Regarding spelling disorder identification, nonword spelling might be considered a reliable tool, since it is expected to reach adequate levels of performance in typically developing LMBC children before word and text spelling; this is in line with the clinical accuracy of nonword repetition tasks in DLD identification (Boerma et al., 2015, Authors, 2020).

Some limitations should be considered when interpreting the results of the study. First, LMBC and monolinguals were not matched for SES. We acknowledge that this lack makes it difficult to discern if differences arise from lack of exposure to the dominant language or poor language input due to low-SES background, limiting the generalizability of the findings. Our results are however in line with previous studies (Authors, 2022) in which low-SES LMBC underperformed in spelling tasks compared to low-and high-SES monolinguals. Secondly, our measure of phonology referred to phonological memory (nonword repetition) rather than phonological awareness. Although it may limit the generalizability of the results, this task was chosen because it is a reliable index of DLD in bilinguals, being freer of lexical knowledge (Boerma et al., 2015, Authors, 2020). In addition, the heterogeneity of linguistic backgrounds does not allow to control for specific L1-L2 connections. Finally, the present study excluded some
variables considered important in previous literature (handwriting, working memory, and reading) in the prediction of writing skills. These aspects in bilingual populations require future study.

Despite these limitations, this study offers a new picture of the pattern of spelling and writing skills among LMBC; it is, to our knowledge, the first study considering different levels of analysis, typologies of errors, and methodologies. Moreover, data collected on LMBC exposed to highly transparent languages may enrich previous literature focused on English as L2.

Finally, there are implications in clinical and educational settings. At a diagnostic level, clinicians should consider that using monolingual standards might strongly increase the chance of misdiagnosis for middle low-SES LMBC. Secondly, different levels of spelling and writing skills should be considered. Evaluating writing production might better aid in identifying strengths than dictation/spelling tasks. In educational contexts, there is a need for educational interventions that aid the development of linguistic and cognitive skills, as is shown by the connection drawn between these skills and spelling and writing in bilinguals. Bilingual children are required to perform academic tasks in a language they are still consolidating and, moreover, have linguistic competence in another language. Therefore, any gaps with respect to monolingual standards should be interpreted from an evolutionary and multidimensional perspective.

REFERENCES


Authors (2018).

Authors (2017).

Authors (2020a).

Authors (2020b).

Authors (2022).

Authors (2016).
Authors (2014).


Hollingshead, A. A. (1975). *Four-factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.


Table 1. Mean scores and standard deviations for matrix, reading and dictation tasks for the LMBC and monolingual groups.

<table>
<thead>
<tr>
<th></th>
<th>Monolinguals</th>
<th>LMBC</th>
<th>Direction of differences</th>
<th>Skewness (SE = 0,170)</th>
<th>Kurtosis (SE = 0,338)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Matrix°</strong></td>
<td>72–137</td>
<td>98,64</td>
<td>13,3</td>
<td>72–127</td>
<td>98,64</td>
</tr>
<tr>
<td><strong>Verbal Knowledge</strong></td>
<td>12–53</td>
<td>39,87</td>
<td>7,77</td>
<td>8–50</td>
<td>36,53</td>
</tr>
<tr>
<td>**Morphosyntactic Comprehension **</td>
<td>11–18</td>
<td>17,05</td>
<td>1,14</td>
<td>9–18</td>
<td>16,28</td>
</tr>
<tr>
<td><strong>§Non-Word Repetition^</strong></td>
<td>-2,00–1,00</td>
<td>0,54</td>
<td>0,6</td>
<td>-2,25–1,00</td>
<td>0,27</td>
</tr>
<tr>
<td><strong>Word Reading Speed</strong></td>
<td>32–68</td>
<td>50,78</td>
<td>8,68</td>
<td>23–68</td>
<td>45,59</td>
</tr>
<tr>
<td><strong>Word Reading Accuracy</strong></td>
<td>16–56</td>
<td>49,25</td>
<td>9,14</td>
<td>22–56</td>
<td>46,08</td>
</tr>
<tr>
<td><strong>Non-Word Reading Speed</strong></td>
<td>29–80</td>
<td>50,26</td>
<td>9,58</td>
<td>28–73</td>
<td>48,31</td>
</tr>
<tr>
<td><strong>Non-Word Reading Accuracy</strong></td>
<td>26–63</td>
<td>51,11</td>
<td>8,12</td>
<td>20–63</td>
<td>48,96</td>
</tr>
<tr>
<td><strong>Word Dictation^</strong></td>
<td>-7,00–1</td>
<td>-0,09</td>
<td>1,01</td>
<td>-11,50–1,00</td>
<td>-1,4</td>
</tr>
<tr>
<td><strong>Non-Word Dictation^</strong></td>
<td>-3,50–1,50</td>
<td>-0,41</td>
<td>1,05</td>
<td>-6,50–1,50</td>
<td>-0,88</td>
</tr>
<tr>
<td><strong>Passage Dictation - Phonological errors^</strong></td>
<td>-6,99–0,88</td>
<td>-0,3</td>
<td>1,45</td>
<td>-24,57–0,88</td>
<td>-2,61</td>
</tr>
<tr>
<td><strong>Passage Dictation - Non-Phonological errors^</strong></td>
<td>-9,52–0,59</td>
<td>-0,31</td>
<td>1,5</td>
<td>-9,35–0,59</td>
<td>-1,46</td>
</tr>
<tr>
<td><strong>Passage Dictation - Accents and Doubles errors^</strong></td>
<td>-3,23–0,66</td>
<td>0,36</td>
<td>0,67</td>
<td>-5,44–0,66</td>
<td>-0,36</td>
</tr>
<tr>
<td><strong>Passage Dictation - Total errors^</strong></td>
<td>-7,00–1,02</td>
<td>0,02</td>
<td>1,29</td>
<td>-14,16–1,02</td>
<td>-1,89</td>
</tr>
</tbody>
</table>

°Standard scores, *T-scores, ^z-scores, **raw scores

In bold, variables that showed significant differences between groups (§ not included in group analysis).
Table 2. Mean scores and standard deviations for writing narrative task for the LMBC and monolingual groups.

<table>
<thead>
<tr>
<th></th>
<th>Monolinguals</th>
<th>LMBC</th>
<th>Direction of differences</th>
<th>Skewness (SE = 0,170)</th>
<th>Kurtosis (SE = 0,338)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
<td>Mean</td>
</tr>
<tr>
<td>Story complexity*</td>
<td>-1.83–0.99</td>
<td>0</td>
<td>1.02</td>
<td>-1.83–0.99</td>
<td>0</td>
</tr>
<tr>
<td>Total story elements*</td>
<td>-3.53–1.56</td>
<td>0</td>
<td>1.05</td>
<td>-3.53–1.56</td>
<td>0</td>
</tr>
<tr>
<td>Number of phrases^</td>
<td>-4.68–2.18</td>
<td>-0.68</td>
<td>1.31</td>
<td>-5.57–2.18</td>
<td>-0.12</td>
</tr>
<tr>
<td>Number of words^</td>
<td>-1.70–3.20</td>
<td>0.16</td>
<td>1.01</td>
<td>-1.80–2.58</td>
<td>-0.28</td>
</tr>
<tr>
<td>Number of types*</td>
<td>-1.77–2.98</td>
<td>0.22</td>
<td>0.98</td>
<td>-1.95–2.04</td>
<td>-0.38</td>
</tr>
<tr>
<td>Word/Type ratio*</td>
<td>-3.72–2.56</td>
<td>0.08</td>
<td>0.92</td>
<td>-3.92–2.27</td>
<td>-0.14</td>
</tr>
<tr>
<td>MLU*</td>
<td>-1.77–3.45</td>
<td>-0.02</td>
<td>0.92</td>
<td>-1.79–3.61</td>
<td>0.05</td>
</tr>
<tr>
<td>Principal Clauses*</td>
<td>-1.51–4.33</td>
<td>0.13</td>
<td>1.02</td>
<td>-1.51–4.33</td>
<td>-0.22</td>
</tr>
<tr>
<td>Coordinate Clauses*</td>
<td>-1.51–3.80</td>
<td>0.1</td>
<td>0.96</td>
<td>-1.88–3.80</td>
<td>-0.19</td>
</tr>
<tr>
<td>Subordinate Clauses*</td>
<td>-1.66–3.33</td>
<td>0.06</td>
<td>1.03</td>
<td>-1.66–2.95</td>
<td>-0.09</td>
</tr>
<tr>
<td>Principal/subordinate</td>
<td>-1.08–5.77</td>
<td>0.01</td>
<td>0.94</td>
<td>-1.02–6.79</td>
<td>-0.03</td>
</tr>
<tr>
<td>ratio*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Phonological errors^</td>
<td>-9.75–0.63</td>
<td>-0.53</td>
<td>1.75</td>
<td>-15.50–0.63</td>
<td>-1.21</td>
</tr>
<tr>
<td>% Non-Phonological errors^</td>
<td>-10,18–0.58</td>
<td>-0.42</td>
<td>1.66</td>
<td>-10,99–0.58</td>
<td>-1.28</td>
</tr>
<tr>
<td>% Accents and Doubles errors^</td>
<td>-4.21–0.80</td>
<td>-0.13</td>
<td>1.18</td>
<td>-8.38–0.80</td>
<td>-1.19</td>
</tr>
<tr>
<td>% Morphosyntactic errors*</td>
<td>-3.43–0.75</td>
<td>0.18</td>
<td>0.85</td>
<td>-4.06–0.75</td>
<td>-0.24</td>
</tr>
<tr>
<td>% Total errors^</td>
<td>-7.08–1.07</td>
<td>-0.57</td>
<td>1.63</td>
<td>-10.62–1.07</td>
<td>-1.92</td>
</tr>
</tbody>
</table>

* z-scores obtained on total sample’s scores
^z-scores obtained from manual
In bold, variables that showed significant differences between groups.