



An analysis of the impact of the L1 on bilinguals' writing competence

Paula Elosua^{*}, David Lasagabaster

University of the Basque Country UPV/EHU, Spain

ARTICLE INFO

Keywords:

Assessing writing
Bilinguals
L1
Alignment

ABSTRACT

The research fields of second language acquisition in general and L2 writing in particular have been dominated by studies in which English is the language under scrutiny. However, many different researchers have claimed that it is high time to consider languages other than English (LOTes), since multilingualism is the norm rather than the exception all over the world. In particular, minority languages have been largely overlooked in the literature on L2 writing, a research gap that this study intends to fill in. With this objective in mind, this research examines the impact of the participants' L1 (Basque or Spanish) on their written production. The sample was made up of 9767 participants from the Basque Autonomous community in Spain, who completed a writing test aimed at assessing the C1 level (Common European Framework of Reference for Languages) in Basque in five different points in time. The results revealed differences in writing competence in favour of L1=Basque participants when compared with their L1=Spanish counterparts, especially as far as the criterion of accuracy was concerned. The study concludes by providing insights that will be useful in other international educational contexts in which a minority and a majority language share co-official status.

1. Introduction

The rapid spread of both theoretical and empirical work on L2 writing has been outstanding during the last few decades. However, in the field of second language acquisition in general and in L2 writing in particular, research has clearly focused on English as a foreign language (EFL). It can be affirmed that the hegemonic position of English as the current lingua franca has profoundly shaped not only research but also current conceptualizations on L2 writing. Authors (e.g. Ushioda & Dörnyei, 2017) from different fields of research have underscored the need to also consider languages other than English (LOTes) in an era in which multilingualism has become a buzzword. In an attempt to contribute to redressing this imbalance between EFL and LOTes, in this paper we intend to focus on a minority language, Basque, whose speakers straddle the Pyrenees, as it is spoken both in the north of Spain and the south of France. Although the status of minority languages in Spain has been overshadowed by Spanish as the majority language, at this moment in time their status is higher than ever in history (Elosua & De Boeck, 2020; Lasagabaster, 2017). The revitalization policies undertaken in those Spanish autonomous communities where they are co-official with Spanish have greatly boosted both the formers' symbolic and economic value (O'Rourke & Ramallo, 2015) and in this process L2 speakers, who are referred to as new speakers in the literature (O'Rourke et al. 2015), are of the utmost importance. Since the survival of a minority language hinges to a great extent on new speakers' ability to use the language in many different domains and whether they succeed in making it part of their everyday lives, the

^{*} Corresponding author at: University of the Basque Country, Avda. Tolosa, 70, San Sebastian 20018, Spain.
E-mail address: Paula.elosua@ehu.es (P. Elosua).

<https://doi.org/10.1016/j.jslw.2025.101182>

Received 18 September 2024; Received in revised form 21 January 2025; Accepted 21 January 2025

Available online 31 January 2025

1060-3743/© 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

analysis of their written production gains momentum due to the fact that it is considered to be “a basic educational skill” (Ushioda & Dörnyei, 2017; 451) but also one of the most challenging linguistic skills for L2 learners to master.

With a LOTE perspective in mind, the objective of this article is to compare the written production of adults for whom Basque is their L1 (also known as traditional speakers) with adults for whom it represents their L2. Since the survival of minority language hinges to a great extent on new speakers’ ability to use the language in many different domains and whether they succeed in making it part of their everyday lives, the analysis of their written production gains momentum due to the fact that it is considered to be “a basic educational skill” (Ushioda and Dörnyei, 2017; 451) but also one of the most challenging linguistic skills for L2 learners to master. From an international perspective, the bilingual situation in the Basque Autonomous community (BAC) in Spain will help to inform research on L2 writing. Despite the potential inherent differences with other international bilingual contexts in which a majority and a minority language coexist, the results found in the Basque context may help other officially bilingual settings to glean useful insights concerning the importance of the impact of the L1 on writing proficiency. The Basque experience may open new research agendas in the L2 writing field of research at a time when the presence of minority languages is steadily increasing in education systems the world over (O’Rourke et al. 2015). Furthermore, it has to be noted that the label new speakers is closely linked to that of heritage speakers (O’Rourke et al. 2015), who can be found in any country in the age of globalization.

2. The Basque context

Minority languages are always defined in terms of power issues, as the use of term implies that there is also a majority language in the context under scrutiny (Pedley & Viaut, 2019). Basque, along with Spanish, holds co-official status in the BAC, encompassing the provinces of Bizkaia, Gipuzkoa, and Araba (in Basque). Whereas Spanish is the majority language and it is spoken all over Spain, Basque is a minority language spoken in only two of the 17 autonomous communities that make up Spain, namely the BAC and Navarre. In this context, the normalization process of the minority language plays a key role:

The process of normalizing Basque is a relatively recent development (Hualde & Zuazo, 2007). In 1918, the Basque Language Academy was established with the goal of standardizing the language, even though the standardized version was not commonly spoken at the time. Significant progress in constructing a standardized Basque took place after the Spanish Civil War (1936–1939). The proposal for *Euskara Batua* (Unified or Standardized Basque) was presented and adopted by the Academy in 1968. The Basque Academy continues its work in codifying standard Basque by establishing rules and creating lists of words in the standard form, among many other tasks (see Hualde & Zuazo, 2007).

However, when Spain came under Franco’s control (1939–1975), the use of the Basque language was forbidden in every single social sphere (including school) and those who dared to violate this ban were persecuted. This had a very detrimental impact on the number of Basque speakers whose numbers fell dramatically. With the advent of democracy, Basque was granted co-official status in 1979, and efforts to revitalize the language gained momentum in the 1980s, supported by the 1982 Act Governing Standardization of the Use of Euskera (Basque). In the BAC, all residents have the right to use Basque alongside Spanish. The use of Euskara has expanded to various domains, including education, administration, and the media. The last sociolinguistic survey carried out in 2021 (Basque Government, 2023) reveals that there are currently 680,629 bilingual speakers (36.2 % of the Basque population). However, a much higher percentage is found in the 16–24 age range (74.5 %) due to the widespread use of Basque as medium of instruction in education.

At this stage a distinction needs to be made between new and traditional speakers of Basque. Traditional or L1 speakers of Basque are those who acquired the language at home and who tend to live in small towns and villages. While they speak a dialectal variety at home and use it for everyday social interactions, they learn standard Basque at school and usually limit its use to more formal situations such as official exams. Interestingly, 44.3 % of Basque speakers express themselves more easily in Spanish, whereas the percentage of those who feel more at home in Basque is much lower, that is, 27.4 % (Basque Government, 2023). The remaining 28.3 % claim to have a high proficiency in both languages.

On the other hand, new speakers are defined as those individuals who learn the language not through family transmission but rather through some form of formal education, albeit not exclusively (O’Rourke et al. 2015). Basque new speakers usually have little exposure to the language in their everyday lives outside the formal contexts where they learn it. In the Basque context new speakers tend to be gathered in largely Spanish-speaking urban areas of big cities. New or L2 speakers of Basque usually exhibit a good command of the standardized variety, although it may sometimes be perceived as artificial and distinct from the vernacular spoken by native speakers (Ortega et al. 2015). It is important to note that new Basque speakers play an essential role in the survival of Basque as they nowadays outnumber traditional speakers, which is why the future of the language largely rests on their shoulders. According to Urla et al. (2018), new Basque speakers’ language development is well worth looking into due to two main reasons: “First, because they constitute the majority of younger Basque speakers today and second, because their formal schooling has been overwhelmingly in the newly created standard Basque” (p. 26). Although there are studies on new speakers’ language attitudes, habits of language use or educational experiences, to our knowledge no previous study has focused on the impact of the L1 (by comparing new and traditional speakers) on official writing tests in Basque.

3. Individual differences: the impact of the L1 on writing

Writing is a very complex task that varies considerably from person to person not only in the L2 but also in the L1 (Schoonen et al., 2003) because it involves the exploitation of both linguistic and (meta)cognitive skills and linguistic resources and it is therefore regarded as multicomponential (Schoonen et al., 2011). In fact, reviews of the literature (e.g. Dabrowska, 2012; Kidd et al. 2018) have confirmed significant individual differences among adult monolingual native speakers of the same language. For example, more

educated speakers have more experience with formal written language, which obviously impacts on their written production (Cumming, 1989).

According to different authors (Kormos, 2012; Ahmadian et al. 2022), the importance of individual differences in second language (L2) speaking has been widely looked into, whereas the impact of learner differences on L2 writing has been overlooked. In the field of L1 writing research particular attention has been paid to particular individual differences such as working motivation, self-efficacy beliefs and working memory capacity (Kormos, 2012). This is also the case in the field of L2 writing, in which factors such as language aptitude, working memory, motivational factors, learner beliefs, and emotions have also been examined (Ahmadian et al. 2022). Whereas cognitive factors such as aptitude and working memory have contributed to our understanding of L2 writing, other non-cognitive factors such as the L1 (Crossley & McNamara, 2009; Ströbel et al. 2020) or the use of the individual's multilingual repertoires (Beiler, 2021) should also be considered, because otherwise we will not be able to "paint a comprehensive picture of L2 writing" (Ahmadian et al. 2022: 383). Hulstijn (2024) points out that BLC (Basic Language Cognition) Theory attempts to explain the role of individual differences in the acquisition and control of both native and non-native languages. Among the main constructs of BLC Theory is the Basic Language Cognition (BLC), which is acquired through massive exposure to input and is common to all adult native speakers but uncommon to the vast majority of non-native speakers and which therefore demands further investigation. To our knowledge, however, the role of students' L1 has been neglected in L2 writing studies in LOTE, which is why little is known about how the L1 may hinge on written production in a minority language. The few studies available tend to focus on foreign languages such as French (e.g. Gabriel et al. 2021).

This demographic variable, namely students' L1, is of the utmost importance in contexts where a minority language is spoken, as new speakers play a vital role in the revitalisation process of the language, as mentioned above. Our study is thus relevant because it aims at comparing L1 and L2 written performance in a high-stakes exam in Basque (see below). Studies have hitherto analysed how previous experience in L1 writing affects the L2 writing process by focusing on how the L1 is used for content generation (Weijen et al. 2009), to generate lexical units, self-questioning and evaluations (Murphy & Roca de Larios, 2010), to transfer metacognitive knowledge –operationalized as knowledge of writing strategies and text characteristics– from L1 to L2 (Schoonen et al. 2011), or how cross linguistic influence is evidenced in L2 written texts (Schoonen & van Vuuren, 2022). Researchers have also observed that L1, L2 and L3 writing performances are highly interdependent, which indicates that there is transfer of writing skills (Angelis & Jessner, 2012; Tullock & Fernández-Villanueva, 2013). In this vein, in a longitudinal case study based on one participant's writing development in three languages, Kobayashi and Rinnert (2013) observed that "boundaries became blurred among both the textual and the linguistic features in the three languages. At the same time, there were multi-directional interactions between languages, affected by such interrelated factors as proficiency level, the amount of writing knowledge acquired, and perceptions of writing in each language." However, to our knowledge the examination of how new or L2 speakers perform when compared to traditional or L1 speakers in a minority language has been largely disregarded.

It is also worth noting that the vast majority of studies focused on written performance have English as the target language and many of them tend to focus on very specific aspects of the written production. Crossley and McNamara (2009) compared the writing samples of L1 and L2 English writers based on lexical features. The L2 writers were native speakers of Spanish whose command of English was high intermediate to advanced and their writings were compared with those produced by North-American undergraduates. These authors relied on a freely available on-line computational tool named Coh-Metrix that "was able to successfully distinguish to a significant degree L1 and L2 texts based solely on lexical features" (p. 132). Lan et al. (2022) carried out a corpus-based investigation on noun-phrase complexity among L1 and L2 English writers. Whereas L1 writings showed diverse patterns of noun phrases, L2 essays had compressed structures of noun phrases, whose use is regarded as an indicator of advanced academic writing. The authors concluded that noun phrase complexity can be differentiated in L1 and L2 writing, results which concur with those of Crossley and McNamara (2009). However, these studies centered on very specific aspects of L2 writing, while we intend to analyse written production from a more holistic perspective by examining five more broad writing domains. Seminal studies such as Cumming (1989) and Sasaki and Hirose (1996), as well as epoch-making studies such as Schoonen et al. (2003), point out that if students are to develop more expert writing performance, they should be prompted to focus beyond single aspects of their writing (e.g. lexical or syntactic errors), because otherwise their attention would be directed away from considerations of the complexity of their writing output.

Since writing abilities are influenced by L2 exposure (Kim et al. 2022), it is important to underscore that the exposure to Basque as L2 is much higher than that of the EFL participants of the studies analysed above. This is a factor that may help to diminish differences when comparing L1 and L2 written texts in Basque. In addition, previous studies relied on small samples, a methodological weakness that we intend to overcome by relying on a much larger number of participants.

In the United States Perin et al. (2017) analysed 16 studies on writing skills covering all grade levels from elementary to post-secondary that compared language minority students, defined as those students who speak a language different from English at home, and observed a clear trend toward differences between L1 and L2 English speakers. The results revealed that the participants who had been educated in English dominant schools for at least 3 years resembled native speakers of English in some ways, although they displayed poorer writing skills than their L1=English peers. The authors conclude that the field is "pending a more nuanced and systematic body of knowledge of LM writing than is currently available" (p. 13). However, a main difference between the studies reviewed by Perin et al. (2017) and ours is that in the former L2 participants were minority language speakers, while in our study many of the participants are majority language speakers (Spanish L1 speakers for whom Basque is the L2).

An interesting study to report on is the one by Phuoc and Barrot (2022), as they did consider whether writing proficiency was a matter of L1 background. These authors explored complexity, accuracy and fluency (CAF measures) as an index of writing proficiency. CAF measures have been thoroughly examined in research on writing, but the focus on L1 backgrounds remains hitherto an untrodden

path (Phuoc & Barrot, 2022). When the argumentative essays of Asian writers with different L1s were compared, the findings evidenced that those L2 participants whose languages were typologically close (e.g. Chinese and Thai L2 writers) exhibited remarkable similarities in CAF measures that those whose L1s were dissimilar (e.g. Chinese and Pakistani L2 writers). Many other studies have examined CAF measures in L2 writing, but only a few have considered different L1 backgrounds. Kuiken and Vedder (2019) analysed complexity measures in argumentative essays at the A2 and B1 levels of the Common European Framework of Reference for Languages (CEFR) written by two groups of participants: Dutch speakers who wrote in Italian (N = 39) and Spanish (N = 23), and 32 Dutch L2 learners who had 27 different L1 backgrounds. The authors found significant correlations between proficiency and complexity measures but acknowledged that, due to the small sample size (and the many diverse L1s among Dutch L2 learners), no definitive conclusion could be drawn. Khushik and Huhta (2020) zoomed in on syntactic complexity at the A1, A2 and B1 of the CEFR in argumentative essays produced by Finnish and Pakistani (whose L1 was Shindi) teenager writers. The results revealed differences between the two groups when similar proficiency levels were compared, as Pakistani learners produced longer sentences and relied more on coordination, while the Finnish participants used more subordination.

In any case, it is important to note that the three studies reviewed in the previous paragraph focused only on complexity. In summary, our study is innovative due to four main reasons: (i) it analyses broad writing domains rather than specific aspects (e.g. complexity measures), (ii) the L2 under scrutiny is not English but rather a minority language, (iii) the L2 speakers are L1 =majority language speakers; most studies involving minority language speakers analyse these students' writing performance in the dominant language, and (iv) the sample is very large; e.g. the majority of the studies reviewed by Perin et al. (2017) had a small number of participants, although the studies by Phuoc and Barrot (2022) and Khushik and Huhta (2020) would be an exception on that score.

4. Research questions

Our approach is twofold, incorporating both the analysis of the internal structure of the writing competence and the study of observed raw scores. The paper addresses the following research questions:

- (1) RQ1: Is the internal structure, i.e., the underlying components, of writing competence equivalent among bilinguals with different first languages (L1s)?
- (2) RQ2: Does bilingual speakers' L1 influence their overall writing competence?
- (3) RQ3: Does bilingual speakers' L1 have an impact on each of the evaluation criterion defined to measure writing competence?

5. Method

5.1. Participants and instrument

The essay. Writing competence is assessed through an essay as part of a test designed to evaluate the C1 level in Basque, in accordance with the CEFR (European Council, 2021). Participants are given 90 min to complete the task, with a minimum word count of 300 words. The essay typically requires candidates to write an opinion article or reflective piece on a given topic, framed within a situational context. Prompts are designed to encourage critical thinking and structured argumentation, often addressing broad societal issues such as technological advancements, environmental challenges, or social changes. Candidates are expected to present a well-organized response, offering a general perspective on the topic and elaborating on specific aspects related to the theme. This is an independent writing task, allowing participants to generate and express their ideas without relying on external sources or integrated reading materials. The testing program usually runs two assessment sessions per year. In this study, five test editions were examined: the two held in 2020 (2020_1, 2020_2), the two from 2021 (2021_1, 2021_2), and the 2022 spring edition (2022_1).

The scoring. The essays were scored analytically using five domains or assessment criteria: (a) adequacy, (b) coherence, (c) cohesion, (d) richness and (d) accuracy (see Table 1). Each domain received a score based on a 5-point scale ranging from 0 (inadequate), 1 (minimal), 2 (sufficient) 3 (good), to 4 (very good). The scores for each domain are combined to calculate the final score with a maximum possible score of 20. This scoring framework was developed to ensure a comprehensive evaluation of both the functional and linguistic aspects of writing competence. While the approach shares similarities with the Functional Adequacy Scale proposed by Kuiken and Vedder (2017), such as the emphasis on coherence and task-related adequacy, it incorporates additional criteria, like richness and accuracy, to capture the linguistic range and grammatical correctness of the essays.

The examinees. The sample comprised the test-takers from the BAC. The number of participants was 9767; the L1 was Basque for

Table 1

Definitions of the five analytical domains.

<i>Adequacy</i>	Whether the subject is adequately approached. Explanations and arguments, not mere examples and anecdotes.
<i>Coherence</i>	Explain the ideas in a clear, concise and appropriately organised manner. The structure of the text is appropriate.
<i>Cohesion</i>	Related ideas and sentences well and appropriately explained; Use of elements of cohesion to give unity to the text.
<i>Richness</i>	General lexicon correctly and appropriately used. Ideas expressed using appropriate complex structures.
<i>Accuracy</i>	Correct spelling and grammar.

59.9 % ($N_{L1\text{Basque}} = 5852$) and Spanish for 40.1 % ($N_{L1\text{Spanish}} = 3917$). The mean age was 22.2 ($SD = 8.80$). The characteristics of the candidates can be seen in Table 3.

The raters. 350 raters took part in the scoring process for each of the editions. The number of essays scored per rater varied from 36 to 40, with 38 being the median. Ten raters took the role of adjudicators, experts with at least 3 years' experience, tasked with evaluating essays that received discrepant scores.

The rating procedure. Two raters independently scored each composition analytically, using five domains or assessment criteria. If the two raters agreed on the pass/fail decision (a score of 10 or higher was considered a passing score), the scoring process was complete. However, if the two raters assigned different pass/fail decisions, the essay was treated as a discrepant paper and was reassessed by an adjudicator. In cases of disagreement, the adjudicator's decision resolved the pass/fail outcome. Once agreement was reached, the candidate's final score was determined as the higher of the two raters' scores, ensuring that candidates received the most favorable evaluation in line with the scoring criteria.

5.2. General data analytic approach

The data analytic approach has three features. First, we analyzed data coming from five consecutive test editions instead of relying on just one. This avoids capitalizing on chance that may be associated with the results of a single edition analysis and achieves a level of generalization that would not be possible with the study of a single call. Second, we will use information regarding writing competence scores as well as evaluation criteria scores. Third, the study of the internal structure allows us to investigate measurement invariance, so that possible attenuating effects stemming from observed scores are dis-attenuated (Millsap, 2011). This aspect of the research is preliminary, as measurement equivalence serves as the foundation for conducting group comparisons. Without measurement invariance, there is a risk of drawing inaccurate conclusions, where observed mean differences may not be attributable to underlying mean differences but rather to irrelevant distinctions between groups, such as variations in test editions, for instance.

Measurement invariance. Alignment. An invariant measure is one in which the distribution of observed scores, conditioned on the underlying space, is independent of the group variable (Meredith, 1993). When invariance exists among groups, individuals belonging to different groups with the same level on the latent variable(s) being measured will obtain the same average observed score. Traditionally, the assessment of measurement invariance is performed using Multiple-Group Confirmatory Factor Analysis (MG-CFA). However, when dealing with a large number of groups to compare, the alignment method is preferred; this procedure operates under the assumption that there may be some degree of non-invariance, and the goal is to minimize it. In this study, groups are formed by combining test editions (five) and candidates' L1 (Basque and Spanish), resulting in a total of 10 groups. To conduct the analysis (Asparouhov & Muthén, 2014; Muthén & Asparouhov, 2018), we followed a two-step process. Initially, a configural model is established to identify the best-fitting model among all groups and where factor loadings and thresholds were set free across groups. In the next step, the alignment optimization is carried out where factor means and variances are freely estimated. Once the point of minimum discrepancy has been reached, the analysis compares factor means and factor variances across groups and identifies the model parameter (loadings and thresholds) that significantly differ from the average of that parameter across all groups. Parameters that are not significantly different are grouped together as invariant, while those that are significantly different are identified as non-invariant. Finally, a Monte Carlo simulation was designed to assess the quality of the alignment output. The number of replication was fixed to 100 and the sample size to the mean values of samples across test editions ($N = 1950$). For each of those conditions the correlation between the true factor means and the estimated factor means were analysed. Correlations values above .98 indicate a good quality of the alignment optimization.

Considering the ordinal nature of the data, the invariance was evaluated for items loadings and thresholds (Elosua, 2011; Millsap & Yun-Tein, 2004). Model fit for the configural model was assessed by the Chi-square statistic, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI) and the Standardized Root Mean Residual (SRMR). SRMR should be less than or equal to .08 for a good model fit, a value equal or higher than .95 is accepted as indicative of good fit for CFI and TLI indices (Hu & Bentler, 1999). The modelling process was conducted in Mplus version 8.9 (Muthén & Muthén, 2023).

Factorial design. The comparative study of observed scores includes six blocks of analysis: one for the overall writing competence scores and one for each of the five assessment criteria. After testing the assumptions of the ANOVA, it was consistently found that homogeneity of variances was violated ($p < 0.01$). Therefore, a non-parametric version of ANOVA was performed: the analysis of variance of aligned transformed data (ART anova; Wobbrock et al. 2011). The impact of test edition and evaluation criteria on writing competence were assessed. In addition to pair-wise comparisons, Cohen's effect sizes were estimated.

Data preparation: Prior to conducting the invariance analysis, the number of candidates in each of the score categories was checked. Due to the limited count of examinees falling into the extreme categories, with only 88 candidates (0.90 %) receiving a score of 0 and 326 candidates (3.34 %) achieving a score of 4, these extreme categories were merged with their adjacent categories. Specifically, scores of 0 and 1 were combined into a single category, as were scores of 3 and 4, while the middle category (score of 2) remained unchanged. Following this adjustment, the initial 5-point scale was compressed into a 3-point scale to ensure a more balanced distribution across categories for the analysis.

6. Results

RQ1. Internal structure across groups

The configural model, which estimated loadings and thresholds freely across groups yielded the following fit indices: χ^2 (86) = 784.79; $p < .001$; CFI = .984; TLI = .981; SRMR = .037). Overall, these values suggest a good fit between the model and the data,

indicating that the unidimensional model presented in the Fig. 1 is tenable for the 10 groups.

The optimization process was carried out using as reference sample the group whose native language is Spanish and who took the test in the autumn of the year 2020. Table 2 displays the estimated parameters and the count of both invariant and non-invariant groups, and Table 3 presents the latent means for group and the ranking according those values. It is worth noting that all loadings were invariant, so each assessment criteria contributes to the latent construct, writing competence, to a similar degree across groups. Regarding thresholds, two of the five assessment criteria were invariant, namely adequacy and coherence. This implies that individuals with the same level of writing competence receive comparable scores in those two domains. The remaining assessment criteria exhibited some form of non-invariance: (a) for the assessment criterion cohesion, the first threshold displayed non-invariance in one of the groups (2021, spring, Basque L1), (b) the first threshold for richness was non-invariant in two groups (2020, first call, Spanish L1; 2022, first call, Spanish L1), and (c) for the accuracy domain non-invariant values were identified in the first threshold for five groups, and in the second threshold in two groups (2021, first call, 2022 first call). Notably, all of non-invariant terms were associated with the groups of candidates whose native language is Basque. In summary, out of the total of 50 estimated loadings, 100 % of them were invariant, and out of the 100 estimated thresholds, 10 exhibited non-invariance, with seven of them being associated with the accuracy criterion. The literature suggests a non-invariance rate of 25 % as the cut-point for conducting mean comparisons (Asparouhov & Muthén, 2014), and in this research, the percentage of non-invariance parameters is 6.66 %, which is well below the acceptable value.

The alignment optimization solution was confirmed by Monte Carlo simulation. The correlation between the true factor means and a single simulation factor means was found to be .986, indicating a strong relationship. Similarly, for factor variances, the correlation value was .96. Across the 100 simulations, the mean correlation values for these two parameters were .99, demonstrating a consistent and high level of agreement.

Summary of the results. The alignment analysis concluded that internal structure of the writing competence remained invariant across test editions and candidates' L1. The latent mean comparisons provide a ranking among groups, showing that individuals whose native language is Basque performed better than those whose native language is Spanish. It is important to highlight that most of the threshold parameters related to the accuracy domain show non-invariance among groups defined by their L1, indicating a distinction based on L1 that will be explored further in the following sections.

RQ2: Global results by L1 and test edition

The analysis revealed significant interaction between L1 and test edition ($F(4,9754) = 3.84; p = .003$) but an associated effect size close to 0 ($\eta^2 = 0.001$). These results allowed us to interpret the main effects ($F_{L1}(1,9754) = 200.53; F_{\text{testedition}}(4,9754) = 14.38$), whose η^2 values were .02 for mother language and .005 for test edition. The examination of marginal means uncovered a mild to moderate effect size ($d = .31$) for the differences attributed to L1. Among the pairwise comparisons of test editions, only three out of the ten comparisons yielded Cohen's d values exceeding .10, although none of them reached the value of .20; the pairs were: $d_{2020_1/2021_1} = -.19; d_{2020_1/2022_1} = -.16; d_{2020_2/2021_1} = -.16$ - Fig. 2 shows the interaction plot between test edition and L1, and Table 4 shows the marginal means. Candidates with Basque as their native language consistently demonstrated higher means, which is concordant with the previous latent analysis. Although the size of the differences showed slight fluctuations across editions, the pattern consistently favoured Basque candidates.

Note. In parenthesis SE. CI = confidence interval; LL = lower limit; UL = upper limit.

RQ3: Assessing criteria results by L1 and test edition.

Adequacy. There is not an interaction effect between L1 and test edition ($F(4,9754) = 2.34; p = .05$), and the main effects are statistically significant ($F(1, 9754) = 167.00; F(4, 9754) = 17.70$) with η^2 values of .016 for L1 and .005 for test edition. The comparison among candidates based on L1 showed a Cohen's d effect size of .27; Among the 10 pairwise comparisons of test editions, all comparisons, except one ($d_{2020_1/2020_2} = .27$), exhibited negligible effects.

Coherence. The interaction and main effects were found to be statistically significant ($F(4, 9754) = 4.66; F(1, 9754) = 104.98; F(4, 9754) = 49.14$), with corresponding η^2 values of .001 for the interaction term, .01 for mother language, and .02 for test edition. The difference in mean values for L1 was associated with a Cohen's d effect size of .22. Among the test editions, significant differences were observed in six cases with effect sizes being negligible for three of them ($d < .20$), and small for the rest of three ($d_{2020_1/2021_1} = .33; d_{2020_2/2021_1} = .35, d_{2021_1/2021_2} = .33$).

Cohesion. The interaction and main effects were found to be statistically significant ($F(4, 9754) = 18.55; F(1, 9754) = 24.36; F(4, 9754) = 21.15$), with corresponding η^2 values of .007 for the interaction term, .002 for mother language, and .02 for test edition. The difference in mean values for L1 was associated with a Cohen's d effect size of .10. Among the test editions, significant differences were observed in six cases; two of them (2021_1 vs. 2021_2 and 2021_2) showed small/medium effect size ($d_{2021_1/2021_2} = .40; d_{2021_2/2022_1} = .46$), the rest of Cohen's d values were smaller than .2.

Richness. The interaction and main effects were found to be statistically significant ($F(4, 9754) = 15.13; F(1, 9754) = 165.51; F(4,$

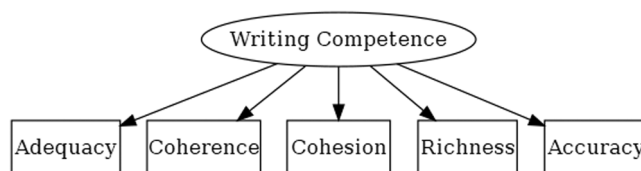


Fig. 1. Internal structure of the writing competence.

Table 2
Parameter estimates.

Assessment criteria	Parameter	Weighted average	Number (percentage) of approx. invariant groups
Adequacy	Loading	2.44	10
	Threshold 1	-4.29	10
	Threshold 2	2.50	10
Coherence	Loading	2,56	10
	Threshold 1	-2,17	10
	Threshold 2	3.26	10
Cohesion	Loading	1.96	10
	Threshold 1	-1,73	9 (90 %)
	Threshold 2	3.86	10
Richness	Loading	0.87	10
	Threshold 1	-1.20	8 (80 %)
	Threshold 2	3.93	10
Accuracy	Loading	0.86	10
	Threshold 1	1.00	5 (50 %)
	Threshold 2	4.50	8 (80 %)

Table 3
Latent mean values.

Test edition	N	L1	N	Latent mean	σ^2	Ranking
2020_1	2285	Basque	1417	0.320 (.06)	2.30	4
		Spanish	868	0.000	2.21	2
2020_2	1269	Basque	724	0.22(.07)	1.19	5
		Spanish	545	0.054 (.07)	0.826	9
2021_1	2752	Basque	1714	0.467(.06)	1.20	2
		Spanish	1038	0.23 (.06)	0.99	6
2021_2	1373	Basque	751	0.320 (.07)	1.03	3
		Spanish	622	0.172 (.07)	0.89	7
2022_1	2088	Basque	1246	0.564 (.07)	1.40	1
		Spanish	842	0.165 (.06)	0.68	8

Note. In parenthesis standard error (SE)

9754) = 50.87), with corresponding η^2 values of .006 for the interaction term, .01 for mother language, and .02 for test edition. The difference in mean values for L1 was associated with a Cohen's *d* effect size of .27. In the case of the test editions, significant differences were observed in four cases, with effect size values nearing .3 in each of them ($d_{2020_1/2021_1} = .36$; $d_{2020_2/2021_1} = .36$; $d_{2021_1/2021_2} = .27$; $d_{2021_1/2022_1} = .25$).

Accuracy. The interaction and main effects were found to be statistically significant ($F(4, 9754) = 5.03$; $F(1, 9754) = 413.59$; $F(4, 9754) = 37.98$), with corresponding η^2 values of .002 for the interaction term, .04 for mother language, and 0.1 for test edition. The difference in mean values for L1 was associated with a Cohen's *d* effect size of .44. As for test editions, significant differences were observed in five comparisons; the biggest Cohen's *d* values were associated with the comparisons 2021_1/2021_2 ($d = .34$), 2020_1/2021_2 ($d = .26$), and 2021_1/2022_1 ($d = .26$).

Summary of results. The non-parametric ANOVAs show significant interactions between L1 and test edition for coherence, cohesion, richness and accuracy, but the associated η^2 values were close to 0. Since the ordinal interactions got a minimal practical significance, we conducted pair-wise comparisons of the main effects to identify any potential patterns. The main results indicated that across all evaluation criteria, candidates whose native language was Basque consistently outperformed candidates whose native language was Spanish (Table 5). This contrast was particularly notable in the accuracy criterion, where the effect size was more substantial ($d = .48$). For the remaining criteria, the respective values were as follows: Adequacy ($d = .19$), Coherence ($d = .18$), Cohesion ($d = .05$), and Richness ($d = .23$). Fig. 3 displays the mean values for each of the evaluation criterion and candidate's L1. Regarding the factor associated with the test edition, while certain specific effects were identified, no discernible or constant patterns were uncovered in the data.

Note. In parenthesis SE. CI = confidence interval; LL = lower limit; UL = upper limit.

7. Discussion and conclusions

To our knowledge, this is the first study to systematically examine the impact of L1 on the writing production of a large sample of minority language learners. Perin et al. (2017: 14) highlight the lack of systematic and rigorous research in the field of writing among minority language speakers, which limits the development of effective interventions. They stress the importance of capturing "specific constructs consistently," a principle that guided our methodology. We achieved this by analyzing data from different cohorts across five test editions, paying close attention to the internal structure of the writing competence test, and utilizing a large sample size to identify group membership as a potential moderator.

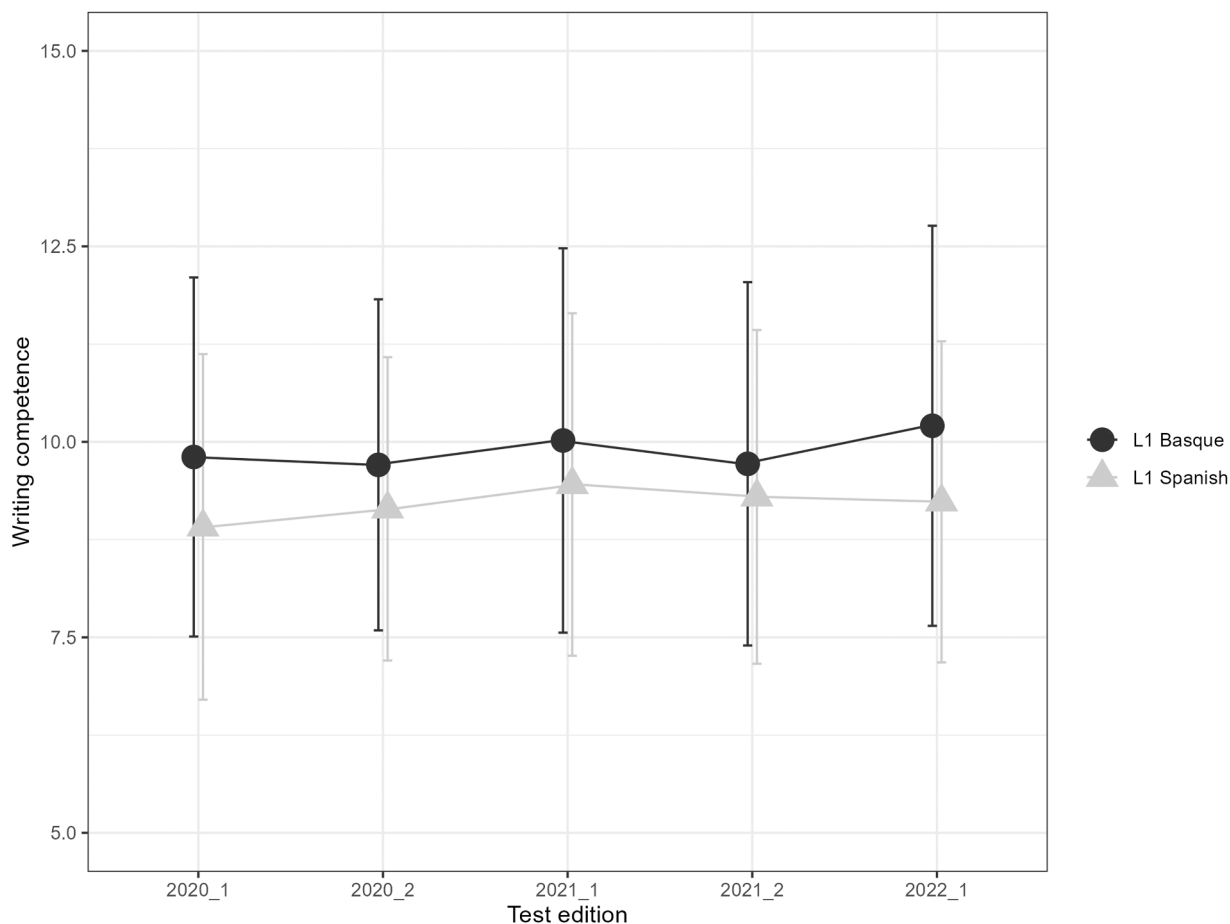


Fig. 2. Mean plots by L1 and test edition.

Table 4
Sample composition and statistics by L1 and test edition.

Test edition	Marginal Means	95 % CI [LL, UL]	SD	L1	Marginal Means	95 % CI [LL, UL]	SD
2020_1	9.38 (.04)	[9.29 – 9.48]	2.30	Basque	9.74 (.05)	[9.64 – 9.84]	2.30
				Spanish	9.03 (.05)	[8.92 – 9.14]	2.21
2020_2	9.41 (.06)	[9.29 – 9.54]	2.06	Basque	9.77 (.06)	[9.64 – 9.90]	2.12
				Spanish	9.06 (.06)	[8.92 – 9.20]	1.94
2021_1	9.72 (.04)	[9.63 – 9.80]	2.38	Basque	10.07 (.04)	[9.98 – 10.06]	2.46
				Spanish	9.37 (.05)	[9.26 – 9.47]	2.19
2021_2	9.49 (.06)	[9.37 – 9.62]	2.25	Basque	9.85 (.06)	[9.72 – 9.98]	2.32
				Spanish	9.14 (.06)	[9.01 – 9.27]	2.13
2022_1	9.74 (.05)	[9.65 – 9.84]	2.41	Basque	10.10 (.05)	[9.99 – 10.20]	2.56
				Spanish	9.39 (.05)	[9.28 – 9.50]	2.05

The methodology employed in this study can be adapted to other minority language contexts by replicating its core design: integrating data from multiple test editions and forming comparison groups based on linguistic backgrounds. This approach minimizes the risk of capitalizing on chance inherent in single-test studies, yielding more reliable and generalizable insights. Additionally, it ensures meaningful comparisons by accounting for the complexities of linguistic diversity.

A key strength of this methodology lies in its emphasis on measurement invariance, ensuring that observed differences in writing performance are attributable to true underlying differences rather than test-specific or group-related factors. The use of alignment optimization further enhances the approach by accommodating some degree of non-invariance while identifying areas requiring closer investigation.

Following the research questions raised in this study, the results demonstrate the writing competence exam holds the condition of measurement invariance across test editions and candidates' L1. The alignment output provided valuable insights, highlighting two

Table 5
Statistics by assessment criteria and L1.

Criterion	Mother language	Marginal Mean	95 % CI [LL, UL]	SD	Cohen's <i>d</i>
Adequacy	Basque	2.28 (.007)	[2.27, 2.30]	0.56	.19
	Spanish	2.18 (.008)	[2.16, 2.20]	0.51	
Coherence	Basque	2.08 (.008)	[2.06, 2.10]	0.66	.18
	Spanish	1.97 (.010)	[1.95, 1.99]	0.60	
Cohesion	Basque	1.92 (.007)	[1.91, 1.94]	0.60	.05
	Spanish	1.90 (.009)	[1.88, 1.92]	0.54	
Richness	Basque	1.89 (.008)	[1.88, 1.91]	0.61	.23
	Spanish	1.76 (.009)	[1.74, 1.78]	0.60	
Accuracy	Basque	1.71 (.008)	[1.70, 1.73]	0.66	.48
	Spanish	1.41 (.010)	[1.39, 1.43]	0.62	
Writing competence	Basque	9.89 (.030)	[9.83, 9.95]	2.39	.31
	Spanish	9.21 (.037)	[9.13, 9.28]	2.13	

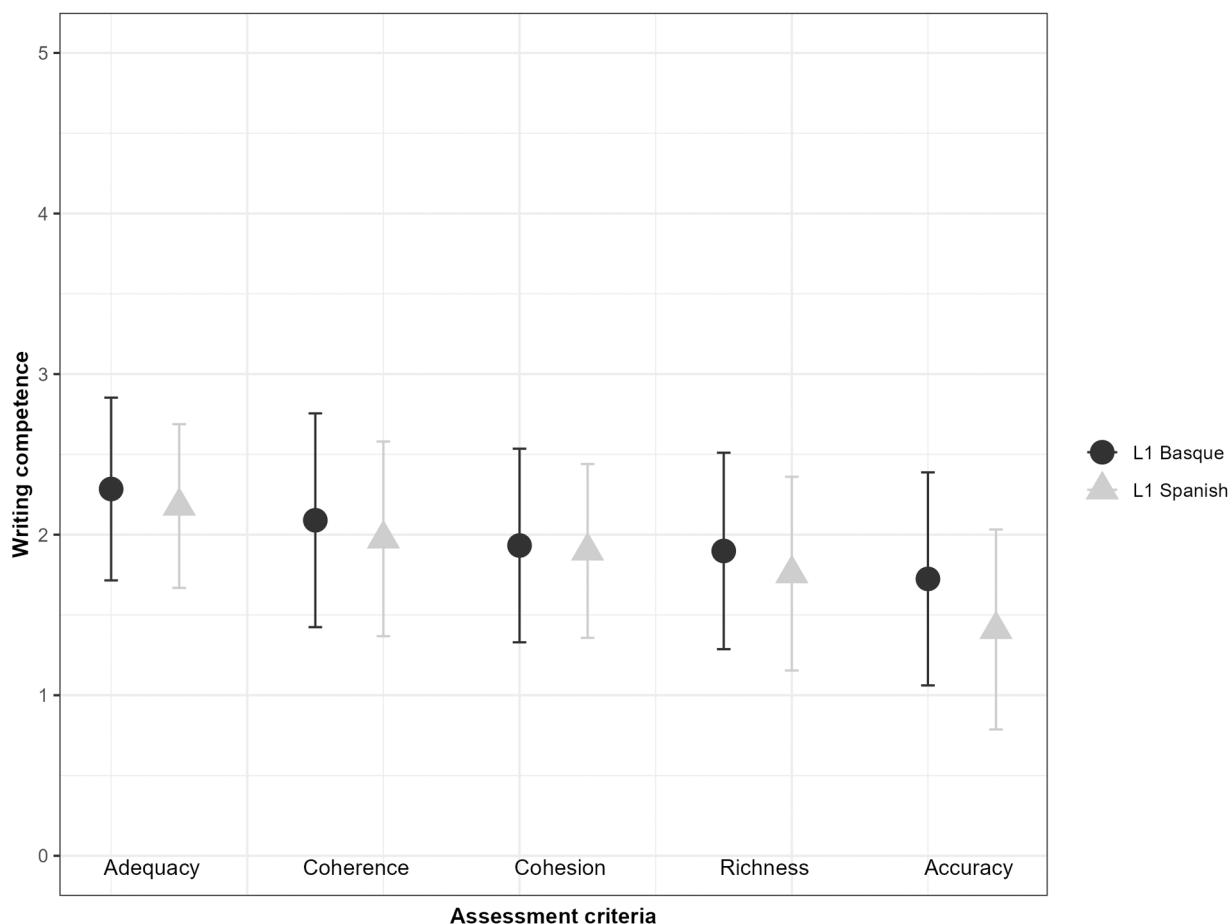


Fig. 3. Mean plots by assessment criteria and L1.

key points: the superior performance of candidates whose native language is Basque and the distinctive nature of the accuracy criterion. This criterion showed the highest thresholds in the two L1 groups, which is interpreted as indicative of its difficulty, and additionally, the alignment procedure concluded lack of invariance with reference to the candidates' mother language.

In terms of the second research question, the results confirm the distinction in writing competence between candidates whose native language is Basque or Spanish, whereas the study of the effects associated to the test edition has shown a negligible impact. Although some differences were detected among the five editions, the effect sizes were minimal for most of them.

Finally, the study of each of the assessment criteria showed that differences between candidates' groups was not equal for all of them. The biggest differences were found in the richness and accuracy criteria, which are the criteria where the mean values were the lowest, indicating they are the most challenging criteria. It is particularly noteworthy that the Cohen's *d* value for accuracy was .48,

results that can be linked with the lack of invariance on thresholds discovered in the measurement invariance study. Candidates whose native language was Spanish tended to achieve lower scores in the richness criterion compared to candidates with the same level of competence who belonged to the group whose native language was Basque.

The recommendation to be drawn from our findings is that L1=Spanish learners should be provided with systematic support if their writing skills are to be improved in the five evaluation criteria analysed in this study, but especially in the case of richness and accuracy. Those learners who have Spanish as their L1 should be invited to pay particular attention to general lexicon and its correct use within appropriate complex structures. Since all these learners share the same L1, customized writing interventions would be easier to implement than in other contexts where the variation of learners' L1s is wider. If they are to become proficient L2 writers, adult Basque learners have to build Basque knowledge, that is, they need to hone their vocabulary knowledge. Since new speakers are fundamental in reversing language shift because they are responsible for the dramatic increase in the number of minority language speakers in Spain in the last few decades (Urla et al. 2018), support measures to hone their writing skills are sorely needed. This is especially important in minority language contexts such as the BAC because of its situation of diglossia throughout history, as most Basques, being basically an illiterate group, have had little contact with the written version of the language until very recently. This past diglossia is in sharp contrast with the current spread of Basque in the educational system that is unparalleled in history. However, although it could be expected that new speakers might have higher accuracy due to the increasing presence of Basque in formal contexts such as school and university, the fact is that for them Basque is a school/academic language rather than a language they use colloquially on a daily basis.

With this context in mind, it has to be noted that all input (including any kind of multimodal input; see Manchón & Roca de Larios, 2023) and output (Swain, 1985) are relevant. However, and taking into account the close connection between L2 reading and writing and that L2 vocabulary knowledge is one of the most important predictors of L2 writing (Kim et al. 2022), a potential intervention should include raising awareness activities about the importance of reading in Basque for L1 =Spanish learners if they are to improve their writing skills. Since the control of the written language becomes automatized through massive reading and writing practice (Hulstijn, 2024), such interventions would have a very positive impact on students' writing proficiency, especially at a time when reading and writing on social media is overwhelmingly limited to short messages, which is why more complex literacy-related practices are sorely needed. Since L2 writing development entails "a gradual progression from declarative knowledge to procedural knowledge and then to the automatization of procedural knowledge" (Roca de Larios, Manchón & Murphy, 2006; 108), we could surmise that not all the L1 =Spanish writers in our study had reached the automatization of procedural knowledge in their L2, Basque.

Therefore, instruction on writing processes in the minority language should be considered, as it could help L2 learners to overcome the artificial separation between language instruction and learners' writing development. Teachers should go beyond the mere transmission of metalinguistic knowledge and strive to foster "students' thinking about the value and formation of their substantive expression" (Cumming, 1989; 83). The differences found in the richness and accuracy of the written texts reveal that L1 =Spanish students need to reflect on the importance of expressing ideas using appropriate vocabulary and complex structures, because this has a significant impact on the ideas they present in their writings. Such reflection should pave the way for greater attention not only to word choices (accuracy) –writers' lexical knowledge influences the quality of their texts and lexical diversity correlates significantly with holistic ratings (Schoonen et al., 2003)–, but also to a greater focus on the interrelation between planning and the actual production processes in a sophisticated manner (richness). This is why, while focusing on the gist of their written text, students also need to become aware of the importance of paying heed to the interaction of multiple aspects of their writing (adequacy, coherence, cohesion, richness and accuracy) from a holistic perspective. In this way, instructional approaches will prompt students not to focus just on a single aspect of their writing but rather on the interplay between multiple aspects (Cumming, 1989; Sasaki & Hirose, 1996; Schoonen et al., 2003),

Future research could embark on analyzing the impact that such interventions may have on the development of learners' writing skills. Replication studies in other contexts involving different L1s but including the same evaluation criteria used in this study would also be welcomed, especially if we take into account that most research on first and second language acquisition has been undertaken in Europe and North America, which may have "created a Western bias in theory construction and empirical research" (Hulstijn, 2024; 8).

This research suffers from some limitations. The study was only focused on writing in Basque, but future studies could examine the relationship between writing in both Spanish and Basque. Last but not least, and due to the complexity of having a very large sample, the study relied on a single writing task to tap into writing ability. Future research may include different writing tasks because this would allow researchers to rest on firmer ground to generalize about the impact of learners' L1 on writing performance.

Funding

This work is supported by the Ministry of Science, Innovation and Universities, Spain, Grant number PID2024-155252OB-100 funded by MICIU/AEI/10.13039/501100011033, and Basque Government, Spain, grant number IT1426-22.

CRedit authorship contribution statement

Elosua Paula: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Lasagabaster David:** Writing – review & editing, Writing – original draft, Validation, Project administration, Investigation, Conceptualization.

Data availability

The authors do not have permission to share data.

References

- Ahmadian, Mostafa Papi, Olena Vasylets, & Mohammad Javad. (2022). Individual Difference Factors for Second Language Writing. *The Routledge Handbook of Second Language Acquisition and Individual Differences*. Routledge.
- Angelis, G., & Jessner, U. (2012). Writing across languages in a bilingual context: A dynamic systems theory approach. In R. M. Manchón. Hrgs. (Ed.), *L2 writing development: Multiple perspectives* (pp. 47–68). Berlin: De Gruyter Mouton.
- Asparouhov, Tihomir, & Muthén, Bengt (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495–508. <https://doi.org/10.1080/10705511.2014.919210>. Routledge.
- Basque Government. (2023). *VII Encuesta sociolingüística*. Vitoria-Gasteiz: Basque Government.
- Beiler, Ingrid Rodrick (2021). Marked and unmarked translanguaging in accelerated, mainstream, and sheltered English classrooms. *Multilingua*, 40(1), 107–138. <https://doi.org/10.1515/multi-2020-0022>
- Crossley, Scott A., & McNamara, Danielle S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18(2), 119–135. <https://doi.org/10.1016/j.jslw.2009.02.002>
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39(1), 81–135. <https://doi.org/10.1111/j.1467-1770.1989.tb00592.x>
- Dąbrowska, Ewa (2012). Different speakers, different grammars. *Linguistic Approaches to Bilingualism*, 2(3). <https://doi.org/10.1075/lab.2.3.01dab>
- Elosua, P. (2011). Assessing measurement equivalence in ordered-categorical data. *Psicológica*, 32(2), 403–421.
- Elosua, P., & De Boeck, P. (2020). Educational assessment issues in linguistically diverse contexts: a case study using a generalised linear mixed model. *Language, Culture and Curriculum*, 33(3), 305–318. <https://doi.org/10.1080/07908318.2019.1662432>
- European Council. (2021). *Common European Framework of Reference for Languages*. European Council.
- Gabriel, Christoph, Klinger, Thorsten, & Usanova, Irina (2021). VOT production, writing skills, and general proficiency in multilingual learners of French: approaching the intertwining of different linguistic levels. *Multilingua*, 40(6), 745–769. <https://doi.org/10.1515/multi-2021-0016>
- Hu, Li-tze & Peter M. Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. Routledge. 6(1), 1–55. doi:10.1080/10705519909540118.
- Hualde, José Ignacio & Koldo Zuazo. 2007. The standardization of the Basque language. *Language Problems and Language Planning*. John Benjamins. 31(2). 143–168. doi:10.1075/lplp.31.2.04hua.
- Hulstijn, J. (2024). Predictions of individual differences in the acquisition of native and non-native languages: An update of BLC theory. *Languages*, 9(5), 173. <https://doi.org/10.3390/languages9050173>
- Khushik, Ghulam Abbas, & Huhta, Ari (2020). Investigating syntactic complexity in EFL learners' writing across common european framework of reference levels A1, A2, and B1. *Applied Linguistics*, 41(4), 506–532. <https://doi.org/10.1093/applin/amy064>
- Kidd, Evan, Donnelly, Seamus, & Christiansen, Morten H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, 22(2), 154–169. <https://doi.org/10.1016/j.tics.2017.11.006>
- Kim, Minkyung, Scott A. Crossley & Byeong-Kweon Kim. 2022. Second language reading and writing in relation to first language, vocabulary knowledge, and learning backgrounds. *International Journal of Bilingual Education and Bilingualism*. Routledge. 25(6). 1992–2005. doi:10.1080/13670050.2020.1838434.
- Kobayashi, H., & Rinnert, C. (2013). L1/L2/L3 writing development: Longitudinal case study of a Japanese multi-competent writer. *Journal of Second Language Writing*, 22(1), 4–33. <https://doi.org/10.1016/j.jslw.2012.11.001>
- Kormos, Judit (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing. Exploring L2 Writing-SLA Interfaces*, 21(4), 390–403. <https://doi.org/10.1016/j.jslw.2012.09.003>
- Kuiken, Folkert, & Vedder, Ineke (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321–336. <https://doi.org/10.1177/0265532216663991>
- Kuiken, Folkert, & Vedder, Ineke (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29(2), 192–210. <https://doi.org/10.1111/ijal.12256>
- Lan, Ge, Zhang, Qiusi, Lucas, Kyle, Sun, Yachao, & Gao, Jie (2022). A corpus-based investigation on noun phrase complexity in L1 and L2 English writing. *English for Specific Purposes*, 67, 4–17. <https://doi.org/10.1016/j.esp.2022.02.002>
- Lasagabaster, D. (2017). Language Learning Motivation and Language Attitudes in Multilingual Spain From an International Perspective. *The Modern Language Journal*, 101(3), 583–596. <https://doi.org/10.1111/modl.12414>
- Manchón, R. M., & Roca de Larios, J. (2023). The study of L2 writing processes: Lines and methods of inquiry. In R. M. Manchón, & J. Roca de Larios (Eds.), *Research Methods in the Study of L2 Writing Processes* (pp. 6–31). Amsterdam: John Benjamins.
- Meredith, William (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- Millsap, Roger E. (2011). *Statistical Approaches to Measurement Invariance*. New York: Routledge. <https://doi.org/10.4324/9780203821961>
- Millsap, Roger E. & Jenn Yun-Tein. 2004. Assessing Factorial Invariance in Ordered-Categorical Measures. *Multivariate Behavioral Research*. Routledge. 39(3). 479–515. doi:10.1207/S15327906MBR3903_4.
- Murphy, Liz, & Roca de Larios, Julio (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing*, 19(2), 61–81. <https://doi.org/10.1016/j.jslw.2010.02.001>
- Muthén, Bengt & Tihomir Asparouhov. 2018. Recent Methods for the Study of Measurement Invariance With Many Groups: Alignment and Random Effects. *Sociological Methods & Research*. SAGE Publications Inc. 47(4). 637–664. doi:10.1177/0049124117701488.
- Muthén, L.K. & B.O. Muthén 2023. Mplus user's guide. Eighth. Muthén and Muthén.
- O'Rourke, Bernadette, Pujolar, Joan, & Ramallo, Fernando (2015). New speakers of minority languages: the challenging opportunity – Foreword. *International Journal of the Sociology of Language*, 2015(231), 1–20. <https://doi.org/10.1515/ijsl-2014-0029> (De Gruyter Mouton).
- O'Rourke, Bernadette, & Ramallo, Fernando (2015). Neofalantes as an active minority: understanding language practices and motivations for change amongst new speakers of Galician. *International Journal of the Sociology of Language*, 2015(231), 147–165. <https://doi.org/10.1515/ijsl-2014-0036> (De Gruyter Mouton).
- Ortega, Ane, Urla, Jacqueline, Amorrtortu, Estibaliz, Goirigolzarri, Jone, & Uranga, Belen (2015). Linguistic identity among new speakers of Basque. *International Journal of the Sociology of Language*, 2015(231), 85–105. <https://doi.org/10.1515/ijsl-2014-0033> (De Gruyter Mouton).
- Pedley, Malika, & Viaut, Alain (2019). What do minority languages mean? European perspectives. *Multilingua*, 38(2), 133–139. <https://doi.org/10.1515/multi-2018-0025>
- Perin, Dolores, Susan De La Paz, Kelly Worland Piantadosi & Megan Madigan Peercy. 2017. The Writing of Language Minority Students: A Literature Review on Its Relation to Oral Proficiency. *Reading & Writing Quarterly*. Routledge. 33(5). 465–483. doi:10.1080/10573569.2016.1247399.
- Phuoc, Vo. Dinh, & Barrot, Jessie S. (2022). Complexity, accuracy, and fluency in L2 writing across proficiency levels: A matter of L1 background? *Assessing Writing*, 54, Article 100673. <https://doi.org/10.1016/j.asw.2022.100673>
- Roca de Larios, J. R., Manchón, R. M., & Murphy, L. (2006). Generating text in native and foreign language writing: A temporal analysis of problemsolving formulation processes. *The Modern Language Journal*, 90(1), 100–114. <https://doi.org/10.1111/j.1540-4781.2006.00387.x>
- Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, 46(1), 137–174. <https://doi.org/10.1111/j.1467-1770.1996.tb00643.x>
- Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning*, 53(1), 165–202. <https://doi.org/10.1111/1467-9922.00213>

- Schoonen, R., & van Vuuren, S. (2022). Transfer, writing, and SLA: L2 writing as a multilingual event. In R. M. Manchón, & C. Polio (Eds.), *The Routledge Handbook of Second Language Acquisition and Writing* (pp. 97–108). New York: Routledge.
- Schoonen, Rob, Amos van Gelderen, Stoel, Reinoud D., Jan Hulstijn, & Kees de Glopper. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning*, 61(1), 31–79. <https://doi.org/10.1111/j.1467-9922.2010.00590.x>
- Ströbel, Marcus, Kerz, Elma, & Wiechmann, Daniel (2020). The relationship between first and second language writing: Investigating the effects of first language complexity on second language complexity in advanced stages of learning. *Language Learning*, 70(3), 732–767. <https://doi.org/10.1111/lang.12394>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, Massachusetts: Newbury House.
- Tullock, B. D., & Fernández-Villanueva, M. (2013). The role of previously learned languages in the thought processes of multilingual writers at the Deutsche Schule Barcelona. *Research on the Teaching of English*, 47, 420–441. <http://www.jstor.org/stable/24397846>.
- Urla, Jacqueline, Estibaliz Amorortu, Ortega, A., Ane, & Goirizolzarri, Jone, J. (2018). Basque Standardization and the New Speaker: political Praxis and the Shifting Dynamics of Authority and Value. In P. Lane, J. Costa, & H. Korne. Hrsg. (Eds.), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery* (pp. 24–46). New York: Routledge.
- Ushioda, Ema, & Dörnyei, Zoltán (2017). Beyond Global English: Motivation to Learn Languages in a Multicultural World: Introduction to the Special Issue. *The Modern Language Journal*, 101(3), 451–454. <https://doi.org/10.1111/modl.12407>
- Weijen, Daphne van, van den Bergh, Huub, Rijlaarsdam, Gert, & Sanders, Ted (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing*, 18(4), 235–250. <https://doi.org/10.1016/j.jslw.2009.06.003>
- Wobbrock, Jacob O., Findlater, Leah, Gergle, Darren, & Higgins, James J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11* (pp. 143–146). New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1978942.1978963>

Paula Elosua PhD in Psychology and full professor of psychometrics at the University of the Basque Country. She heads a psychometrics research group specialising in language, educational and psychological assessment, test and questionnaire construction/adaptation, and validation studies.

David Lasagabaster is Full Professor of Applied Linguistics at the University of the Basque Country (UPV/EHU), Spain. He has published widely in international journals and books on second/third language acquisition, multilingualism, attitudes and motivation, and EMI (English-Medium Instruction).