

Multilingual disciplinary literacies: exploring developmental patterns of science writing across secondary education

Roberto Arias-Hermoso^{*1}, Ainara Imaz Agirre¹ and Eneritz Garro Larrañaga¹

* Roberto Arias-Hermoso (corresponding author). rarias@mondragon.edu, <https://orcid.org/0000-0002-0780-7320>

Ainara Imaz Agirre. aimaz@mondragon.edu, <https://orcid.org/0000-0001-6415-1586>

Eneritz Garro Larrañaga. egarro@mondragon.edu, <https://orcid.org/0000-0002-6148-8383>

¹ Faculty of Humanities and Education Sciences, Mondragon Unibertsitatea

This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Multilingual and Multicultural Development on March 24, 2025. It is available at: <https://doi.org/10.1080/01434632.2025.2481200>

Full citation:

Arias-Hermoso, R., Imaz Agirre, A., & Garro Larrañaga, E. (2025). Multilingual disciplinary literacies: exploring developmental patterns of science writing across secondary education. *Journal of Multilingual and Multicultural Development*, 1–19. <https://doi.org/10.1080/01434632.2025.2481200>

Multilingual disciplinary literacies: exploring developmental patterns of science writing across secondary education

This study examines multilingual disciplinary literacies from a developmental perspective, a still relatively underexplored matter. Using a cross-sectional design, 535 multilingual learners enrolled in Basque immersion programmes across the four years of secondary education (12-16 year old students) completed a disciplinary source-based writing task in the three languages of the curriculum: Basque, Spanish and English. They were asked to produce two cognitive discourse functions: *argue* and *compare*. Texts were coded with an analytic rubric consisting of 11 items, and linear mixed models were used to assess the effect of Year, Testing Language and L1. Findings showed that all items improved in all three languages as students progressed through secondary education, with non-linear and gradual patterns in certain moments. While diverse patterns across languages were found, parallel developmental patterns took place, particularly between Basque and Spanish. Some measures showed varying developmental patterns for L1 groups. In conclusion, results support multilingual writing development and language transfer theories, and provide evidence for the development of multilingual disciplinary literacies. This study presents important theoretical and practical implications for multilingual education and disciplinary literacies.

KEYWORDS

Cognitive discourse functions, disciplinary literacy, writing development, multilingualism, language acquisition

1. Introduction

Language plays a crucial role in the learning-teaching process, as it is through language that students understand, construct and negotiate new academic content (Li et al., 2022). It has been highlighted that learners require sufficient academic language competence to navigate the learning process (Cummins, 2021; Lorenzo & Trujillo, 2017). In other words, students must deal with multiple language-related structures that are specific to academic contexts, i.e., cognitive academic language proficiency (CALP, Cummins, 1980). Recently, research has shifted from a general vision of academic language to exploring the specific needs of each discipline, understanding that each discipline constructs and communicates knowledge in subject-specific ways (Dalton-Puffer et al., 2024; Moje, 2015; Shanahan & Shanahan, 2008): disciplinary literacies (DLs). DLs have been proposed as key to the integration of content and language, which is a complex matter (Nikula et al., 2016). Said integration has often been studied from a second language (L2) perspective, however, DLs must be developed in all languages of the curriculum. In line with Pascual i Granell (2024), we understand integration as the combined teaching-learning of content and the language needed to express it, irrespective of it being in a first (L1), second or third language (L3). In fact, it is common that learners in bi/multilingual education programmes use various languages in the learning process (Cenoz, 2023; Lorenzo & Trujillo, 2017), and therefore, educational approaches that foster students' multilingual competence need to be considered (Cenoz & Gorter, 2011; Usanova & Schnoor, 2022).

There is an overall agreement that learners' DLs develop over time (Granados et al., 2023), and that they might be transferred across languages (Arias-Hermoso & Imaz Agirre, 2024; Granados et al., 2021). However, research on multilingual disciplinary literacies (MDLs, Hüttner & Llinares, forthcoming) including more than two languages is still scarce, as well as MDLs' developmental patterns, which is an issue of interest for the international community (Ting et al., 2024). The present article, therefore, focuses on studying developmental patterns in MDLs by analysing scientific students' production of cognitive discourse functions (CDFs, Dalton-Puffer, 2013) in a source-based writing (SBW) task in their three languages of schooling: Basque, Spanish and English. We adopt a cross-sectional and multilingual approach for studying language development in MDLs by exploring developmental patterns across secondary education.

2. Cognitive discourse functions: *argue* and *compare*

The construct of CDFs (Dalton-Puffer, 2013) was initially created by considering both educational and linguistic theories, and, after a review of the operative verbs used in educational-academic contexts, a seven-category construct was created (see Dalton-Puffer, 2013 for more information). The seven categories group together operative verbs that share the same underlying communicative intention, and include *categorise*, *define*, *describe*, *evaluate*, *explore*, *explain* and *report*. These verbs represent the linguistic realisations of cognitive actions that are crystallised in classroom contexts and are present in diverse aspects of the teaching-learning process, such as classroom interaction (Lersundi et al., forthcoming), assessment (e.g., del Pozo, 2024; Llinares & Morton, 2024) or written exams (Breeze & Dafouz, 2017). Learners are often required to perform CDF tasks (Gerns, 2023), which could be defined as a type of task that prompts students to perform one or more CDFs. Using the prompt in the study by Evnitskaya and Dalton-Puffer (2023) as an example, a CDF task could look like this: *Compare that age [the historical period learners had chosen to describe in their blog entry] with our life today*. CDF production is closely related to the formulation of the prompt, although students do not always produce the type of CDF the prompt intended to elicit (Hüttner & Smit, 2024; Llinares & Nikula, 2023). Among these CDFs, this study focuses on two: *argue* (part of *evaluate*) and *compare* (part of *categorise*).

Arguing has been studied in depth in the literature, and has usually been defined as defending one's position using evidence (Dalton-Puffer, 2016; Toulmin, 1958), often including empirical or experimental data (Jiménez-Aleixandre & Erduran, 2007; Polias, 2016). Learning to produce successful arguments is crucial for reflective and critical thinking (Pylonitis & Meyer, 2024), however, it might be challenging both cognitively and linguistically (Crossley et al., 2022). Student argumentation has typically been operationalised following Toulmin's (1958) proposal (TAP), which includes elements basic to argumentation: claim, supporting data, warrants, backings, rebuttals and qualifiers. Previous research using TAP, however, have modified or adapted some of the elements by leaving backings and qualifiers aside, or by including counterarguments, a crucial component of argumentation (Liu & Stapleton, 2014). According to several studies (Arias-Hermoso et al., 2024; Crammond, 1998; Crossley et al., 2022; Liu & Stapleton, 2014), every argumentative sequence needs to include a claim and supporting data to show some extent of reasoning. However, other elements (warrants, counterarguments or rebuttals) could be seen as complementary and a sign of more complex argumentation skills.

Comparing is a less studied CDF, and previous research has debated its complexity, with some scholars arguing that it is a relatively easy function (Anderson & Krathwohl, 2001; Llinares & Nikula, 2023) while others highlighting that it is challenging and demanding (Chen & Zhou, 2022; Tullis & Goldstone, 2016). Still, it has been argued that it is crucial in education, especially to connect previous knowledge with new concepts (Coyle & Meyer, 2021; Gerns, 2023a). It has been defined as establishing connections and/or differences between two or more concepts by exploring certain criteria. Similarly to argumentation, some elements for the comparisons have been emphasised as basic elements, such as the inclusion of the concepts being compared or the criterion according to which they are being compared (Chen & Zhou, 2022, Evnitskaya & Dalton-Puffer, 2023, 2024; Gerns, 2023a, 2023b). Gerns (2023a, 2023b) also highlights the importance of the comparative structure, i.e., how the information about the items being compared is organised. In her studies, she underscored two main methods: point-by-point (with an emphasis on the criteria) or block-method (with an emphasis on the comparative items). A final remark of comparing is the precedence of difference-based comparisons over similarity-based comparisons, a result suggested by most studies (Chen & Zhou, 2022, Evnitskaya & Dalton-Puffer, 2023, 2024; Gerns, 2023a, 2023b).

3. Multilingual developmental patterns and CDFs

Building on research on CDFs and DLs reviewed above, our study also integrates theoretical

perspectives on writing development and multilingual writing, and more importantly, their interplay: the progression of writing across multiple languages. A critical gap in the literature emerges when considering the interplay among these, as much research on writing development has focused on L2 development (Durrant et al., 2021), and multilingual writing research including more than two languages still needs to be explored further (Gogolin et al., 2022). Studies on writing development, in general, indicate that writers' performance becomes increasingly complex, with longer language units and fewer errors, and display more varied grammatical structures (see, for example, Durrant et al., 2021; Egli Cuenat, 2017; Staples et al., 2022; Whittaker et al., 2011). Nevertheless, the multifaceted, multilayered and non-linear nature of writing development has also been highlighted (Larsen-Freeman, 2009; Michel, 2017), showing that not all measures of writing development might develop in parallel, and some might show stagnation or decrease points (e.g., Lorenzo & Rodríguez, 2014; Lahuerta, 2018). Research on multilingual writing shows that, on the whole, writers' texts present across-language similarities (Granados et al., 2021; Arias-Hermoso & Imaz Agirre, 2024; Orcasitas-Vicandi, 2021; Usanova & Schnoor, 2022, 2024; Rinnert & Kobayashi, 2016; Forbes, 2021). The studies support Cummins' Common Underlying Proficiency (CUP, 1980), which understands that academic language skills are positively transferred across one's languages in multiple directions.

This study, therefore, draws on both developmental and multilingual research on writing to examine writing development and assess similar or differential patterns in students' development, as research has suggested that some aspects might develop in parallel across one's languages, but some others might not (e.g., Egli Cuenat, 2017; Granados et al., 2021), as they may be influenced by other factors, such as learners' L1 (Staples et al., 2022). The above-mentioned studies have significantly contributed to our understanding of multilingual writing development, but the majority has approached it from a complexity, accuracy and fluency (CAF) perspective, often not including all languages of the curriculum, not focusing on development, or failing to address DLs. However, to understand the development of MDLs, a more functional perspective is needed, which is where the CDF construct plays an important role.

3.1. The production of CDFs in more than one language

Regarding the production of CDFs in more than one language, studies point out crosslinguistic similarities in bi/multilingual production of CDFs (Arias-Hermoso et al., 2024, 2025; Evnitskaya & Dalton-Puffer, 2023; Llinares & Morton, 2024; Llinares & Nashaat-Sobhy, 2021; Nashaat-Sobhy & Llinares, 2023), closely aligning with Cummins' CUP (1980). Nevertheless, some nuanced differences across languages also emerged. For example, when comparing L1 Spanish and L2 English *definitions* in science or history, students tended to use relatively more expansions in Spanish, and more circumstantial expansions to the definition in English (Llinares & Nashaat-Sobhy, 2021; Nashaat-Sobhy & Llinares, 2023). Similar findings were reported by Llinares and Morton (2024) when analysing students' production of the CDF *explore*, using data from the same context, indicating similarities in the frequencies of students' explorations and modalization across L1 and L2. Evnitskaya and Dalton-Puffer (2023) analysed oral categorisations in history and science in L1 Spanish and L2 English and reported similarities across languages regarding the structural features and lexicogrammar used. Arias-Hermoso et al. (2024, 2025), comparing the production of scientific arguments by the same group of Year 9 students, found no difference between languages between L1 Spanish, L2 Basque and L3 English, as well as positive significant correlations across languages, even though students used more varied lexicogrammar choices in Spanish and Basque than in English.

3.2. The production of CDFs across ages

While the development of MDLs and CDFs is a topic that has drawn international attention (Ting et al., 2024), research is still scarce. It is generally believed that students' proficiency when performing CDFs will become more complex and accurate over the ages (Granados & Lorenzo, 2024). However, to our knowledge, this has received little attention in the literature, with few studies exploring CDF

production by learners of different ages. In a longitudinal setting focusing on the writing by bilingual children from the Madrid region in Spain, Nashaat-Sobhy and Llinares (2023) reported that Year 8 students outperformed Y6 students in their general production of definitions, as well as in producing formal (vs. semi-formal) definitions. Similarly, the study by Whittaker and McCabe (2023) focused on student evaluations across different disciplines and age groups, and suggested that learners' evaluations become increasingly justified in secondary (Year 8), leaving evaluative language behind, a characteristic of primary education learners' productions (Year 6). Evnitskaya and Dalton-Puffer (2024), focusing on a similar context, recently showed that older students (Year 8) produced generally more implicit comparisons in history than younger children in L2 English, probably due to conceptual abstraction. Finally, del Pozo (2024) and del Pozo and Llinares (2021) found that students in L2 English CLIL history improved from Y7 to Y9 when producing the CDF *report*. On the whole, the studies suggest that there are indeed certain developmental patterns, with students producing better, more complete and more complex CDFs, in line with Granados and Lorenzo (2024), and general writing development research (Durrant et al., 2021; Granados et al., 2021).

However, most of the studies reviewed examining CDFs do not consider both the production in more than one language and age-related research, and some do not provide inferential statistics to back their results. Therefore, the production of multilingual CDFs is still an underexplored topic, and even more so their developmental trajectories. Although research has provided evidence to support that one's academic and disciplinary literacy skills evolve over time (del Pozo, 2024; Durrant et al., 2021; Granados et al., 2023; Lorenzo, 2017; Lorenzo & Rodríguez, 2014) and can be taught (Garro Larrañaga et al., 2024; Gerns, 2023a, 2023b) the specific patterns that show change still need to be addressed in different languages to respond to current multilingual educational contexts. Consequently, it is crucial to adopt a multilingual perspective to MDL development (Granados et al., 2021; Usanova & Schnoor, 2024). In order to explore the aforementioned gaps, this study examines developmental patterns in MDLs by focusing on students' CDF production in three languages in a disciplinary source-based writing task. Therefore, one research objective motivates this study: *To explore multilingual parallel developmental patterns in students' disciplinary writing and the extent to which these vary among L1s*. To this end, the main objective has been divided into three specific research questions (RQs). For each of the RQs an hypothesis has been formulated, which are provided below as contextual information:

1. **RQ1. Does student disciplinary writing show developmental patterns across secondary education?** It is hypothesised that there will be developmental patterns in students' performance in the different dimensions explored (CDF performance and source-based writing skills), in line with previous studies addressing age-related differences in CDFs and MDLs (del Pozo, 2024; Granados et al., 2021, 2023; Whittaker & McCabe, 2023) or academic writing in general (Durrant et al., 2021; Staples et al., 2022). Non-linear development is expected (Larsen-Freeman, 2009; Michel, 2017), as well as different slopes for different items, with some displaying year-to-year differences and others gradual change.
2. **RQ2. Are the developmental patterns parallel across students' languages?** Drawing from Cummins' CUP and multilingual theories of writing (Cenoz & Gorter, 2011; Rinnert & Kobayashi, 2016), it is understood that crosslinguistic transfers occur in students' writing. Therefore, following Usanova and Schnoor (2024), as well as Granados et al. (2021), it is expected that students' development will show parallel patterns across their three languages, especially between the minority and majority languages, i.e., Basque and Spanish, as suggested by multilingual writing research in the same context (see Arias-Hermoso & Imaz Agirre, 2024; Orcasitas-Vicandi, 2021; Sagasta, 2003; Martín-Laguna, 2020).
3. **RQ3. Do these developmental patterns vary based on learners' L1?** We assume certain differences in students' performance according to their L1s (Basque, Spanish or both). In line with previous studies in the Basque context (Elosua & de Boeck, 2020; Elosua & Lasagaster,

2025), we expect different performances depending on students' L1, which will lead to varying developmental patterns, as also suggested by previous studies (Alzahrani, 2024; Durrant et al., 2021; Shen et al., 2023; Staples et al., 2022).

4. Methods

4.1. Sample

This study analyses secondary education students ($n = 535$) enrolled in five different schools in the Basque Autonomous Community (BAC), in which Basque holds a co-official status alongside Spanish. Despite its official status, Basque can still be considered a minority language due to the high presence of Spanish in society. The five schools were selected in order to represent the different sociolinguistic contexts, including schools from the three territories in the BAC (Bizkaia, Araba and Gipuzkoa). Two schools were from Spanish-dominant contexts, one from a Basque-dominant context and two from mixed contexts. This permitted including students from different sociolinguistic backgrounds and L1 profiles, with different exposures to each language. All learners were studying in a Basque Model D immersion programme (see Cenoz, 2023 for a description of the models), in which Basque is the language of instruction for all subjects, except for Spanish and English as a FL. Students, therefore, receive approximately 26 teaching hours in Basque (including Basque language classes), 3 in Spanish and 3 in English.

In order to track developmental patterns in writing, students from Years 7, 8, 9 and 10 participated (age range: 12-16), i.e., the four years of compulsory secondary education in the BAC. In each school, at least two classes of students took part. A sociolinguistic background questionnaire was used to elicit students' reported L1, i.e., defined here as the language(s) they most frequently used at home when all family members were present: Basque (BA), Spanish (SP) or both (BASP) (see Table 1).

Table 1. Number of participants (n) per school, year and reported L1

	Year 7			Year 8			Year 9			Year 10			TOTAL
	BA	SP	BASP	BA	SP	BASP	BA	SP	BASP	BA	SP	BASP	
School A	7	15	5	14	20	4	4	5	3	7	11	6	101
School B	9	12	0	13	11	9	16	10	8	5	3	5	101
School C	2	24	2	2	21	1	3	18	1	8	10	7	99
School D	16	11	5	15	10	5	16	10	6	12	7	5	118
School E	1	23	1	0	30	1	0	34	0	0	26	0	116
TOTAL	133			156			134			112			535

4.2. The corpus: collection and procedure

Each student completed three written essays in a SBW task, one in Basque, one in Spanish and one in English. SBW tasks are defined as a type of task that requires learners to produce written texts after reading or listening to a certain (multimodal) source(s), and it is a common practice in tertiary education. SBW tasks, therefore, activate both learners' comprehension and production sub-processes (van Weijen et al., 2019; Zhu et al., 2024). Therefore, the final corpus comprised 1605 texts (535 in each language). Students were asked to write about a discipline-related topic, i.e., Nutrition. The topic was selected following the educational curriculum of the BAC, as it is a topic that is worked across secondary education at different levels. The task was designed as a complex CDF-eliciting task (Breeze & Dafouz, 2017), in which two CDFs are connected, in this case, *argue* and *compare*. Students were thus required to take a stance on which one of the two foods selected (chicken and chocolate) was healthier, by comparing them and using arguments to provide their claims.

A SBW task was used to ensure that all students shared a baseline regarding the content knowledge

(Read, 1990). Consequently, a handout including written information in the form of tables and lists was presented to students. The handout described the main nutritional values of food, as well as a couple of ideas about their impact on the body and society/nature. Students were asked to take notes for three minutes on a separate sheet, as they were not allowed to use the sources during composition, but they could use their notes. After the three minutes for notes, they were given a maximum of 30 minutes to write a 150-word long essay. In order to control for potential task repetition effects, the order of the *test language* (TL) was counterbalanced in a pseudo-random order, with students completing the task in either of these sequences: Basque - Spanish - English, Spanish - English - Basque, English - Basque - Spanish. Each student was randomly assigned to one of these orders, with similar proportions of students within classes. It is important to note that students completed the task in the same language as the source, i.e., they received the source in Basque and wrote the composition in Basque.

4.3. The rubric

For the purposes of this study, an analytical rubric was developed through expert judgment (see Arias-Hermoso et al., under review, for a description of the validation process). Shortly described, an iterative process was used, in which a first version of the rubric was created based on the existing literature, and the version was submitted to 13 expert judges to assess each item. Regarding the expertise of the experts, all judges were university researchers holding a PhD degree, with a background in applied linguistics and/or educational science. All of them had experience researching and publishing in peer-reviewed journals about disciplinary literacies, CDFs and bilingual education. Additionally, two of them were Science Education experts. After their initial ratings, the rubric was modified and a second round of revisions was carried out, leading to the final version of the rubric¹. In the revision process, some items were removed, merged, or added, and the majority of items were reformulated to ensure coherence. For instance, the first version of the rubric did not include the *comparative structure* item, which was raised as an important aspect of the compare CDF. Also, the items *degree of transformation* and *inclusion of original ideas* in D3 were proposed as a single item in the first version, but they were split into two as suggested by the judges. Most items were reformulated to maintain consistency in the formulation patterns.

The rubric included three dimensions, each comprising 3 or 4 items (see Figure 1). In the first and final versions, D1 and D2 both included 4 items, while D3 only included 2 items in the first version, however, one of them was split into two separate items. The items included descriptors between 0 (the lowest score possible) and 3 (the highest score). The descriptors from 0 to 3 were selected due to being able to capture enough variability in students' performance, and most judges commented positively on the scale. The rubric was designed so that it could be used to analyse CDF productions and SBW in any language or discipline. It is important to note that the three dimensions of the rubric are included in the present paper: argumentation skills, comparison skills and SBW. Although the latter is not one of the main theoretical foci of this article, it remains a crucial methodological component of the study, as well as an important academic language competence (Cheong et al., 2021). Its inclusion in the article thus helps us provide a more detailed and comprehensive analysis of students' writing. For a theoretical contextualisation of this dimension, see van Weijen et al. (2019) or Zhu et al. (2024). After a second round of judgment by the experts, considering their qualitative and quantitative assessment, only some minor formulation changes were made. Therefore, the final version of the rubric was created. The three authors performed initial analyses with 12 texts, and while no inter-coder agreement quantitative measures were analysed, total qualitative agreement was reached through discussion. The remaining texts were coded by the first author.

¹Due to the word limit, it is not feasible to describe the rubric design process in depth in this paper. The readers are encouraged to read Arias-Hermoso et al. (under review) for a detailed description and the final version of the rubric including 0-3 descriptors.

Figure 1. Analytic rubric items, divided by dimensions (adapted from Arias-Hermoso et al., under review).

D1: ARGUMENTATION SKILLS		D2: COMPARISON SKILLS		D3: SOURCE-BASED WRITING	
Stance-taking (D1_ST)	This is defined as the writer's ability to express their claim or position towards the topic and take a stance.	Comparative items (D2_CI)	This is defined as the writer's ability to produce comparisons between the two concepts, by including differences and/or similarities, and elaborating on them.	Use and comprehension of sources (D3_UC)	This is defined as the writer's ability to convey their comprehension of the sources through the use of sources accurately in a relevant way.
Reasoning (D1_R)	This is defined as the writer's ability to provide supporting data and warrants to justify their stance.	Comparative criteria (D2_CC)	This is defined as the writer's ability to explicitly express the comparative criteria, i.e., the bases of comparison (the criterion according to which the similarities/differences are being made).	Degree of transformation (D3_DT)	This is defined as the writer's ability to transform the information from the sources via paraphrasing (lexical and structural changes).
Counterargumentation and rebuttal (D1_CR)	This is defined as the writer's ability to acknowledge and refute opposing views to their claim, by providing counterarguments and rebuttals based on data.	Comparative structure (D2_CS)	This is defined as the writer's ability to organise the comparisons logically, at sentence, paragraph and text levels.	Inclusion of original ideas (D3_O)	This is defined as the writer's ability to include ideas not present in the sources, (original ideas from the writer's background) by appropriately integrating external knowledge into the text.
Language for arguing (D1_LA)	This is defined as the writer's ability to use appropriate, accurate and varied lexicogrammatical resources to produce argumentation elements, including connectives, words to express cause and effect, evaluative language, modalisation...	Language for comparing (D2_LC)	This is defined as the writer's ability to use appropriate, accurate and varied lexicogrammatical resources to produce comparisons, including compare/contrast connectives, lexical and grammatical forms to express difference/similarity, juxtapositions...		

4.4. Analytical approach

Linear Mixed Models (LMMs) were used to answer the research questions, which support ordinal and non-parametric variables and have proven very useful when considering repeated measures (see Gries, 2015). LMMs include both fixed and random effects, as well as hierarchical or nested data. For this study, using the *lme4* package in R, we included *Year*, *L1* and *TL* as the fixed effects, and *sociolinguistic context*, *school* and the *participants* themselves as random effects. The dependent variables were the 11 items from the rubric. Wald tests were used to test the overall effect of the fixed effects. After that, post hoc Tukey comparisons were conducted using the *emmeans* function to test pairwise differences across Years within each TL and L1, including Contrast Estimates (CEs), effect sizes (Cohen's *d*) and significance (*p*) (see Supplementary materials for the results).

For the purposes of this study, it is crucial to define (*parallel*) *developmental patterns*. Developmental patterns are defined as significant differences across Years, including year-to-year comparisons or gradual comparisons (Y7-Y10 or two-year cycles). Research on *parallel* developmental patterns, however, is scarce and might be challenging to define and operationalise. Staples et al. (2022) examined convergence in confidence intervals (CIs) across L1/L2 learners' productions of grammatical aspects. Granados et al. (2021) used a Pearson correlation between the degree of evolution across L1 and L2 after a LMM. These approaches, nevertheless, present challenges for our data, as 1) we do not focus on divergence/convergence in CIs/means, but on the *change* across years, and 2) our dataset is not longitudinal. Usanova and Schnoor (2024) proposed that parallel patterns could be seen as comparable or divergent based on descriptive change and visual graphs.

Considering the complexity of the issue, a novel approach was used. Cohen's *d* was calculated to estimate change in year-to-year comparisons, using Plonsky and Oswald's (2014) 0.4 threshold for a small effect. Then, the differences between effect sizes were calculated, and a threshold of ± 0.4 was used to estimate whether the patterns could be parallel across L1s and/or TLs. For instance, if L1 Spanish learners showed an effect size of 0.8 and 0.7 respectively, when comparing Year 7 and Y8 scores in Basque and English, it was assumed that parallel development in these two languages occurred. We acknowledge that this threshold might be arbitrary to a certain extent, as it is used as a measure of effect size, and not of similarity/divergence *between* effect sizes. However, using this cutoff could help shed light on whether the patterns were similar using a consistent method, in

addition to visual inspection of the graphs (see Appendix 2).

5. Results

5.1. Developmental Patterns Across Years

The results from the LMM model showed a significant effect of the fixed effects (Year, L1, TL) for most items. However, the random effects models including *school*, *sociolinguistic context* and *participants* presented singularity issues due to insufficient variance in the data. To address this, alternative models were tested, including both *school* and *sociolinguistic context*, only one of these factors, or none. The model excluding both factors (AIC = 3620.4) was selected as the final model, as it had the lowest AIC value compared to the full model (AIC = 3623.5), the model including only *school* (AIC = 3621.5) or only *sociolinguistic context* (AIC = 3622.4). As the exclusion of these factors did not significantly impact model fit ($ps > 0.3$), the simplest model, which excluded both factors, was selected as the final model. This model included the three main fixed effects and *participants* as the random effect. Wald III tests were conducted to assess the overall impact of each fixed effect and their interactions (Table 2).

Table 2: Results of Wald III Tests for all variables (χ^2 values, p-values represented by asterisks).

Item	Year	L1	TL	Year*L1	Year*TL	TL*L1	Y*TL*L1
D1_ST	34.85***	8.413*	24.555***	5.95	13.402*	2.59	12.2
D1_R	55.124***	6.46*	20.14***	9.91	2.482	2.661	12.56
D1_CR	47.269***	10.278**	10.478**	5.53	7.567	10.03*	15.8338
D1_LA	141.766***	8.63*	21.14***	17.073**	16.99**	5.14	7.53
D2_CI	68.008***	3.827	4.927	11.731	8.984	10.884*	18.394
D2_CC	70.566***	2.14	4.858	6.716	9.675	2.048	15.958
D2_CS	29.551***	0.152	0.2571	24.0364***	19.879**	11.9782*	29.43**
D2_LC	106.743***	16.533***	10.917**	10.837	14.36*	14.1**	6.949
D3_UC	50.137***	6.473*	2.564	12.323	2.484	3.88	2.704
D3_DT	68.575***	2.42	9.604**	20.978**	9.196	0.95	9.683
D3_O	9.717*	2.3871	9.991**	14.391*	8.701	4.906	13.969

Note: ***p < 0.001; ** p < 0.01, * p < 0.05

Year had a significant effect on all 11 items. Appendix 1 shows the mean scores for the 11 items, divided by Years, L1 and TLs. Overall, results indicate that there is a significant increase from Y7 to Y10 in all items (see Table 3 for an example), as confirmed by the average improvement (average of Y7-Y10 CEs within each dimension), with an improvement of 0.831 in Basque, 0.601 in Spanish and 0.611 in English for argumentation skills. Similarly, comparison skills showed an improvement of 0.809 in Basque, 0.784 in Spanish, and 0.652 in English. Finally, the use of sources also improved significantly, with average values of 0.652, 0.659 and 0.529, respectively. As can be seen in the mean scores in Appendix 1, all items did not show the same developmental patterns, with some being relatively high already at Y7 (e.g., *stancetaking* or the *use of sources*), while others showed a steeper change from Y7 to Y10 (e.g., D1_LA and D2_LC, *language for arguing* and *language for comparing*). In the case of *inclusion of original ideas*, despite the significant effect of Year, the change was smaller, since students at Y10 still showed difficulties in producing original ideas.

Table 3: Two examples from the corpus, from Year 7 and Year 10

Year 7 text (003_E1_z_EN)	Year 10 text (100_E4_z_EN)
The chicken is healthier than the chocolate because: 100g of chicken has 99,2 kcal, 1 g total fat, 0,46 g Sugars, 21,9g Proteines, 11 mg calcium, 0'36g salt. 100 g of chocolate has 531 kcal, 29g Total fat, 57g sugars, 7,4g Proteins, 247 mg Calcium, 0,27 g salt. I like the chicken that prepare my grandmother it has carrot, potatoes, onions and aceite a lot of salt. The chicken prepare a lot of diferent types. I like petxuga rebozada with ketchup. I like the chocolate in bread to breakfast it is very good por ejemplo I ate chocolate with bread yesterday it was excelent. I like liquid chocolate too. I don't know any persone that don't like chocolate. I prefere the chocolate.	In my opinion the chicken is the best chose if you are indecis between the chicken and the chocolate. For one hand, the chicken is the best option because it hasn't to many salt and sugar, conretly it has 0,49 g of sugar and 0,36 g of salt. The chicken it is very used for any type of diets. Also, it helps to the nervous system and digestive system. Such as, if you have nerviosism the chicken can help to you. It is very recomendated to eat chicken before or after you do sport, it gives to you to mach energy. However it has a bad detail if you don't cook the chicken it can intoxicate you, but I think that it is healthier than the chocolate, also it is very nutritious. For other hand, I don't think that the chocolate is the best option, I know that it is more delicious than the chicken but the animals are killed in the process. It control our cholesterol and affect our state of mind. But it also has negative consecuencias, it has too many calories, concretly 531 kcal and if you eat too many chocolate you can take too much weight in a short period. So the conclusion is that for my the chicken is healthier than the chocolate.

However, as shown in Table 2, the Wald III Test held a significant effect of L1 and TL in some measures, as well as interactional effects. As displayed in the graphs in Appendix 2 (see also Supplementary materials for post hoc test results), significant differences were not always found in year-to-year comparisons, but often between two-year cycles (between Y7 and Y9, or Y8 and Y10). This is also reflected in the effect size coefficients, which tend to be small in year-to-year comparisons, but medium-strong in two-year cycles and from Y7 to Y10. This pattern was especially salient in BASP learners in all three dimensions, and BA learners in D1 and D3, while SP learners tended to show significant year-to-year differences in all three dimensions. Nevertheless, it is important to note that SP learners showed stagnation patterns in most items after Y9, reaching a plateau in their development, something that BA and BASP learners did not systematically present. Therefore, although it can be concluded that all measures showed positive improvement from Y7 to Y10, it is necessary to further explore the effects of TLs and L1s to better understand parallel developmental patterns.

5.2. Parallel Developmental Patterns across TLs

To assess whether developmental patterns are parallel across the three languages, the CEs between consecutive years were compared across TLs, and a matrix was constructed indicating parallel (differences in Cohen's d below ± 0.4) or non-parallel (differences in Cohen's d above ± 0.4) patterns. The matrix in Figure 2 summarises parallel developmental patterns across the three possible consecutive-year pairs (Y7-Y8, Y8-Y9 and Y9-Y10). For each L1 profile group (BA L1 learners), the matrix shows whether the slopes are parallel between language pairs (BA-SP, BA-EN, SP-EN) or all three languages (3L). A total of 33 comparisons (3 consecutive year pairs x 11 items) are presented in the matrix.

Figure 2. Matrix of parallel developmental patterns ($\pm 0.4 \Delta d$) across TLs

		D1_ST			D1_R			D1_CR			D1_LA			D2_CI			D2_CC			D2_CS			D2_LC			D3_UC			D3_DT			D3_O						
L1	Years	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	P	NP	% Ps				
BA	BA-SP	P	P	P	P	P	NP	NP	P	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	26	7	78.79			
	BA-EN	P	P	NP	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	NP	NP	P	P	28	5	84.85		
	SP-EN	NP	NP	P	P	P	NP	P	P	P	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	27	6	81.82		
	3 L	NP	NP	NP	P	P	NP	NP	P	P	P	P	NP	NP	P	P	P	P	P	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	20	13	60.61	
BASP	BA-SP	P	P	P	P	NP	P	P	NP	NP	P	NP	P	P	NP	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	NP	NP	P	NP	NP	P	22	11	66.67
	BA-EN	P	P	NP	P	NP	P	P	P	P	NP	P	P	P	NP	P	P	NP	P	P	NP	NP	P	P	NP	P	P	P	P	P	NP	P	P	NP	P	23	10	69.70
	SP-EN	P	P	NP	P	P	P	P	NP	NP	P	P	P	P	P	P	NP	P	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	26	7	78.79	
	3 L	P	P	NP	P	NP	P	P	NP	NP	NP	NP	P	P	NP	P	NP	NP	NP	NP	P	P	NP	P	P	P	P	P	P	NP	NP	P	NP	NP	P	16	17	48.48
SP	BA-SP	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	30	3	90.91	
	BA-EN	NP	P	P	P	P	P	P	P	NP	P	P	P	P	NP	P	P	P	NP	P	NP	P	NP	P	P	P	P	P	P	P	P	P	P	P	26	7	78.79	
	SP-EN	P	P	P	P	P	P	P	P	P	P	NP	P	P	P	P	P	P	P	P	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	30	3	90.91	
	3 L	NP	NP	P	P	P	P	P	NP	NP	P	P	NP	P	P	NP	P	P	NP	P	NP	NP	NP	P	P	P	P	P	P	P	P	NP	P	P	22	11	66.67	
P		7	8	7	12	9	9	9	9	7	10	8	10	9	7	12	10	10	7	6	4	10	10	10	8	12	12	12	10	9	10	4	7	12				
NP		5	4	5	0	3	3	3	3	5	2	4	2	3	5	0	2	2	5	6	8	2	2	2	4	0	0	0	2	3	2	8	5	0				
% Ps		61.11			83.33			69.44			77.78			77.78			75.00			55.56			77.78			100.00			80.56			63.89						

In general, Basque L1 and Spanish L1 learners showed more parallel developmental patterns than BASP learners, especially when considering parallel development across the three languages or in between Basque and Spanish texts. As can be seen in the matrix, the development of D3_UC (*use of sources*) was parallel for all possible combinations. As expected, the patterns were more similar between Basque and Spanish texts for Basque L1 and Spanish L1 learners (78.79 % and 90.91 %, respectively), and also showed similarities between English and the other two languages (around 80% parallel patterns for BA learners, and 78.79% BA-EN and 90.91% SP-EN for SP learners). However, surprisingly, BASP students presented more parallel patterns between English and Spanish (79.79 %) than between the combinations including Basque (66-69 %).

Regarding the dimensions that presented the highest occurrence of parallel developmental patterns, all measures displayed total percentages above 50%, with specific items having higher values. This suggests that certain items (such as *use of sources*, *degree of transformation* or *language for arguing*) might be more likely to develop simultaneously across languages. Notably, the *use of sources* presented 100% similar patterns across all combinations. When examining specific comparisons, some noteworthy observations emerge. Except for *the use of original ideas*, and *stance-taking* to some extent, the majority of parallel patterns occurred between Y7 and Y8. This might indicate a critical period in the development of MDLs.

5.3. Parallel Developmental Patterns across L1s

Similarities in the developmental patterns across L1 profiles were also included in a similar matrix as the one described above (Figure 3). As observed in the matrix, instances of parallel development across L1s were present to varying degrees when examining TLs. Surprisingly, more similar patterns were found in English writing between BASP students and the other two (81%), but this was not the case between BA and SP learners (57%). When focusing on TL Spanish, parallelisms across L1s were found to range between 69% and 72%. The parallel patterns in Basque texts were the most irregular, with BASP learners displaying around 60% parallelism with BA or SP L1 learners.

When examining parallel development patterns across L1s per item, all items but *original ideas* showed values above 50%. In fact, *stance-taking*, *language for comparing* and *use of sources* showed high values of parallelism, especially salient in Y7-Y8 and Y8-Y9. In addition, although some items did not reach high percentages on the whole, some parallel patterns were found for specific year combinations. For example, *comparative items* in Y9 to Y10, *comparative criteria* in Y7-8 and Y8-Y9, or

degree of transformation in Y7-Y8.

Figure 3. Matrix of parallel developmental patterns ($\pm 0.4 \Delta d$) across L1 profiles

		D1_ST			D1_R			D1_CR			D1_LA			D2_CI			D2_CC			D2_CS			D2_LC			D3_UC			D3_DT			D3_O			P	NP	% Ps	
TL	Years	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10	Y7-Y8	Y8-Y9	Y9-Y10				
		BA	BA-BASP	P	P	P	P	NP	P	P	NP	P	NP	NP	P	NP	NP	P	P	P	P	NP	P	P	P	P	P	P	P	P	NP	NP	P	P	NP	NP	21	12
BA-SP	NP		P	P	P	P	NP	P	P	NP	P	NP	NP	P	P	P	P	NP	P	P	P	NP	P	P	NP	NP	P	P	P	NP	P	P	NP	21	12	63.64		
SP-BASP	P		P	P	P	NP	P	P	P	NP	NP	P	NP	NP	P	P	P	P	NP	NP	P	P	P	P	P	P	P	P	P	NP	NP	NP	NP	NP	20	13	60.61	
SP	BA-BASP	P	P	P	NP	P	P	NP	P	NP	NP	P	P	NP	P	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	NP	NP	23	10	69.70	
	BA-SP	P	NP	P	P	P	P	P	P	P	P	NP	NP	NP	P	P	P	P	P	P	P	P	P	NP	P	P	NP	P	P	NP	P	P	NP	24	9	72.73		
	SP-BASP	P	P	P	NP	P	P	NP	P	NP	P	P	P	NP	P	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	NP	NP	24	9	72.73	
EN	BA-BASP	P	P	P	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	NP	P	NP	P	P	P	P	P	P	P	P	NP	NP	NP	P	P	P	27	6	81.82
	BA-SP	P	P	P	P	P	NP	P	NP	NP	P	NP	NP	P	P	NP	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	NP	P	NP	NP	19	14	57.58	
	SP-BASP	P	P	P	P	P	P	NP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	NP	NP	NP	P	P	27	6
P		8	8	9	7	7	7	5	6	5	6	4	6	4	6	8	7	8	5	5	4	6	8	9	6	9	9	6	7	5	4	6	4	2				
NP		1	1	0	2	2	2	4	3	4	3	5	3	5	3	1	2	1	4	4	5	3	1	0	3	0	0	3	2	4	5	3	5	7				
% Ps		92.59			77.78			59.26			59.26			66.67			74.07			55.56			85.19			88.89			59.26			44.44						

6. Discussion

Our study has sought to understand the developmental patterns of MDLs across secondary education by considering students' three languages of the curriculum (Basque, Spanish and English) and varying L1 profiles (Basque, Spanish or both). To do that, by employing a rubric consisting of 11 items, students' argumentation skills, comparison skills and source-based writing were assessed.

RQ1 explored developmental patterns from Y7 to Y10, hypothesising that the development would be non-linear (Larsen-Freeman, 2009; Michel, 2017). The hypothesis was confirmed by the results from the present study, which showed that all 11 items increased significantly from the beginning of secondary education to the end. The findings are consistent with previous studies on general writing (Durrant et al., 2021; Shen et al., 2023; Staples et al., 2022; Usanova & Schnoor, 2024) and MDL/CDF development (del Pozo, 2024; Evnitskaya & Dalton-Puffer, 2024; Granados et al., 2021, 2023; Lorenzo, 2017; Whittaker & McCabe, 2023), which in general suggest that learners' writing abilities evolve over time. Nevertheless, as expected, the change was non-linear and gradual in some cases. In fact, post hoc pairwise tests (Supplementary material 1) showed significant differences were not always present in year-to-year comparisons, but rather in two-year cycle comparisons or when comparing Y7 and Y10. These findings hint at a gradual development of MDLs, as also emphasised by Granados et al. (2021, 2023). The non-linear nature was especially salient in BASP learners' development, who showed various decreases and increases, while BA and SP learners had a more linear performance. Furthermore, our study also supported the theoretically grounded multilayered and multidimensional nature of DLs (Nikula et al., 2024; Dalton-Puffer et al., 2024), as different items showed different developmental patterns, with some having steeper change, such as both items referring to the lexicogrammar of CDFs (*language for arguing* and *language for comparing*).

It was the focus of RQ2 to analyse parallelism of developmental patterns across students' languages. The findings are partially in line with our hypotheses, as developmental patterns common to the three languages were found indistinctly, with improvement and/or stagnation patterns often being parallel across the three languages (Granados et al., 2022; Usanova & Schnoor, 2024). The results are supported by multilingual writing theories positing a common set of multilingual writing skills (Cenoz & Gorter, 2011; Rinnert & Kobayashi, 2016) or transfer effects across languages (Arias-Hermoso & Imaz Agirre, 2024; Forbes, 2023; Martín-Laguna, 2020; Usanova & Schnoor, 2022, 2024). The findings from the present study also showed that certain items were more likely to show parallel development across languages (Egli Cuenat, 2017), for example, the ability to *use and comprehend information from the sources*, which displayed a 100% parallelism across all Year, TL and L1 combinations. In addition,

as also suggested by other trilingual studies in the BAC (Arias-Hermoso & Imaz Agirre, 2024; Orcasitas-Vicandi, 2021; Sagasta, 2003) and other bilingual communities with a minority language (e.g., Catalan and Spanish, Martín-Laguna, 2020), stronger parallel patterns were observed between the minority and majority languages, Basque and Spanish, for L1 Basque and L1 Spanish speakers. This is probably due to the higher levels of proficiency in said languages, representing their main language of instruction and/or dominant languages in society. However, BASP learners showed more frequent parallel patterns between English and either Basque or Spanish, than between Basque and Spanish. This finding is unexpected, given that BASP learners have both Basque and Spanish as their home languages, and it could be expected for them to show more closely aligned developmental patterns between these two languages, probably due to having more balanced cognitive abilities in both languages. All in all, our findings reveal that MDL development might be parallel across one's languages to a great extent.

Our final research question, RQ3, focused on the parallel developmental patterns across L1 profiles. In line with previous studies on writing development, both varying and parallel developmental patterns caused by students' L1 profiles were found (Alzahrani, 2024; Durrant et al., 2021; Shen et al., 2023; Staples et al., 2022). As a whole, texts in Basque presented approximately 60% of parallelism across L1 profiles, 70% in Spanish, and between 57-81% in English. Therefore, it can be said that some aspects of MDLs develop differently depending on individual factors such as L1. The higher overlap in English texts could be explained by similar learning approaches across students, as it could be argued that their exposure to English within school is similar, but their exposure to Basque and Spanish differs. If specific year combinations are examined, it could be observed that parallelism occurred at different points for different measures, with some showing stronger parallel patterns in the early years of secondary schooling (*reasoning or comparative criteria*, for example), while others might be more parallel in the final years (e.g., *stance-taking*). The findings support previous studies' claims that L1 background plays a role in language development, and that, even though some aspects might be shared and developed equally across L1 groups, some others might differ (Alzahrani, 2024; Durrant et al., 2021; Elosua & deBoeck, 2020; Elosua & Lasagabaster, 2025; Shen et al., 2023; Staples et al., 2022). An unexpected finding needs to be addressed—the stagnation of Spanish L1 learners after Y9 in all three languages. While the existence of plateaus in academic language development is well-established (Granados et al., 2021, 2023), it is surprising that only Spanish L1 learners displayed this stagnation consistently. Considering that schools share similar instructional approaches and that all L1 profiles follow the same curriculum, analysing other factors (e.g., out-of-school exposure, motivation, executive functions) could shed light on this pattern. A tentative hypothesis could be linked with either lack of motivation or cognitive overload derived from studying in a minority (L2) language. However, this hypothesis requires further research, possibly supported with a qualitative design such as teacher interviews or think aloud protocols.

The findings from this study have important theoretical and pedagogical implications. First, theoretically, the results contribute to the emerging field of MDL development by indicating the trajectories of trilingual writing by multilingual learners. In this regard, our study provides empirical evidence of theoretical constructs, such as Cummins' CUP (1980) or the multilayered and multidimensional conceptualisation of DLs (Dalton-Puffer et al., 2024; Nikula et al., 2024). Second, in a more pedagogical vein, our study complements that of Granados et al. (2022) by providing a functional description of DL development, which could be beneficial to understanding how these abilities are developed throughout secondary education, and therefore, designing pedagogical interventions or improving integrated assessment of content and language. The identification of challenging aspects of MDLs that might not develop alone might inform practitioners to design explicit interventions that foster students' DLs (Garro Larrañaga et al., 2024; Gerns, 2023a). In addition, our study highlights the importance of using and teaching students' whole linguistic repertoire in their schooling by implementing multilingual approaches to education, particularly in minority language contexts, which can benefit the development of the multilingual repertoire of students (Cenoz &

Gorter, 2017).

Despite the significant contributions of this research, some limitations need to be addressed. No inter-coder agreement analyses were carried out when using the rubric, as qualitative total agreement was reached in a first round of analyses. The present study adopted a cross-sectional approach to language development, and, although it can provide a picture of developmental patterns (Evnitskaya & Dalton-Puffer, 2024), the development of the same students could not be traced. In addition, task design might have influenced the possible outcomes, by integrating two CDFs in a single SBW task. Future studies should consider investigating CDFs in combination, but also separately, both in independent and SBW tasks. A further limitation of this study is the exclusion of students having a language other than Basque, Spanish or both as their L1s, which could have provided interesting findings in terms of language acquisition of heritage speakers. Future research should try to shed light on these aspects, as well as expand the use of the rubric in other language combinations, educational contexts (CLIL, monolingual teaching) or disciplines (history, mathematics) to achieve a deeper understanding of MDL development in a broader context.

7. Conclusion

In sum, the results of the present study have led us to conclude that there is strong evidence to support the development of MDLs and CDFs, which is currently a matter of special interest in the field (Ting et al., 2024). This study indicates that students' argumentation skills, comparison skills and their ability to write from sources develop across the years of secondary education. Therefore, it can be said that learners become increasingly subject-literate in all of their languages. These insights underscore the significance of considering the whole linguistic repertoire of students to foster their multilingual writing development, as both parallel and non-parallel developmental patterns might arise in the language acquisition and learning process.

REFERENCES

- Alzahrani, A. (2024). Utility of Kolmogorov complexity measures: Analysis of L2 groups and L1 backgrounds. *PLoS ONE*, 19(4), e0301806. <https://doi.org/10.1371/journal.pone.0301806>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Arias-Hermoso, R., Imaz Agirre, A., & Garro Larrañaga, E. (under review). Assessing academic and disciplinary literacies: rubric validation to measure argumentation, comparison and source-based writing skills.
- Arias-Hermoso, R., Imaz Agirre, A., & Garro Larrañaga, E. (2024). A comparison between input modalities and languages in source-based multilingual argumentative writing. *Assessing Writing*, 60, 100813. <https://doi.org/10.1016/j.asw.2024.100813>
- Arias-Hermoso, R., Lersundi, A., Imaz Agirre, A., & Garro Larrañaga, E. (2025). Bigarren Hezkuntzako ikasleen argudiaketa zientifikoa euskaraz, gaztelaniaz eta ingelesez idazten. [Translated title: Secondary Education students' scientific argumentation: writing in Basque, Spanish and English]. *Euskera*.
- Arias-Hermoso, R. & Imaz Agirre, A. (2024). Exploring multilingual writers in secondary education: insights from a trilingual corpus. *European Journal of Applied Linguistics*, 12(2), 282-307. <https://doi.org/10.1515/eujal-2023-0022>
- Breeze, R., & Dafouz, E. (2017). Constructing complex Cognitive Discourse Functions in higher education: An exploratory study of exam answers in Spanish- and English-medium instruction settings. *System*, 70, 81-91. <https://doi.org/10.1016/j.system.2017.09.024>
- Cenoz, J., & Gorter, D. (2011). Focus on multilingualism: A study of trilingual writing. *The Modern Language Journal*, 95(3), 356-369. <https://doi.org/10.1111/j.1540-4781.2011.01206.x>
- Cenoz, J. (2023). Plurilingual education in the Basque Autonomous Community. In J. M., Cots (Ed.). *Profiling plurilingual education: A pilot study of four Spanish autonomous communities* (pp. 33-53). Edicions de la Universitat de Lleida.
- Chen, M., & Zhou, H. (2022). Comparison-and-contrast in research articles of applied linguistics: A frame-based analysis. *Lingua*, 276, 103387. <https://doi.org/10.1016/j.lingua.2022.103387>
- Cheong, C. M., Zhu, X. & Xu, W. (2021). Source-Based Argumentation as a Form of Sustainable Academic Skill:

- An Exploratory study comparing Secondary School Students' L1 and L2 Writing. *Sustainability*, 132, 2869. <https://doi.org/10.3390/su132212869>
- Cheong, C. M., Zhu, X. & Xu, W. (2021). Source-Based Argumentation as a Form of Sustainable Academic Skill: An Exploratory study comparing Secondary School Students' L1 and L2 Writing. *Sustainability*, 132, 2869. <https://doi.org/10.3390/su132212869>
- Coyle, D., & Meyer, O. (2021). *Beyond CLIL: Pluriliteracies Teaching for Deeper Learning*. Cambridge. Cambridge University Press.
- Crammond, J. (1998). The uses and complexity of argument structures in expert and student persuasive writing. *Written Communication*, 15, 230–268. <https://doi.org/10.1177/0741088398015002004>
- Crossley, S. A., Tian, Y., & Wan, Q. (2022). Argumentation features and essay quality: Exploring relationships and incidence counts. *Journal of Writing Research*, 14 (1), 1–34.
- Cummins, J. (1980). The exit and entry fallacy in bilingual education. *NABE Journal*, 4(3), 25-60. <https://doi.org/10.1080/08855072.1980.10668382>
- Cummins, J. (2021). *Rethinking the education of multilingual learners: A critical analysis of theoretical concepts*. Multilingual Matters.
- Dalton-Puffer, C., Hüttner, J., & Nikula, T. (2024). The conceptualisation of disciplinary literacies in CLIL. In J. Hüttner & C. Dalton-Puffer (Eds.), *Building disciplinary literacies in content and language integrated learning* (pp. 1-25). Routledge.
- Dalton-Puffer, C. (2013). A construct of cognitive discourse functions for conceptualising content-language integration in CLIL and multilingual education. *European Journal of Applied Linguistics*, 1(2): 216-253. <https://doi.org/10.1515/eujal-2013-0011>
- Dalton-Puffer, C. (2016). Cognitive discourse functions: Specifying an integrative interdisciplinary construct. In T. Nikula, E. Dafouz, P. Moore, & U. Smit. (Eds.), *Conceptualising Integration in CLIL and Multilingual Education* (pp. 29–54). Multilingual Matters.
- del Pozo, E., & Llinares, A. (2021). Assessing students' learning of history content in Spanish CLIL programmes: a content and language integrated perspective, in C. Hemmi, & D. L. Banegas (Eds.), *International Perspectives on CLIL. International Perspectives on English Language Teaching*. Palgrave MacMillan. https://doi.org/10.1007/978-3-030-70095-9_3
- del Pozo, E., Llinares, A. (2021). Assessing Students' Learning of History Content in Spanish CLIL Programmes: A Content and Language Integrated Perspective. In: C. Hemmi & D.L. Banegas (Eds.) *International Perspectives on CLIL. International Perspectives on English Language Teaching*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-70095-9_3
- del Pozo, E. (2024). Assessment in CLIL: the pending subject in bilingual education? A case study. *Revista De Educación*, 1(403), 231–257. <https://doi.org/10.4438/1988-592X-RE-2024-403-605>
- Durrant, P., Brenchley, M., & McCallum, L. (2021). *Understanding Development and Proficiency in Writing: Quantitative Corpus Linguistic Approaches*. Cambridge University Press. <https://doi.org/10.1017/9781108770101>
- Egli Cuenat, M. (2017): Dreisprachige Textproduktion bei Sekundarschülerinnen und -schülern mit unterschiedlichen Lernbedingungen. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 22(1), 152-166.
- Elosua, P., & De Boeck, P. (2019). 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>
- Elosua, P., & Lasagabaster, D. (2025). An analysis of the impact of the L1 on bilinguals' writing competence. *Journal of Second Language Writing*, 67, 101182. <https://doi.org/10.1016/j.jslw.2025.101182>
- Evniktskaya, N., & Dalton-Puffer, C. (2024). CLIL learners' categorizations: Writing about history in English across three grade levels in Spanish bilingual schools. In J. Hüttner & C. Dalton-Puffer (Eds.), *Building disciplinary literacies in content and language integrated learning* (pp. 60-79). Routledge.
- Evniktskaya, N., & Dalton-Puffer, C. (2023). Cognitive discourse functions in CLIL classrooms: Eliciting and analysing students' oral categorizations in science and history. *International Journal of Bilingual Education and Bilingualism*, 26(3), 311–330. <https://doi.org/10.1080/13670050.2020.1804824>
- Forbes, K. (2021). *Cross-Linguistic Transfer of Writing Strategies: Interactions between Foreign Language and First Language Classrooms*. Multilingual Matters. <https://doi.org/10.21832/9781788929752>
- Garro Larrañaga, E., Antón, E., Arias-Hermoso, R., & Carmona, U. (2024). The effects of instruction on students' argumentative scientific writing in a Basque medium of instruction setting. *Journal of Immersion and Content-Based Language Education*. Online Version. <https://doi.org/10.1075/jicb.23014.gar>
- Gerns, P. (2023a). Building scientific knowledge in English: Integrating content, cognition and communication in

- secondary school CLIL biology. *Journal of Language and Education*, 9(3), 52–78. <https://doi.org/10.17323/jle.2023.17569>
- Gerns, P. (2023b). Qualitative insights and a first evaluation tool for teaching with cognitive discourse function: “Comparing” in the CLIL science classroom. *Porta Linguarum*, 40, 161–179. <https://doi.org/10.30827/portalin.vi40.26619>
- Granados, A., & Lorenzo, F. (2024). A functional description of disciplinary literacy in history: Applications of the Common European Framework of Reference for Languages to content and language integrated learning courses. In J. Hüttner & C. Dalton-Puffer (Eds.), *Building disciplinary literacies in content and language integrated learning* (pp. 83–100). Routledge.
- Granados, A., Lorenzo-Espejo, A., & Lorenzo, F. (2021). Evidence for the interdependence hypothesis: a longitudinal study of biliteracy development in a CLIL/bilingual setting. *International Journal of Bilingual Education and Bilingualism*, 25(8), 3005–3021. <https://doi.org/10.1080/13670050.2021.2001428>
- Granados, A., Lorenzo-Espejo, A., & Lorenzo, F. (2023). A portrait of academic literacy in mid-adolescence: a computational longitudinal account of cognitive academic language proficiency during secondary school. *Language and Education*, 37(3), 288–307. <https://doi.org/10.1080/09500782.2022.2079951>
- Gries, S. T. (2015). The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora*, 10(1), 95–125. <https://doi.org/10.3366/cor.2015.0068>
- Hüttner, J., & Llinares, A. (forthcoming). Bi- and Multilingual Disciplinary Literacies at School: a Challenge for 21st Century Citizens. In A. Imaz Agirre & R. Arias-Hermoso (Eds.), *Insights into Disciplinary Literacies: Multilingual Perspectives across Educational Levels*. Multilingual Matters.
- Hüttner, J., & Smit, U. (2024). Collaborative student writing in English-medium business studies: On the use of Cognitive Discourse Functions as an analytic tool. In M. Reisigl, J. Spitzmüller, F. Grosser, J. Hassemer, C. Lozo, & V. Ntouvli (Eds.), *Reden, Schreiben, Handeln: Festschrift für Helmut Gruber* (pp. 343–362). Wiener Linguistische Gazette.
- Jiménez-Aleixandre, M. P., & Erduran, S. (2007) Argumentation in science education: An overview. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education* (pp. 3–27). Springer.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579–589. <https://doi.org/10.1093/applin/amp043>
- Lersundi, A., Nikula, T., & Garro Larrañaga, E. (forthcoming). Scaffolding of Cognitive Discourse Functions through classroom interaction. In A. Imaz Agirre & R. Arias-Hermoso (Eds.), *Insights into Disciplinary Literacies: Multilingual Perspectives across Educational Levels*. Multilingual Matters.
- Li, D. (2022). A review of academic literacy research development: From 2002 to 2019. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), 1–22. <https://doi.org/10.1186/s40862-022-00130-z>
- Liu, F., & Stapleton, P. (2014). Counterargumentation and the cultivation of critical thinking in argumentative writing: Investigating washback from a high-stakes test. *System*, 45, 117–128. <https://doi.org/10.1016/j.system.2014.05.005>
- Llinares, A., & Morton, T. (2024). Assessing CLIL students' expression of Explore across languages and school disciplines: An interdisciplinary approach. In J. Hüttner & C. Dalton-Puffer (Eds.), *Building disciplinary literacies in content and language integrated learning* (pp. 101–117). Routledge.
- Llinares, A., & Nashaat-Sobhy, N. (2021). What is an ecosystem? Defining science in primary school CLIL contexts. *Language Teaching for Young Learners*, 3(2), 337–362.
- Llinares, A., & Nikula, T. (2023). CLIL students' production of cognitive discourse functions: Comparing Finnish and Spanish contexts. *Language and Education*, 38(3), 381–400. <https://doi.org/10.1080/09500782.2023.2211049>
- Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System*, 47, 64–72. <https://doi.org/10.1016/j.system.2014.09.016>
- Lorenzo, F. (2017). Historical literacy in bilingual settings: Cognitive academic language in CLIL history narratives. *Linguistics and Education*, 37, 32–41. <https://doi.org/10.1016/j.linged.2016.11.002>
- Lorenzo, F. & Trujillo, F. (2017). Languages of schooling in European policymaking: present state and future outcomes. *European Journal of Applied Linguistics*, 5(2), 177–197. <https://doi.org/10.1515/eujal-2017-0007>
- Martín-Laguna, S. (2020). *Tasks, pragmatics and multilingualism in the classroom: A portrait of adolescent writing in multiple languages*. Multilingual Matters.
- Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50–68). Routledge.
- Moje, E. B. (2015). Doing and teaching disciplinary literacy with adolescent learners: A social and cultural enterprise. *Harvard Educational Review*, 85(2), 254–278. <https://doi.org/10.17763/0017-8055.85.2.254>

- Nashaat-Sobhy, N., & Llinares, A. (2023). CLIL students' definitions of historical terms. *International Journal of Bilingual Education and Bilingualism*, 9, 1–14. <https://doi.org/10.1080/13670050.2020.1798868>
- Nikula, T., Dalton-Puffer, C., Llinares, A., & Lorenzo, F. (2016). More than content and language: The complexity of integration in CLIL and bilingual education. In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising Integration in CLIL and Multilingual Education* (pp. 1–25). Multilingual Matters. <https://doi.org/10.21832/9781783096145-004>
- Nikula, T.; Nashaat-Sobhy, N.; Minardi, S.; Gölle, T.; Yalçın, S.; Duman, S.K.; Bozbiyık, M.; Delibegović Džanić, N.; Ellison, M.; Gerns, P.; Gómez, E.; Hanušová, S.; Kováčiková, E.; Kääntä, L.; Lin, A.; Llinares, A.; Yi Lo, Y.; Lopriore, L.; Meyer, O.; Morton, T.; Neville, C.; Rannu, K.; Sağlamel, H.; Sula, G.; Sulkunen, S.; Pavičić Takač, V.; Tiermas, A.; Ting, T.; Tsagari, D.; Yüksel, G.; Žmavc, J. (2024): Towards an initial operationalisation of disciplinary literacies: A paper by CLILNetLE Working Group 1. V1. PHAIDRA repository (University of Vienna). <https://hdl.handle.net/11353/10.2050621>
- Orcasitas-Vicandi, M. (2021). Towards a multilingual approach in assessing writing: holistic, analytic and cross-linguistic perspectives. *International Journal of Bilingual Education and Bilingualism*, 25(6), 1-22. <https://doi.org/10.1080/13670050.2021.1894089>
- Pascual i Granel, V. (2024). *Tractament Integrat de Llengües i Continguts (TILC) en un context de minorització lingüística: De la teoria a la pràctica*. Octaedro Editorial.
- Plonsky, L., & Oswald, F. L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning*, 64(4), 878-912. <https://doi.org/10.1111/lang.12079>
- Polias, J. (2016). Apprenticing students into science: doing, talking & writing scientifically. Lexis Education.
- Pylonitis, C., & Meyer, O. (2024). Arguing for Global Citizenship: Mapping Deeper Learning in the Language-as-Discipline Classroom. In S. Greco & L. Cinganotto (Eds.), *Innovation in Education for Deeper Learning* (pp. 55–78). INDIRE-IUL Press.
- Read, J. (1990). Providing relevant content in an EAP writing test. *English for Specific Purposes*, 9(2), 109–121. [https://doi.org/10.1016/0889-4906\(90\)90002-T](https://doi.org/10.1016/0889-4906(90)90002-T)
- Rinnert, C. & Kobayashi, H. (2016). Multicompetence and multilingual writing. In R. M. Manchón & P. Matsuda (Ed.), *Handbook of Second and Foreign Language Writing* (pp. 365-386). De Gruyter Mouton. <https://doi.org/10.1515/9781614511335-020>
- Sagasta, M. P. (2003). Acquiring writing skills in a third language: The positive effects of bilingualism. *International Journal of Bilingualism*, 7(1), 27-42. <https://doi.org/10.1177/13670069030070010301>
- Schnoor, B. & Usanova, I. (2022). Multilingual writing development: Relationships between writing proficiencies in German, heritage language and English. *Reading and Writing*, 36(3), 599–623. <https://doi.org/10.1007/s11145-022-10276-4>
- Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40-59.
- Shen, C., Guo, J., Shi, P., Qu, S., & Tian, J. (2023). A corpus-based comparison of syntactic complexity in academic writing of L1 and L2 English students across years and disciplines. *PLoS ONE*, 18(10), e0292688. <https://doi.org/10.1371/journal.pone.0292688>
- Staples, S., Gray, B., Biber, D., & Egbert, J. (2023). Writing trajectories of grammatical complexity at the university: Comparing L1 and L2 English writers in BAWE. *Applied Linguistics*, 44(1), 46–71. <https://doi.org/10.1093/applin/amac047>
- Ting, T., Rieder-Marschallinger, S., & Dalton-Puffer, C. (2024). Overview of curricular demands of disciplinary literacies: An exploratory survey. A report by CLILNetLE Working Group 3 Deliverable 7. Retrieved from <https://phaidra.univie.ac.at/o:2082897>
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press.
- Tullis, J. G., & Goldstone, R. L. (2016). Comparison versus reminding. *Cognitive Research: Principles and Implications*, 1(20). <https://doi.org/10.1186/s41235-016-0020-8>
- Usanova, I., & Schnoor, B. (2024). Die Entwicklung von Schreibfähigkeit in mehreren Sprachen im Verlauf der Sekundarstufe – eine sprachpragmatische Sicht. *DDS – Die deutsche Schule*, 116(2), 11–31. <https://doi.org/10.31244/dds.2024.02.02>
- van Weijen, D., Rijlaarsdam, G. & van den Bergh, H. (2019) Source use and argumentation behavior in L1 and L2 writing: a within-writer comparison. *Reading and Writing*, 32, 1635–1655. <https://doi.org/10.1007/s11145-018-9842-9>
- Whittaker, R., & McCabe, A. (2023). Expressing evaluation across disciplines in primary and secondary CLIL writing: A longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 5(2), 1–18. <https://doi.org/10.1080/13670050.2020.1798869>
- Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school.

Language Teaching Research, 15(3), 343-362. <https://doi.org/10.1177/1362168811401154>
Zhu, X., Yao, Y., Liao, X., Xu, W., Zhao, P. & Jiang, Y. (2024). Exploring relationships between L1 and L2 writing strategy use and integrated writing performance: A cross-linguistic perspective. *International Review of Applied Linguistics in Language Teaching*. Online advance publication. <https://doi.org/10.1515/iral-2023-0330>

Funding

This work was supported by the Basque Government under Grants [IT1664-22 & PRE_2021_1_0001] and by the Spanish Ministry of Science, Innovation and Universities [PID2019-111655RA-I00].

Declaration of interest

The authors report there are no competing interests to declare.

Data availability statement

Due to the nature of the research and to ethical restrictions, supporting data is not available.

Appendix 1. Means across Years, L1s and TLs

Item	L1	L1_BA				L1_SP				L1_BASP			
		Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10
D1_ST	BA	1.895	2.023	2.179	2.563	1.512	1.913	2.026	2.228	1.923	1.800	1.944	2.391
	SP	2.132	2.273	2.308	2.500	1.963	1.967	2.377	2.439	2.154	2.100	2.222	2.522
	EN	2.053	1.884	2.179	2.219	1.915	1.891	2.250	2.351	2.154	2.150	2.222	2.174
D1_R	BA	1.658	1.977	2.077	2.500	1.293	1.707	2.039	1.982	1.385	1.450	1.944	2.261
	SP	1.868	2.159	2.231	2.219	1.707	1.967	2.247	2.263	1.846	1.850	2.056	2.304
	EN	2.053	1.884	2.179	2.219	1.512	1.826	2.184	2.123	1.692	1.950	2.000	2.043
D1_CR	BA	1.395	1.500	1.692	2.094	0.939	1.283	1.649	1.702	1.154	1.100	1.611	1.826
	SP	1.211	1.705	1.872	2.063	1.256	1.554	1.896	1.860	1.615	1.300	1.278	2.174
	EN	1.105	1.488	1.795	2.094	1.037	1.261	1.895	1.579	1.231	1.100	1.556	2.000
D1_LA	BA	1.289	1.773	2.026	2.625	0.902	1.489	2.078	2.070	1.231	1.150	1.778	2.304
	SP	1.263	1.727	1.974	2.281	1.220	1.707	2.078	2.000	1.308	1.450	1.778	2.130
	EN	1.026	1.279	1.590	2.156	0.827	1.152	1.868	1.596	0.923	1.150	1.500	1.913

Item	L1	L1_BA				L1_SP				L1_BASP			
		Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10
D2_CI	BA	1.421	1.773	1.974	2.281	1.207	1.641	1.818	1.895	1.385	1.400	2.111	2.174
	SP	1.211	1.909	1.897	2.031	1.963	1.967	2.377	2.439	1.692	1.900	2.111	2.261
	EN	1.368	1.767	1.821	2.063	1.915	1.891	2.250	2.351	1.385	1.650	1.889	2.130
D2_CC	BA	0.842	1.068	1.590	1.969	0.768	1.120	1.442	1.228	0.846	0.650	1.722	1.826
	SP	0.605	0.977	1.385	1.594	1.707	1.967	2.247	2.263	0.769	0.800	1.333	1.217
	EN	0.579	0.930	1.256	1.313	1.512	1.826	2.184	2.123	0.462	0.800	1.056	1.478
D2_CS	BA	0.553	0.750	1.308	1.406	0.354	0.772	1.052	1.263	0.385	0.500	1.111	1.478
	SP	0.711	1.432	1.641	2.063	1.256	1.554	1.896	1.860	1.231	1.250	1.611	1.826
	EN	1.105	1.349	1.615	1.969	1.037	1.261	1.895	1.579	0.692	1.050	1.611	1.826
D2_LC	BA	1.263	1.455	1.769	2.188	0.817	1.337	1.662	1.719	1.231	1.450	1.611	2.043
	SP	1.053	1.455	1.769	2.063	1.220	1.707	2.078	2.000	1.077	1.350	1.611	2.087
	EN	0.974	1.093	1.538	1.656	0.827	1.152	1.868	1.596	0.846	1.000	1.500	1.739

Item	L1	L1_BA				L1_SP				L1_BASP			
		Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10	Y7	Y8	Y9	Y10
D3_UC	BA	1.842	2.045	2.128	2.500	1.561	1.870	2.143	2.035	1.692	1.850	2.111	2.391
	SP	1.711	1.977	2.128	2.375	1.646	1.935	2.234	2.070	1.615	1.850	2.222	2.435
	EN	1.658	1.814	2.051	2.250	1.537	1.793	2.092	1.842	1.692	1.700	2.000	2.217
D3_DT	BA	0.711	0.977	1.256	1.875	0.720	0.913	1.312	1.211	0.769	0.650	1.222	1.696
	SP	0.763	1.227	1.538	1.938	0.720	1.228	1.662	1.544	0.846	1.150	1.389	1.913
	EN	0.632	0.860	1.333	1.625	0.476	0.728	1.368	1.123	0.692	0.800	0.889	1.478
D3_O	BA	0.237	0.545	0.615	0.813	0.341	0.587	0.688	0.596	0.615	0.150	0.611	1.091
	SP	0.342	0.295	0.615	0.781	0.524	0.457	0.688	0.509	0.462	0.500	0.500	1.000
	EN	0.342	0.302	0.282	0.719	0.220	0.250	0.539	0.474	0.462	0.100	0.278	0.739

Appendix 2. Line graphs for each item, divided by Year, TL and L1



