

English L2 connectives in academic bilingual discourse: A longitudinal computerised analysis of a learner corpus*

Conectores en el inglés como L2 del discurso académico bilingüe: Análisis computarizado longitudinal de un corpus de aprendices

Adrián Granados

UNIVERSIDAD PABLO DE OLAVIDE
ESPAÑA
agranav@upo.es

Francisco Lorenzo

UNIVERSIDAD PABLO DE OLAVIDE
ESPAÑA
fjlorber@upo.es

Received: 27-IX-2019 / **Accepted:** 30-XI-2020

DOI: 10.4067/S0718-09342021000200626

Abstract

This paper aims to describe the longitudinal evolution in the use of English L2 connectives made by students enrolled in a bilingual CLIL programme in the Andalusian secondary education system (Southern Spain) over three years of formal instruction. The automated tool Coh-Metrix has been used to approach a learner corpus produced by students as part of the school subject of bilingual history, which is taught in English as an L2. The overall evolution of connectives has been analysed, as well as the evolution of each of the connectives' categories measured by Coh-Metrix (causal, logical, adversative/contrastive, temporal, extended temporal and additive connectives). Results have then been interpreted in order to pinpoint the developmental stage of students' L2 written proficiency and analyse their degree of historical literacy. Over the three years of our study, the students have been found to increase their overall use of connectives in 15%, indicating that they are becoming more proficient L2 writers. Furthermore, there is a particular increase in their use of causal and adversative/contrastive connectives, and a decrease in extended temporal connectives, which points to the development of their historical literacy and their transition from narrative to expository texts.

Key Words: Connectives, English L2, CLIL, historical literacy, Coh-Metrix.

Resumen

Este artículo pretende describir la evolución longitudinal en el uso de conectores en inglés como segunda lengua (L2) por parte del alumnado del programa bilingüe (AICLE) de la educación secundaria de Andalucía (sur de España) a lo largo de tres cursos académicos. Se ha empleado la herramienta lingüística Coh-Metrix para procesar un corpus de aprendices producido por el alumnado como parte de la asignatura de historia bilingüe, que se imparte en inglés como L2. Se ha analizado la evolución global del uso de conectores, así como la evolución de cada una de las categorías de conectores analizadas por Coh-Metrix (causales, lógicos, adversativos/contrastivos, temporales, temporales extendidos y aditivos). Los resultados han sido interpretados para describir la fase de desarrollo de la competencia escrita en inglés como L2 del alumnado y analizar su nivel de literacidad histórica. A lo largo de los tres años de nuestro estudio, los sujetos han aumentado su uso global de conectores en un 15 %, lo que indica un aumento de su competencia escrita en inglés como L2. Además, se constata un particular aumento de los conectores causales y adversativos/contrastivos, y una disminución de los temporales extendidos, lo que apunta al desarrollo de su literacidad histórica y a la transición de textos narrativos a textos expositivos.

Key Words: Conectores, inglés como L2, AICLE, literacidad histórica, Coh-Metrix.

INTRODUCTION

The development of academic second language skills has been described from a number of angles (Wolfe-Quintero, Inagaki & Hae-Youn, 1998; Ortega & Byrnes, 2008; Christie, 2012, etc.). All language levels—including lexicon, syntax and pragmatics—evolve over the years as individuals face more cognitively demanding communicative situations. This advanced literacy process may finally result in individuals either struggling with language in non-colloquial circles or proving a firm and solid grasp of the tongue in academic settings. In a global multilingual world, the development of advanced skills may take place in more than one language. In these contexts, full literacy implies biliteracy and entails the gradual incorporation of advanced academic aspects into the multilingual repertoire of each individual. As a result of their training, individuals position themselves at a determinate point along the continuum between semilingualism and multilingualism, depending on their academic language skills. This means that the well-known distinction between BICS (basic interpersonal communicative skills) and CALP (cognitive academic language proficiency) needs to be accounted for in more than one language (see Bernstein, 1990, for a more classical restricted-elaborate code dyad).

However, the BICS/CALP dichotomy does not provide a full account of academic language. Instead, CALP arbitrates its own form and shape in each discipline. In other words, the knowledge structures of every academic discipline (mathematics, biology, history, etc.) display their specific language representations (Shanahan & Shanahan, 2008). Each field triggers particular language structures (not only lexicon) which mould a disciplinary language of its own. Equation, distribution and similarity are

mathematical notions. Revolution, agreement or corroboration (fact-checking) have special significations in history. Each discipline constitutes its own properties and produces peculiar disciplinary genres (see Biber, 1992; Lorenzo, 2013). Peculiarities can be traced down to levels below the sentence. In history, the abundance of causal relations is accompanied by particular grammar and syntax structures: ambiguous use of conjunctions, nominalisations, implicit causal and temporal organisation, cause-effect relations within clauses, etc. In fact, these recurrent linguistic representations have led historiographers to declare a 'language turn' in their discipline: history has been transformed into a science as a result of a 'literary artefact' (see White, 2010, for the language turn; Schleppegrell & Colombi 2002; Coffin, 2006, 2009; Nokes, 2013; Achugar & Carpenter, 2014, for a description of the language in history both as an L1 and an L2; and Llinares, Morton & Whittaker, 2012, for a precise description of languages in the subjects).

With all this, this paper intends to track changes in the composition of historical texts in a second language in bilingual programmes at a critical moment in language development, a three-year period in puberty when academic language should thrive as students engage in the reception and production of academic texts. Rather than the description of language per se, this research relates to a number of practical problems in bilingual education, namely, if the development of CALP is similar in monolingual settings and L2 contexts, and if potential language deficits hinder content learning, here historical content. This is an issue of utmost importance at a time when immersion programs (CLIL, content-based and suchlike) replace mainstream EFL teaching.

Of course, the evolution of historical academic discourse needs to be narrowed down to some particular language aspect. A crucial element in academic language is the use of connectives. Connectives are responsible for suprasentential cohesion. As any other part of speech, the elements which link sentences and clauses evolve as individuals mature. A contrafactual sentence with a conditional linker should not be expected from a primary schooler, whereas a continuous sequence of juxtapositions in advanced academic texts would be inappropriate on style and content-representation grounds (Hornberger, 2004). Furthermore, key historical notions (causes, time, manner, consequences, etc.) may or may not be verbalised with L2 linkers, and the distribution of connectives for the expression of intra-sentential relations may vary over time. Research questions will be formulated in this regard and will be applied to a longitudinal learner corpus gathered during a three-year timespan. In short, we intend to understand the interface between language and content, for which a description of the limits and development rates of historical literacy—regarding its use of connectives, in particular—will be undertaken.

1. Language development and the role of connectives

In the field of history (the school discipline in which this study is framed), students need to develop generalisation, abstraction, evaluation and argument in order to succeed (Christie & Maton, 2011; Lorenzo, 2017). The required historical literacy for these processes to be undertaken is said to occur during mid-adolescence: when adolescent learners are asked to compose historical narratives, they employ more morphological narrative complexities, more present and past participles and a higher concentration of nominals as they grow more proficient (Asención-Delaney & Collentine, 2011; Lorenzo, Granados & Ávila, 2019). This is part of the overall language development which has been described as a six-step process. First, speakers' lexis increases and diversifies. Then, syntax becomes more complex, with fewer errors and a higher incidence of dependent clauses and verb tense variation (especially present perfect and progressive tenses). After speakers familiarise with the new complexity of their grammar system, their attention is turned to lexis again, focusing this time on frozen expressions and learned chunks. This last level would represent the highest point in longitudinal development of language syntax (Biber, 1992; Spoelman & Verspoor, 2010; Gené-Gil, Juan-Garau & Salazar-Noguera, 2015).

It is generally acknowledged that connectives are a resource to create cohesive links between ideas and clauses and to provide clues about text organisation (Vande Kopple, 1985; Crismore, Markkanen & Steffensen, 1993; Longo, 1994; Sanders & Noordman, 2000; Cain & Nash, 2011; McNamara, Louwerse, Cai & Graesser, 2014. See the seminal work by Halliday & Hasan, 1976, for a detailed classification of connectives). Regarding their relation to language development, research shows that the use of local cohesive devices traces an inverse u-shaped curve as children grow older and become more proficient L1 writers. From second to eighth grade, children increase their use of connectives (King & Rentel, 1979; McCutchen & Perfetti, 1982; Akiguet & Piolat, 1996) in order to produce texts which are easier to comprehend and process (Just & Carpenter, 1980; Rayner & Pollatsek, 1994; in Crossley, Roscoe, McNamara & Graesser, 2011). According to Crossley et al. (2011: 303), they create less sophisticated texts with "more explicit cohesive devices to guide the reader". From this age on, the use of connectives would stabilise and then decrease (Freedman & Pringle, 1980; Haswell, 1986; in Crossley et al., 2011). This way, more skilled writers would create texts that are more sophisticated and less cohesive (Crossley et al., 2011). In fact, the use of local cohesive devices by mature L1 writers has been found to be negatively correlated to essay quality: the more they use, the lower mark they are awarded by expert evaluators (Evola, Mamer & Lentz, 1980; McCulley, 1985; Crossley & McNamara, 2010, 2011).

That being said, differences between L1 and L2 writers need to be acknowledged, as L1 and L2 writings follow different developmental rates (e.g. the evolution of verb tenses, syntax sophistication, sentence-level errors and cohesive devices differ from L1 to L2 development). The extent to which they are different is, however, still unknown. As Crossley, Kyle & McNamara (2016: 3) affirm:

“studies that examine the use of cohesive devices over time for L2 learners are less frequent and those available focus on the development of cohesive devices in speech, which may not overlap with developments in writing”.

Some longitudinal studies have been conducted in this regard (e.g. Crossley, Salsbury & McNamara, 2010a, 2010b; Crossley et al., 2016; all of them covering only one year and focusing on other types of cohesive devices), and they seem to find that there is an increase in the use of cohesive devices by L2 writers over time. Moreover, instead of having a negative impact on writing quality (as it was the case for the L1), some studies show that there is a positive correlation between the use of cohesive devices in an L2 and writing quality (Jafarpur, 1991; Yang & Sun, 2012, Crossley et al., 2016).

This difference may lay in the level of the writers' proficiency: beginning and intermediate L2 learners would still be expected to use many cohesive devices regardless of their age (and therefore a high incidence of these would be very positively assessed), whereas advanced and proficient writers would be expected to achieve more native-like levels of explicit cohesion and would therefore reduce their use of local cohesive devices. Then, the above-mentioned u-shaped curve would only culminate for L2 writers who achieve a very high level of proficiency. For the rest, the evolution might as well trace a diagonal straight line until becoming stable.

In order to shed some more light on the evolution of connectives, it would be very interesting to conduct a continuous observation of L2 writers' language production over a time span covering several years. Once the longitudinal evolution is described, a relation between connectives use and language proficiency will be evaluated.

2. Research Design and Methodology

2.1. Background

A series of studies conducted by the authors in the field of CLIL in Europe precede the present research (Lorenzo, Casal & Moore, 2010; Pérez, Lorenzo & Pavón, 2016; Lorenzo & Dalton Puffer, 2016; Lorenzo et al., 2019; Lorenzo, Granados & Rico, 2020). Some of these studies addressed issues like the description of incidental learning and positive transfer between L1 and L2, the benefits of CLIL over monolingual education, and the cooperation of linguist experts and content teachers for curriculum organisation and materials design. Nevertheless, as it is

generally acknowledged, the same learning mechanisms are activated in most academic bilingual environments due to the presence of the same linguistic and cognitive demands. For instance, Cenoz, Genesee and Gorter (2014) and Lorenzo (2017) agree that research in immersion, CLIL, EMI or content-based learning should monitor each other's progress, since mutual crossovers in research findings are feasible (Macaro, 2018). For all these reasons, this paper aims to obtain results on mid-adolescents' biliteracy which are applicable to all types of academic bilingual programs regardless of their denomination.

In the domain of Corpus Linguistics, recent empirical research using bilingual corpus produced in formal learning environments is mostly cross-sectional and decontextualised: the students' linguistic competence is only defined as it is at a given time and the language production under examination is not content-bound. However, content plays such an important role in language development that Cummins' distinction between 'basic interpersonal communicative skills' (BICS) and 'cognitive academic language proficiency' (CALP) (Cummins, 2008) is claimed not to be fully comprehensive (Harwood & Hadley, 2004; Dressen-Hammouda, 2008; Heine, 2014). For these authors, a third dimension should be added to the dichotomy: specialised academic register specific for each subject.

In this regard, the present paper focuses on historical literacy, which is crucial to the languages of schooling. For some authors, historical thinking follows a three-stage model (e.g. Vollmer, 2006; Beacco, 2010). The first stage would be 'recording', a purely narrative period. It would be followed by 'explaining', in which causes and consequences—including multifactorial causality—are explored. Finally, personal judgement together with the development of a personal stance would be achieved in 'arguing'. These stages are usually related to ages 11-13, 14-16 and 16-18, respectively.

2.2. Research questions

Within the research context described above, the present paper aims to address the following questions:

Question 1) Is there an evolution in the overall use of connectives made by the students over the three years of the study as measured by Coh-Metrix?

This would help to pinpoint the stage of evolution in which our students are framed regarding the inverse u-shaped curve that the use of local cohesive devices traces from beginner to proficient writers.

Question 2) Is there an evolution in the use of individual categories of connectives made by the students over the three years of the study as measured by Coh-Metrix?

Of all the general categories of connectives measured by Coh-Metrix—causal (e.g. *because, so*), logical (e.g. *and, or*), adversative/contrastive (e.g. *although, whereas*), temporal (e.g. *first, until*) and additive (e.g. *and, moreover*)—, there might be categories with no evolution or with an incidence score of 0. This would indicate that the connectives of said categories are not produced by individuals with an L2 level similar to the students in the study.

Question 3) Is the potential evolution in the use of individual categories of connectives made by the students constant over the three years of the study as measured by Coh-Metrix?

The constant increase or decrease in categories of connectives may indicate students' new abilities to produce certain language functions. Additionally, irregularities in the evolution of certain connectives will be examined, as they may highlight their sensitivity to texts' topics.

2.3. Sampling and participants

A public secondary school from southern Spain was contacted and selected on the grounds that it had already run an optional English-Spanish bilingual programme for more than fifteen years (in line with the Spanish and European trend towards CLIL multilingual education). The focus was then narrowed to a group of students from the same grade and class, so that many of the variables operating in learning settings (e.g. quantity and quality of language exposure, amount and nature of the units taught in an L2, teaching methodology, etc.) could be neutralised. Finally, a population of 20 students from ninth grade (ages 14-15) was selected, what ensured that these students would remain in school for at least two more years (16 years old marks the end of compulsory education in Spain). All these conditions provided the perfect environment for the execution of a longitudinal study.

Consent was granted by parents, teaching staff and school administration, and students were explained that they would participate in a bilingual study in connection with their English and history lessons. Since school teachers would supervise and coordinate the participation in the study, it was clarified that, regardless of the students' performance, their academic records would remain unaffected. As it has already been said, students were ninth-graders (ages 14-15) when the first sample was collected, which means that they had already received two years of training in the bilingual system. They had an English level of A2 according to the Common European Framework of Reference for Languages (CEFR). By the time the last sample was taken, students had become eleventh-graders (ages 16-17) and they were expected to reach a B1 level within the next academic year. However, the population of our research was forcibly reduced to 15 students, since two of them had to retake one year (ninth grade) and other three decided to abandon the bilingual programme in tenth grade.

Unlike cross-sectional research to date (e.g. the description of the evolutionary differences in academic writing between ninth-graders and tenth-graders), this study produces results which are not group-bound. It analyses the evolution of every student and accounts for particularities like individual writing style or attrition. In fact, even if a genuine longitudinal improvement of every individual is registered, 5 of the students abandoned the bilingual programme, be it for the extra cognitive demands, the workload or other academic reasons.

2.4. Instruments and data collection

Tests were designed in line with the history content that was being taught in class as imposed by the official curriculum. Over the three years of the research, the history teacher informed of the topic which was being taught in the pre-arranged test periods, which coincided respectively with 9/11 and The Clash of Civilisations, The Avant-Gardes, The Industrial Revolution, The American Revolution and The Spanish Civil War.

Tests were administered with no previous notice to the students, as the resulting texts should constitute an account of what students had learned and memorised during the explanations and activities in the history class. The duration of tests was 40 minutes. The authors and the English teacher were present during the tests, and no reference material or questions were allowed. The only directions consisted of the text's minimum length and the following guidelines on topic:

1. Define the given concept or historical period.
2. Explain its causes and consequences.
3. Give your opinion on its historical implications.

The longitudinal study was set to measure a timespan of three years, with two writing tests being performed each year. However, as opposed to the original plan, only one test was administered in the final year in order to avoid further dropout rates that would have seriously compromised the study: fatigue among students was observed and test 5 had already signalled evolution regarding tests 1-4.

The data collection process resulted in a learner corpus composed of 75 texts and amounting to 12,000 words. Narratives were organised in three periods of composition (P1, P2 and P3), corresponding to the three years of the study. The mean results of tests 1 and 2 were merged for P1; and so were the mean results of tests 3 and 4 for P2.

2.5. Analysis

The texts produced by the students were tagged by test and student and introduced in Coh-Metrix (version 3.0). Coh-Metrix is an automated web tool that generates

indices of the discourse and linguistic representations of texts within five major dimensions: “narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep (causal) cohesion” (McNamara et al., 2014). For the authors, computerisation therefore replaces other methods used in the past to measure L1/L2 language development, like linguistic analysis provided by hand and other traditional classifications (e.g. Hunt analysis). Other automated tools like Synlex, which has been previously used for the cross-sectional description of bilingual discourse (Lorenzo, 2017; Lu, 2010, for further information on this tool) and other types of software like Trijamod, Childes and Freeling were considered, but they were found less appropriate to the ends of this study.

Many validation studies support Coh-Metrix’s reliability. For instance, Polio and Yoon (2018) compared hand-coding with Coh-Metrix’s results and determined that the automated-tool was a non-redundant and rather transparent method to measure text complexity, cohesion and coherence metrics. In addition, they found it was able to reflect genre differences among ESL writers. Parallels in results and metrics were found with other automated tools previously used to measure the lexical complexity of history texts (Lorenzo, 2017).

Even if second language acquisition research usually employs ethnography in longitudinal case studies (Duff, 2008; Yin, 2014), the results provided by Coh-Metrix were all statistically processed. This way, the reader-evaluator is replaced as the main source of text quality assessment, and statistical analysis becomes the main source instead (Crossley & McNamara, 2012). A previous paper by Lorenzo et al. (2019) examined the four Coh-Metrix’s dimensions with statistical significance (fluency metrics, syntactic complexity, syntactic pattern density and text easability). In this case, the means of every category of connectives in each of the sampling periods will be subject to a descriptive analysis in which significance levels are not targeted.

3. Results

Drawing on Halliday and Hasan’s (1976) and Louwse’s (2000, 2001) taxonomies, Coh-Metrix measures the incidence score (occurrence per 1000 words) of all connectives together and of five general categories of connectives individually (McNamara et al., 2014): causal, logical, adversative/contrastive, temporal, and additive. Additionally, it provides information about extended temporal connectives. The meaning and evolution of each category will be addressed in turn, and an excerpt from our learner corpus (with the original spelling and grammar, as produced by the students) will be provided as an example of usage.

Table 1. Evolution of Coh-Metrix's indices for connectives.

Coh-Metrix's Indices			Mean scores		
No.	Acronym	Name	P1	P2	P3
52	CNCAll	'All connectives incidence'	86.842	94.869	101.311
53	CNCCaus	'Causal connectives incidence'	28.351	34.313	35.600
54	CNCLogic	'Logical connectives incidence'	37.170	35.580	39.150
55	CNCADC	'Adversative and contrastive connectives incidence'	9.670	9.902	13.564
56	CNCTemp	'Temporal connectives incidence'	10.590	9.408	17.126
57	CNCTempx	'Extended temporal connectives incidence'	13.301	16.061	8.296
58	CNCAdd	'Additive connectives incidence'	48.877	55.092	50.915

All connectives: CNCAll (Coh-Metrix's index 52)

This is the incidence score (the occurrence per 1000 words) of all connectives. As it can be seen in Table 1, the overall use of connectives evolves in a constant manner, increasing in almost 15‰ (15 more connectives per 1000 words used) from period 1 to period 3.

(1) As a consequence of the battles many people died, **but** many others were killed by their neighbours, who acused each other **as a result of** past disagreements. **In addition**, republicans were shoted **when** a new town was occupied.

Causal Connectives: CNCCaus (Coh-Metrix's index 53)

This is the incidence score of causal connectives (e.g. 'because', 'so'). Causal connectives also evolve in a constant manner, increasing in more than 7‰ from period 1 to period 3.

(2) Therefore, it quickly spreaded across the country, having under control a third of the country after a few days.

Logical Connectives: CNCLogic (Coh-Metrix's index 54)

This is the incidence score of logic connectives (e.g. 'and', 'or'). They experience a slight decrease (-1.59‰) from P1 to P2, but an overall increase of 1.98‰ from P1 to P3.

(3) Many of countries like Germany **or** Italy helped **and** supported with army and weapons to Spain, **and** countries like USA, Russia **or** France were against that.

Adversative/Contrastive connectives: CNCADC (Coh-Metrix's index 55)

This is the incidence score of adversative/contrastive connectives (e.g. 'although', 'whereas'). They remain almost stable from P1 to P2 (+0.23‰) and then experience an increase of 3.662‰ from P2 to P3.

(4) **On the other hand**, the republican army did not get the support of other democratic countries **but** did get it from the Soviet Union, due to the numerous communists politicians that conformed the government.

Temporal Connectives: CNCTemp (Coh-Metrix's index 56)

This is the incidence score of temporal connectives (e.g. 'first', 'until'). Similarly to logical connectives, they experience a small decrease from P1 to P2 (-1.182‰) and an overall increase of 6.536‰ from P1 to P3. This parallelism will be addressed in the discussion.

(5) For example, the cubism tried to represent an object in movement so to do that the artist turned the object into different geometric shaped objects and **then**, he represented the different surfaces.

Extended Temporal Connectives: CNCTempx (Coh-Metrix's index 57)

This is the incidence score of extended temporal connectives (e.g. 'at this time', 'sooner or later'). Extended temporal connectives are the only category decreasing in the study, with an overall evolution of -5.005‰.

(6) Therefore, it quickly spreaded across the country, having under control a third of the country **after a few days**.

Additive Connectives: CNCAdd (Coh-Metrix's index 58)

This is the incidence score of additive connectives (e.g. 'and', 'moreover'). Contrary to logical and temporal connectives, they experience a positive evolution from P1 to P2 (+6.215) and then a negative evolution from P2 to P3 (-4.177). Nevertheless, they experience an overall increase of 2.038 from P1 to P3. The interrelation between logical, temporal and additive connectives will be addressed in the discussion.

(7) **In addition**, two industrial revolutions have happened and **apart from** the changes in inventions, the society changed **too**.

4. Discussion

The answer to the first research question seems clear: the students in our study increase their use of connectives in almost 15‰ (15 more connectives per 1000 words used) from period 1 to period 3. These results coincide with those from the studies performed by Crossley et al. (2010a, 2010b) and Yang & Sun (2012), indicating that L2 writers create less sophisticated texts with "more explicit cohesive devices to guide the reader" over time (Crossley et al., 2011: 303).

To fully understand the implications of this evolution in terms of proficiency, a distinction should be made between L1 and L2 development. In L2 development, many studies have shown that there is a positive correlation between the use of

cohesive devices in an L2 and writing quality (Jafarpur, 1991; Yang & Sun, 2012; Crossley et al., 2016). Therefore, it could be affirmed that the students in our study are becoming more proficient L2 writers.

If this evolution is compared with that in L1 development, the stage of evolution in which our students are framed can be pinpointed regarding the inverse u-shaped curve that the use of local cohesive devices traces from beginner to proficient writers in an L1. Research shows that children increase their use of connectives from second to eighth grade (King & Rentel, 1979; McCutchen & Perfetti, 1982; Akiguet & Piolat, 1996), i.e. from 7-8 years old to 13-14 years old. From this age on, the use of connectives has been found to stabilise and then decrease (Freedman & Pringle, 1980; Haswell, 1986; in Crossley et al., 2011). This means that the students in our study were just about reaching the top of the connectives curve in their L1 when the study commenced (they were 14-15 years old). Therefore, it comes as no surprise that their L2 connectives development falls certain years behind their L1 development and is still framed in the ascending phase.

The second research question enquired about the manifestation of all the connectives measured by Coh-Metrix in the students' narratives, and the data show that all the categories of connectives are present and signal evolution. Finally, the third research question addressed the constant or irregular evolution of these connectives and its implications. In this regard, results show a constant increase in both causal connectives (+7‰) and adversative/contrastive connectives (+3.662‰). An increase in causal connectives may prove that, apart from becoming more proficient L2 writers, our students are also developing historical literacy, as this discipline is characterised by a pronounced incidence of causality (Achugar & Schleppegrell, 2005). It would also indicate that students have reached the second stage of historical thinking: *explaining*, an exploration of the causes and consequences of historical events, including multifactorial causality (Vollmer, 2006; Beacco, 2010). The increase in adversative/contrastive connectives would in turn point to the onset of a shift in style, in which students develop a less dogmatic description of history. As put forth by Wagner (2011), the end point of the process of grammaticalisation is concessives, a complex form of non-entailment which contradicts ideas conveyed previously and, accordingly, demands cognitive sophistication. Narratives seem thus to include multiple sides of a same story, in which arguments are more nuanced and counterarguments are refuted.

Results also register the constant decrease in extended temporal connectives (-5.005‰). This implies the decrease in the narrativity of texts, as "a narrative text tells a story, with characters, events, places, and things that are familiar to the reader" (McNamara et al., 2014). The evolution of this index therefore points to the transition from narrative to expository texts, in which historical discourse is not constructed

from personal accounts, but from multisided expositions (Martin & White, 2005). This would also reinforce the perception of students' maturity in the rendition of academic historical information.

Finally, some irregularities have been found in the evolution of three types of connectives: logical, temporal and additive. There seems to be a correlation in the evolution of logical and temporal connectives, which decrease slightly from P1 to P2 (-1.59% and -1.182% , respectively) but experience an overall increase from P1 to P3 ($+1.98\%$ and $+6.536\%$). Furthermore, they may be interrelated with additive connectives, which increase from P1 to P2 ($+6.215$) and decrease from P2 to P3 (-4.177). All these fluctuations might be due to the special sensitivity of these connectives to text topics. However, their potential interrelation is still unknown and further analysis would be required to venture any explanation.

In addition to the sensitivity to text topics, another limitation of our study is the sample size. Keeping the teaching variables constant during the 3 years of the study (e.g. the same instruction style, language exposure or assessment) was a priority and it came at the price of a limited number of subjects. Moreover, attrition—a permanent feature of bilingual education (Nikula, Dalton-Puffer, Llinares & Lorenzo, 2016)—further reduced our sample. All this prevented us from targeting significance levels in this descriptive analysis. The findings discussed here should pave the way for future large-scale, cross-sectional analyses aimed at testing their generalizability. Nevertheless, results coincide with and complement those obtained in Lorenzo et al. (2019) and Lorenzo et al. (2020), which did target significance levels, the latter having a sample of over 3,800 students.

CONCLUSION

Over the three years of our study, the students have been found to increase their overall use of connectives in 15%, confirming that they are becoming more proficient L2 writers. The particular increase in their use of causal and adversative/contrastive connectives points to the development of their historical literacy, including a more nuanced conception of historical processes. The increase in their use of extended temporal connectives also signals the transition from narrative to expository texts.

These advances are supported by development outside the domain of connectives. Lorenzo et al. (2019) explored different linguistic domains from the same learner corpus and found a significant increase of nominalisation, abstraction and syntactic complexity, together with a significant decrease in text easability.

Seeing the bigger picture, this evolution is not only important in academic terms, i.e. in students' transition towards CALP (Cummins, 2008), but also in cognitive terms, i.e. in their transition towards the elaborated code (Bernstein, 1990). For students (and citizens) to be able to explore the causes and consequences of events, to

see the two sides of the story, and to judge and take an ideological stance on the past and the present, they need to master some indispensable language structures (Martin & White, 2005). Without them, their language capital (Bourdieu, 1992; Blommaert, 2010) and their freedom is inevitably constrained.

REFERENCES

- Achugar, M. & Carpenter, B. (2014). Tracking movement toward academic language in multilingual classrooms. *Journal of English for Academic Purposes*, 14, 60-71.
- Achugar, M. & Schleppegrell, M. J. (2005). Beyond connectors: The construction of cause in history textbooks. *Linguistics and Education*, 16(3), 298-318.
- Akiguet, S. & Piolat, A. (1996). Insertion of connectives by 9- to 11-year-old children in an argumentative text. *Argumentation*, 10, 253-270.
- Asención-Delaney, Y. & Collentine, Y. (2011). A multidimensional analysis of a written L2 Spanish corpus. *Applied Linguistics*, 32, 299-322.
- Beacco, J. D. (2010). *Items for a description of linguistic competence in the language of schooling necessary for teaching/ learning history (end of obligatory education)* [on line]. Retrieved from:
http://www.coe.int/t/dg4/linguistic/Source/Source2010_ForumGeneva/1_LIS-History2010_en.pdf
- Bernstein, B. (1990). *The structuring of pedagogic discourse Volume IV: Class, codes and control*. London: Routledge.
- Biber, D. (1992). The multi-dimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. *Computers and the Humanities*, 26(5/6), 331-345.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.
- Bourdieu, P. (1992). *Language and symbolic power*. Cambridge: Polity Press.
- Cain, K. & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, 103(2), 429-441.
- Cenoz, J., Genesee, F. & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics*, 35, 243-262.
- Christie, F. (2012). *Language education: A functional perspective. Language learning monograph series*. Chichester: Wiley-Blackwell.

- Christie, F. & Maton, K. (2011). *Disciplinary: Functional linguistic and sociological perspectives*. London: Continuum.
- Coffin, C. (2006). 'Reconstructing personal time as collective time: Learning the discourse of history.' In R. Whittaker, M. O'Donnell & A. McCabe (Eds.), *Language and literacy: Functional approaches* (pp. 15-45). London: Continuum.
- Coffin, C. (2009). *Historical discourse*. London: Continuum.
- Crismore, A., Markkanen, R. & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication*, 10, 39-71.
- Crossley, S. A. & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 984-989). Austin, TX: Cognitive Science Society.
- Crossley, S. A. & McNamara, D. S. (2011). Text coherence and judgments of essay quality: Models of quality and coherence. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 29th annual conference of the cognitive science society* (pp. 1236-1241). Austin, TX: Cognitive Science Society.
- Crossley, S. A. & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35, 115-135.
- Crossley, S. A., Kyle, K. & McNamara, D.S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1-16.
- Crossley, S. A., Roscoe, R. D., McNamara, D. S. & Graesser, A. (2011). Predicting human scores of essay quality using computational indices of linguistic and textual features. In G. Biswas, S. Bull, J. Kay & A. Mitrovic (Eds.), *Proceedings of the 15th international conference on artificial intelligence in education* (pp. 438-440). New York, NY: Springer.
- Crossley, S. A., Salsbury, T. & McNamara, D. S. (2010a). The development of polysemy and frequency use in English second language speakers. *Language Learning*, 60, 573-605.
- Crossley, S. A., Salsbury, T. & McNamara, D. S. (2010b). The development of semantic relations in second language speakers: A case for Latent Semantic Analysis. *Vigo International Journal of Applied Linguistics*, 7, 55-74.

- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopaedia of language and education: Vol. 2. Literacy* (pp. 71-83). New York: Springer.
- Dressen-Hammouda, D. (2008). From novice to disciplinary expert: Disciplinary identity and genre mastery. *English for Specific Purposes*, 27(2), 233-252.
- Duff, P. (2008). *Case study research in applied linguistics*. New York/London: Lawrence Erlbaum.
- Evola, J., Mamer, E. & Lentz, B. (1980). Discrete point versus global scoring for cohesive devices. In J. Oller Jr. & K. Perkins (Eds.), *Research in language testing* (pp. 177-181). Rowley, MA: Newbury House.
- Freedman, A. & Pringle, I. (1980). Writing in the college years: Some indices of growth. *College Composition and Communication*, 31, 311-324.
- Gené-Gil, M., Juan-Garau, M. & Salazar-Noguera J. (2015). Development of EFL writing over three years in secondary education: CLIL and non-CLIL settings. *Language Learning Journal*, 43, 286-303.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Harwood, N. & Hadley, G. (2004). Demystifying institutional practices: Critical pragmatism and the teaching of academic writing. *English for Specific Purposes*, 23(4), 355-377.
- Haswell, R. H. (1986). *Change in undergraduate and post-graduate writing performance: Quantified findings* (Technical Report). Retrieved from: ERIC database. (ED269780)
- Heine, L. (2014). Models of the bilingual lexicon and their theoretical implications for CLIL. *The Language Learning Journal*, 42(2), 225-237.
- Hornberger, N. (2004). The Continua of biliteracy and the bilingual educator: Educational linguistics in practice. *International Journal of Bilingual Education and Bilingualism*, 7(2), 155-171.
- Jafarpur, A. (1991). Cohesiveness as a basis for evaluating compositions. *System*, 19, 459-465.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- King, M. & Rentel, V. (1979). Toward a theory of early writing development. *Research in the Teaching of English*, 13, 243-253.

- Llinares, A., Morton, T. & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge: Cambridge University Press.
- Longo, B. (1994). Current research in technical communication: The role of metadiscourse in persuasion. *Technical Communication*, 41, 348-352.
- Lorenzo, F. (2013). Genre-based curricula: Multilingual academic literacy in content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*, 16(3), 375-388.
- Lorenzo, F. (2017). Historical Literacy in bilingual settings: Cognitive academic language in L2 History Narratives. *Linguistics and Education*, 37, 32-41.
- Lorenzo, F. & Dalton-Puffer, C. (2016). Historical literacy in CLIL: Telling the past in a second language. In T. Nikula, E. Dafouz, P. Moore & U. Smit (Eds.), *Conceptualising Integration in CLIL and Multilingual Education* (pp. 55-72). Bristol: Multilingual Matters.
- Lorenzo, F., Casal, S. & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied Linguistics*, 31(3), 418-442.
- Lorenzo, F., Granados, A. & Ávila, I. (2019). The development of cognitive academic language proficiency in multilingual education: Evidence of a longitudinal study on the language of history. *Journal of English for Academic Purposes*, 41, 100767.
- Lorenzo, F., Granados, A. & Rico, N. (2020). Equity in bilingual education: Socioeconomic status and content and language integrated learning in monolingual Southern Europe. *Applied Linguistics*, aaaa037.
- Louwse, M. (2000). *From cohesion in text to coherence in comprehension*. Ph.D. Dissertation, University of Edinburgh, Edinburgh, United Kingdom.
- Louwse, M. (2001). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics*, 12, 291-315.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.
- Macaro, E. (2018). *English medium instruction: Content and language in policy and practice*. Oxford: Oxford University Press.
- Martin, J. R. & White, P. R. R. (2005). *The language of evaluation. Appraisal in English*. New York: Palgrave Macmillan.

- McCulley, G. A. (1985). Writing quality, coherence, and cohesion. *Research in the Teaching of English*, 19, 269-282.
- McCutchen, D. & Perfetti, C. A. (1982). Coherence and connectedness in the development of discourse production. *Text-Interdisciplinary Journal for the Study of Discourse*, 2(1-3), 113-140.
- McNamara, D. S., Louwse, M., Cai, Z. & Graesser, A. (2014). *Automated evaluation of text and discourse with Cob-Matrix*. Cambridge: Cambridge University Press.
- 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*. Bristol: Multilingual Matters, 1-26.
- Nokes, J. D. (2013). *Building students' historical literacies: Learning to read and reason with historical texts and evidence*. New York: Routledge.
- Ortega, L. & Byrnes, H. (Eds.). (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.
- Pérez, A., Lorenzo, F. & Pavón, V. (2016). European bilingual models beyond *linguafranca*. Key findings from CLIL French programs. *Language Policy*, 15(4), 485-504.
- Polio, C. & Yoon, H. J. (2018). The reliability and validity of automated tools for examining variation in syntactic complexity across genres. *International Journal of Applied Linguistics*, 28(1), 165-188.
- Rayner, K. & Pollatsek, A. (1994). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
- Sanders, T. J. M. & Noordman L. G. M. (2000). The role of coherence relations and their linguistic markers in processing. *Discourse Processes*, 29(1), 37-60.
- Schleppegrell, M. J. & Colombi, M. C. (2002). *Developing advanced literacy in first and second languages*. Mahwah: Lawrence Erlbaum.
- Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents. Rethinking content-area literacy. *Harvard Educational Review*, 78, 40-59.
- Spoelman, M. & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31(4), 532-553.
- Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 36, 82-93.

- Vollmer, H. J. (2006). *Language across the curriculum*. Brussels: Language Policy Division, Council of Europe.
- Wagner, S. (2011). Concessives and contrastives in student writing: L1, L2 and genre differences. In J. Schmied (Ed.), *Academic Writing in Europe: Empirical Perspectives* (pp. 23-49). Göttingen: Cuvillier.
- White, H. (2010). *The fiction of narrative: Essays on History, Literature, and Theory, 1957-2007*. Baltimore: The Johns Hopkins University Press.
- Wolfe-Quintero, K., Inagaki, S. & Hae-Youn, K. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawaii Press.
- Yang, W. & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23, 31-48.
- Yin, R. K. (2014). *Case study research design and methods (5th ed.)*. Thousand Oaks, CA: Sage.

* **ACKNOWLEDGEMENTS**

This research was funded by the Spanish Ministry of Economy & Competitiveness (research grant ref. FFI2016-74885-P EU ERDF) and by the Spanish Ministry of Science (research grant ref. FPU18/02673).