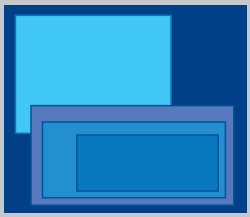


VIAL

VIGO INTERNATIONAL JOURNAL of APPLIED LINGUISTICS

Number 20-2023

Universidade Vigo

Servizo
de Publicacións

VIAL

Vigo International Journal
of Applied Linguistics

Universidade Vigo

VIAL. Vigo International Journal of Applied Linguistics.

VIAL is an international and peer-reviewed journal which is indexed and abstracted in CINDOC (Centro de Información y Documentación Científica), EBSCO, ERIH PLUS, JCR, Web of Knowledge (Arts and Humanities Citation Index, Social Sciences Citation Index), Linguistics Abstracts, Linguistics and Language Behavior Abstracts, Latindex, MLA Directory of Periodicals, PsychoInfo, Scimago Journal Rank Scopus, Ulrich's Database

Webpage: www.webs.uvigo.es/vialjournal/

Printed in Spain - Impreso en España
I.S.S.N.: 1697-0381 - Electronic ISSN: 2660-504X - Depósito Legal: VG-935-2003
Imprime e maqueta: Tórculo Comunicación Gráfica, S.A. Deseño da portada: Elena Gómez.

VIAL

Vigo International Journal of Applied Linguistics

Number 20 - 2023

Contents:

An analysis of language anxiety in English and Basque-Medium Instruction: A study with primary school students
Irene Alonso Arruabarrena and David Lasagabaster Herrarte 9

Collaborative writing and patterns of interaction in young learners: The interplay between pair dynamics and pairing method in LRE production
María Basterrechea and Francisco Gallardo-del-Puerto 49

Metaphors set in motion in the context of L2 academic spoken discourse
Irene Castellano-Risco, Marta Martín-Gilete, Alberto Hijazo-Gascón and Iraide Ibarretxe-Antuñano 77

Computational analysis of adjuncts in ASD-STE100 for the NLP parser ARTEMIS
Francisco José Cortés-Rodríguez and Carolina Rodríguez-Juárez 107

How complex is professional academic writing?
A corpus-based analysis of research articles in 'hard' and 'soft' disciplines
Javier Pérez-Guerra and Elizaveta A. Smirnova 149

Polarization in the Spanish press: A study of political sectarianism during the COVID-19 pandemic
Roberto A. Valdeón 185

An analysis of language anxiety in English and Basque-Medium Instruction: A study with primary school students

Irene Alonso Arruabarrena
Department of English, German and Translation and Interpretation Studies
University of the Basque Country UPV/EHU
irene.alonsoa@ehu.eus

David Lasagabaster Herrarte
Department of English, German and Translation and Interpretation Studies
University of the Basque Country UPV/EHU
david.lasagabaster@ehu.eus

Abstract

In an era of globalisation, acquiring communicative competence in foreign languages has become an educational priority, and in this scenario, Content and Language Integrated Learning (CLIL) emerged as a European endeavour to improve students' foreign language proficiency. In the Basque Autonomous Community (BAC), learners are required to deal with the coexistence of three languages in the school curriculum, which may be regarded as an anxiety-provoking process. Through a mixed method approach, this study explores the interaction between language anxiety and the language of instruction in the case of 89 primary education students in the BAC, along with the effects of students' gender, linguistic repertoire and achievement on anxiety means. Results indicated that participants experienced significantly more anxiety in their English subjects than in their Basque subjects. Moreover, students speaking Spanish at home or scoring lower grades in both instruction types exhibited higher anxiety levels, and this trend diminished in the case of females. These findings reveal the challenging nature of anxiety, an affective factor that should be taken into consideration in the implementation of language teaching methodologies.

Keywords: CLIL, mixed method approach, language anxiety, language of instruction, primary education.

Resumen

En plena era de globalización, adquirir competencias comunicativas en lenguas extranjeras se ha convertido en una prioridad educativa y, en este escenario surgió el Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE) como un

intento por mejorar el dominio de la lengua extranjera de los estudiantes europeos. En la Comunidad Autónoma Vasca (CAV), el alumnado debe afrontar la coexistencia de tres lenguas en el currículo escolar, lo cual puede resultar en un proceso que genere ansiedad. Mediante un enfoque metodológico mixto, este estudio explora la interacción entre la ansiedad y la lengua de instrucción en el caso de 89 estudiantes de educación primaria en la CAV, junto con los efectos del género, el repertorio lingüístico y los resultados académicos de los estudiantes en sus niveles de ansiedad. Los resultados indicaron que los participantes experimentaron significativamente más ansiedad en sus asignaturas impartidas en inglés que en las de euskera. Además, el alumnado que hablaba español en casa o que obtenía calificaciones más bajas en ambos tipos de instrucción exhibía niveles más altos de ansiedad, aunque esta tendencia disminuyó en el caso de las mujeres. Estos datos revelan la naturaleza desafiante de la ansiedad, un factor afectivo que debe considerarse en la implementación de metodologías de enseñanza de idiomas.

Palabras clave: AICLE, enfoque metodológico mixto, ansiedad, lengua de instrucción, educación primaria.

1. Introduction

Considering the complexities and the globalisation-related trends of the multilingual society of today, the importance of learning foreign languages (FL) and reaching good standards has been recognised across the globe (European Commission, 2012). In fact, data gathered by the European Commission (2012) indicates that there is a limited number of Europeans who are able to communicate in a language other than their mother tongue. In view of such results and the linguistic consequences of globalisation, developing adequate methodological approaches and efficient environments for FL acquisition has become a social concern (Lasagabaster, 2008).

In an attempt to improve learners' FL proficiency (predominantly English), Content and Language Integrated Learning (CLIL) programmes were implemented (Dalton-Puffer, 2011; Lasagabaster, 2008; Seikkula-Leino, 2007). Defined as "a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language" (Coyle et al., 2010, p. 1), CLIL provides an innovative pedagogical approach in which the FL acts as a medium of instruction in some curricular subjects (Dalton-Puffer, 2011; Eurydice, 2006; Seikkula-Leino, 2007), while content learning is not negatively affected (Admiraal et al., 2006).

In Spain, as a result of the unsatisfactory outcomes linked to the early teaching of the FL (García Mayo, 2003), new proposals centred on providing meaningful exposure to English, and this produced a rapid increase in the proportion of Spanish

institutions offering CLIL (Lasagabaster, 2009; Ruiz de Zarobe, 2008). However, in bilingual communities such as the Basque Autonomous Community (BAC) learners have to deal with the learning of the two official languages (i.e. Spanish and Basque) along with the augmenting presence of English through CLIL programmes (Ruiz de Zarobe & Lasagabaster, 2010). Therefore, assigning the precise time devoted to each language of instruction in the curriculum becomes a key issue in these officially bilingual communities.

Despite the social demand for learning multiple languages, language learning can be regarded as an intimidating experience for some students, who may face severe difficulties in the classroom (Onwuelgbuzie et al., 1999). In the field of second language acquisition (SLA), the research on individual differences explores the uniqueness of the individual mind, and it has been the dominant paradigm for several decades (Daubney et al., 2017). Within the affective dimension, anxiety has especially aroused experts' attention as a relevant psychological factor that may impact on learners' language achievement (Daubney et al., 2017; Teimouri et al., 2019). As a matter of fact, the negative role of anxiety has been extensively examined, and detrimental effects have been found on language achievement (Aida, 1994; MacIntyre & Gardner, 1991; Teimouri et al., 2019; Young, 1990), attitudes towards the target language (Onwuelgbuzie et al., 1999; Philips, 1992), and motivation (Somers & Llinares, 2021), among others. Moreover, this affective variable can influence FL students to the extent of blocking their learning, and subsequently leading them to abandon their language studies (Dewaele & Thirtle, 2009).

While FL learning can be considered to be anxiety-provoking (Lasagabaster & Doiz, 2017; Horwitz et al., 1986; Onwuelgbuzie et al., 1999), CLIL programmes not only extend FL exposure, but also endeavour to promote motivation and positive attitudes (Coyle et al., 2010; Pladevall-Ballester, 2019). Even though CLIL provision may be offered in primary and secondary education, the real picture is that, as Dalton-Puffer recognises, "CLIL is usually implemented once learners have already acquired literacy skills in their L1, which is more often at the secondary than the primary level" (2011, p. 184). This may have led to the tendency of the existing literature to focus on CLIL in secondary education, neglecting the research at other instruction levels, particularly in primary education.

In an attempt to fill this gap, this study aims to shed light on the impact of CLIL on fifth-grade primary school students' anxiety levels in the BAC. In this sense, we will delve into the interaction between anxiety and two instruction types, i.e. English-medium instruction (EMI) and Basque-medium instruction (BMI). Furthermore, learners' linguistic backgrounds will be taken into account in order to spot possible differences in their anxiety levels depending on the languages spoken in their daily

life. Moreover, the anxiety-achievement correlation as well as the effects of gender will be examined, to which little attention has been paid in CLIL settings in relation to anxiety. Accordingly, the novelty of this study lies in the comparison of this affective variable in two instructional settings (i.e. EMI and BMI) at primary education in a bilingual community, in which a minority language coexists with a majority language and a main FL.

2. Contextualising CLIL

It was not until the early 1990s that CLIL began to be promoted within the European Union to fill a need for greater levels of multilingualism and innovative methods in FL teaching (Eurydice, 2006). In CLIL classrooms, it is not the L1 of learners, but an additional language (not regularly spoken in their community) which is used as a medium of instruction to develop proficiency in both the content subject and the target language itself (Coyle et al., 2010; Dalton-Puffer, 2011; Eurydice, 2006; Seikkula-Leino, 2007).

In line with European policies, Spanish educational systems increasingly began to implement CLIL programmes to optimise FL learning conditions. In this linguistically diverse scenario, due to the several co-official languages spoken in some regions, various CLIL models were developed and adapted to each community's linguistic background (Ruiz de Zarobe & Lasagabaster, 2010). In monolingual regions, where CLIL has been implemented, curricular subjects are taught partly in Spanish and the FL, whereas in bilingual communities, the incorporation of the FL into the curriculum has to coexist with the two official languages, i.e. Spanish and the co-official regional language (Basque, Catalan/Valencian, and Galician).

The setting of the present study is the BAC, a bilingual community, in which Spanish and Basque are the official languages and the medium of instruction, and although English is hardly used outside the classroom, it represents the main FL in compulsory education. By the time CLIL was introduced into formal instruction in the BAC, three linguistic models were already established: model A, model B and model D (Ruiz de Zarobe & Lasagabaster, 2010). In model A, Spanish is used as the main language of instruction except for the Basque and English languages that are taught as subjects, while in model B, school subjects are taught proportionally in Basque and Spanish, excepting the English class. Finally, model D is a Basque immersion context as all the subjects are exclusively taught in Basque with the exception of the Spanish and English subjects. CLIL was generally implemented in compulsory secondary education, and initially, students learnt one or two curricular subjects in English. Nonetheless, the current picture is that CLIL is extended through the Basque educational system and varies from school to school in terms of the number

of subjects, hours of instruction per week, timing and organisation, albeit with a common objective: to improve learners' communicative competence in the FL across the school curriculum.

The CLIL approach offers richer opportunities than the early introduction of English and creates real communicative situations for using the FL in a more meaningful and efficient way, even in a context in which the FL has little social presence (Ruiz de Zarobe & Lasagabaster, 2010). When it comes to research, there is evidence confirming the linguistic benefits of the approach in grammar (Lasagabaster, 2008; Pérez Cañado, 2018a), vocabulary (Pérez Cañado, 2018a), listening (Lasagabaster, 2008; Pérez Cañado, 2018a), reading comprehension (Admiraal et al., 2006; Pérez Cañado, 2018a), writing production (Lasagabaster, 2008), and in oral production (Admiraal et al., 2006; Lasagabaster, 2008; Pérez Cañado, 2018; Ruiz de Zarobe, 2008). With respect to content learning, previous studies (Admiraal et al., 2006; Pérez Cañado, 2018b; Seikkula-Leino, 2007) found no statistically significant differences between CLIL and non-CLIL students, suggesting that learning subject matter through the FL does not hamper content learning. Likewise, CLIL does not entail negative effects on the development of L1 literacy skills either (Pérez Cañado, 2018b). In the BAC in particular, similar L1 proficiency levels were found after comparing CLIL and the traditional EFL approaches (Ruiz de Zarobe & Lasagabaster, 2010).

Furthermore, CLIL-type instruction is also believed to positively influence students' affective stance as it strengthens their motivation levels to learn FLs (Doiz et al., 2014; Pladevall-Ballester, 2019; Seikkula-Leino, 2007) and fosters positive language attitudes (Lasagabaster, 2009). Described as "a relatively anxiety-free environment" by Muñoz (2002, p. 36), CLIL has also been expected to reduce learners' anxiety, owing to the focus on meaning rather than on language form, as is usually the case in EFL teaching methods. However, findings related to the interaction between CLIL and anxiety are more disparate; in fact, some researchers (Lasagabaster & Doiz, 2017; Papaja, 2019; Somers & Llinares, 2021) reported that this approach can create high-anxiety learning environments, considering the difficulty that studying already-complex subjects in the FL may pose.

3. Literature review

3.1. Foreign Language Anxiety

At the most general level, anxiety has been characterised as the subjective feeling of tension, fear, uneasiness, nervousness and worry, closely related to "an arousal of the autonomic nervous system" (Horwitz et al., 1986, p. 125). Language learning is prone to anxiety-arousal, as experiencing anxiety is believed to be part of the learning

process (Daubney et al., 2017). As a result, Foreign Language Anxiety (henceforth, FLA) was conceptualised as a distinct variable and unique to the FL learning context, and is one of the most prominent individual difference factors that may explain differential success in language acquisition (Daubney et al., 2017). A widely-accepted definition of FLA was provided by Horwitz and colleagues (1986), who conceived this situation-specific anxiety as "a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (p. 128). To measure the scope and severity of FLA, they also developed the Foreign Language Classroom Anxiety Scale (FLCAS), a 33-item survey with a Likert type scale, which succeeded in identifying significant FLA levels among university students. Owing to its reliability and internal consistency, the FLCAS became a well-established instrument and was extensively used in later studies (Aida, 1994; De Smet et al., 2018; Dewaele & MacIntyre, 2014; Philips, 1992).

A large body of literature has investigated the debilitating impact of FLA on language acquisition not only in various domains of language performance, but also in non-linguistic aspects of the learning process. Indeed, high levels of FLA in the classroom may induce unfavourable emotional reactions and negative attitudes towards the learning of the language, as well as discourage students from continuing their language studies (Daubney et al., 2017; Dewaele & Thirtle, 2009). In addition, the enjoyment and motivation that are essential to engage in language learning may be adversely affected (Dewaele & MacIntyre, 2014; Somers & Llinares, 2021).

FLA has been claimed to be a significant predictor of language achievement, as it was found to be negatively correlated with course grades or standardised language proficiency tests (Aida, 1994; Elkhafaifi, 2005; Horwitz et al., 1986; MacIntyre & Gardner, 1991; Philips, 1992; Saito et al., 1999). In other words, higher FL anxiety levels were associated with a poorer performance in the FL and vice versa. Nonetheless, a meta-analysis conducted by Teimouri et al. (2019) found that correlations vary dramatically across studies, and the overall correlation found was low to moderate ($r = -.36$). Considering the small magnitude of the negative correlations and the complex, multidimensional nature of anxiety, cause-effect relationships between language anxiety and achievement should be interpreted with caution, because higher anxiety levels do not automatically imply a lower language achievement (Lasagabaster & Doiz, 2017). Likewise, students with a higher propensity for anxiety may achieve outstanding results, whereas there may be others who perform more poorly without experiencing anxiety (Onwuelgbuzie et al., 1999; Park & French, 2013).

The sources for such a negative emotion in the language classroom are multiple, and may be associated with the learner, teacher and the specific learning context. On the one hand, some relevant learner-internal variables include: (i) students' age of

onset of learning, with late starters suffering more from FLA; (ii) FL learning type, particularly higher FLA levels are reported when the language is exclusively learnt through formal instruction without exposure to the FL outside the classroom; (iii) number of languages known, as the more languages learners know, the lower the FLA they feel; (iv) frequency of FL use, being more frequent users less anxious; and (v) gender, in which discrepant results have been found (Dewaele & Al-Saraj, 2015). Further causes related to the learner may lie in their classroom performance, fear of negative evaluation or making mistakes, learning difficulties and some personality traits, in particular perfectionism, extraversion, psychoticism, tolerance of ambiguity, emotional stability and social initiative (Dewaele & Al-Saraj, 2015; Dewaele & Shan Ip, 2013; Onwuegbuzie et al., 1999).

On the other hand, the causes centred on the teacher characteristics may be linked with their teaching methods, level of supportiveness and emotional involvement, personality, unpredictability, and evaluation procedures (Onwuegbuzie et al., 1999; Young, 1990).

When it comes to the factors that emerge in the learning context, peers play a major role since their opinions, laughter and negative judgements can provoke anxious feelings in learners (Horwitz et al., 1986; Young, 1990). The complexity of the target language, classroom atmosphere, class size, exam pressure, methodology and tasks, especially those in which oral communication is involved, may also contribute to feelings of anxiety (Aida, 1994; De Smet et al., 2018; Papaja, 2019; Philips, 1992). In particular, researchers (Horwitz et al., 1986; Young, 1990) have observed that performing orally in a language in which learners do not have full competence is one of the most anxiety-provoking situations due to their fears of embarrassment, speaking publicly or being mistaken.

3.2. FLA in CLIL settings

Within the aforementioned myriad of factors, the methods employed in the language classroom, especially those involving communication, have been considered to potentially contribute to furthering anxiety (Arnold & Brown, 1999). CLIL serves as a clear example of a communication-based approach, which is simultaneously expected to have a positive impact on the learner's affective dimension, namely motivation, language attitudes and FLA (Admiraal et al., 2006; Doiz et al., 2014; Lasagabaster, 2008; Pladevall-Ballester, 2019; Ruiz de Zarobe, 2008). Nevertheless, a few scholars have claimed that CLIL could even exacerbate learners' anxiety, given the demanding process of learning subjects, already complex or with a heavier academic load, in the FL (Lasagabaster & Doiz, 2017; Papaja, 2019; Somers and Llinares, 2021).

The influence of CLIL in FLA was regarded as positive when CLIL learners were found to experience less FLA than their non-CLIL counterparts in many studies (De Smet et al., 2018; Heras Aizpurua, 2016; Simons et al., 2019; Thompson & Sylvén, 2015). In Spain, Heras Aizpurua (2016) observed that CLIL secondary education students felt less anxiety using English when compared to their EFL peers. Similarly, Doiz et al. (2014) evidenced how the levels of anxiety exhibited in CLIL by first-year secondary school students vanished among third-year secondary CLIL students, which may be due to their increased familiarity with using English in this learning context after two academic years of CLIL-type exposure. However, CLIL learners still exhibited a slightly higher level of anxiety than their non-CLIL peers, which was attributed to the more demanding process that learning subjects in the FL entails since CLIL learners run the risk of facing more lexical difficulties.

In other settings like Sweden, even before the start of their CLIL experience, less anxious reactions were perceived among CLIL pupils in a secondary school (Thompson and Sylvén, 2015). Instead of the effect of the methodology itself, other pre-existing individual factors may have come into play given that CLIL was a voluntary option. However, the authors acknowledged the need of a longitudinal study to better illustrate the influence of CLIL over time. In Belgium, De Smet et al. (2018) ascribed the lower anxiety levels in CLIL to the more favourable profiles (i.e. higher socio-economic status and nonverbal intelligence, and lower school failure) of the learners participating in this approach in primary and secondary education. Simons et al. (2019) also found that feelings of anxiety in a group of secondary CLIL students decreased after one year of CLIL education, albeit slightly, due to their growing confidence to use the FL.

Adverse or non-satisfactory effects of CLIL were also encountered in previous research. Lasagabaster and Doiz (2017) longitudinally investigated the impact of CLIL on affective factors in two age groups in the BAC, and observed that unexpectedly CLIL learners' anxiety was non-significantly lower in third grade than in first grade in the case of the younger students. Conversely, levels of anxiety of both CLIL and non-CLIL older participants significantly increased from third to fourth school year. A plausible explanation for this increase of anxiety in later grades was linked to the complexity and cognitive demand of subjects delivered. But they concluded that higher FLA levels may not cause detrimental effects on language learning, as anxiety does not necessarily involve a poorer achievement. In another Spanish community (Madrid), significantly higher feelings of anxiety were reported by secondary education students in their CLIL classes than in non-CLIL classes, due to a more demanding level of teaching involved in secondary education and CLIL lessons (Somers & Llinares, 2021). These findings led the authors to conclude that CLIL may have a negative impact on anxiety.

Similarly, Papaja (2019) observed that learners of secondary schools from three different countries felt more tense and nervous in CLIL subjects than in their regular English subjects, especially when discussing specific contents in English because of their struggles to express what they meant or fear of making mistakes. Most of the participants also considered that other CLIL pupils were more competent in the FL, and felt overwhelmed with the amount of material to be covered in the CLIL curriculum as a result of its dual focus on content and FL learning.

3.3. FLA related to learners' linguistic repertoire

As in the case of methodology, prior linguistic knowledge has also been argued to be determinant in the learning of a new language in that it may assist or hinder its acquisition, in view of the cognates shared by the languages (Odlin, 1989).

As regards FLA, similar scores have been found in previous studies with different native language-target language pairings (Aida, 1994; Saito et al., 1999), suggesting that general FLA is somewhat independent of the distance between the L1 and the L2/FL. Nonetheless, some skill-specific anxieties such as reading anxiety were argued to potentially depend on the L1-L2 distance (Elkhafaifi, 2005; Saito et al., 1999). In Saito and colleagues' (1999) study, reading anxiety of English learners of three different FLs (French, Russian and Japanese) was lower when a FL with a closer proximity to English was targeted (e.g. French) as compared to a more distant FL (e.g. Japanese). The fact that reading Japanese created significantly higher anxiety was explained by the difficulties to read the language due to its unfamiliar, non-Roman writing system in contrast to the L1.

Not only typologically, but the degree of FLA could also fluctuate in terms of the amount of exposure to the languages at the learner's disposal. Despite the internationalisation of English and its ever-growing role in education in the BAC, the opportunities to use the language beyond the school walls are few. For example, Santos Berrondo (2017) observed that there were higher anxiety levels in English than in the L2, Spanish or Basque, among university students owing to the limited exposure to the FL outside the classroom. Likewise, Dewaele (2007) tested among multilingual speakers that FLA gradually increased over the languages learnt later in life, since speakers were typically more proficient in the languages acquired earlier and used more frequently. Later, Dewaele and Al-Saraj (2015) reasoned that FLA was somewhat associated with the frequency of FL use, indicating that the speakers who used the language more frequently were less susceptible to suffer from anxiety in the FL.

Furthermore, Santos Berrondo (2017) found that there was no variation in FLA levels with respect to participants' L1, as they had a similar academic background and

exposure to English. However, they differed in their L2, in other words, L1-Spanish speakers felt more anxious in Basque than L1-Basque speakers in Spanish, because of the minority status of Basque and the fewer opportunities to use the language in daily communication in the Basque society. While Spanish is the dominant language in most areas of the BAC, Basque remains a minority language and is less frequently used than Spanish, notwithstanding the language plans for the consolidation and promotion of Basque in different sectors such as education, society, public administration and media –Basque television, radios and newspapers– (Cenoz, 2009).

Following these patterns, our study expects higher levels of anxiety in the FL (English) than in students' L2 (Basque) irrespective of their L1 (Basque, Spanish or both), because exposure to the former is usually higher. In Basque, the principal language of instruction and our participants' L1 or L2, L1-Spanish students are expected to suffer more from anxiety than their L1-Basque or Basque-Spanish bilingual counterparts.

3.4. Gender differences in FLA

The study of the role of gender in language learning has recently received a great deal of attention, and research has detected a statistically significant trend, albeit with small effect sizes, showing a consistent female advantage over males (Aida, 1994; Park & French, 2013). Not only do females outperform their male counterparts in language achievement and skills, but they have also been reported to manifest a stronger motivation to learn the target language, and a more positive disposition and attitudes towards it (Dewaele et al., 2016; Dewaele & MacIntyre, 2014).

With reference to gender differences in FLA, it was female learners who also scored higher levels of FLA in previous studies. Many researchers (Elkhafaifi, 2005; Park & French, 2013; Santos Berrondo, 2017) examined the effect of gender on FLA among university students, and significant differences were perceived in favour of females in the different linguistic contexts. Dewaele and MacIntyre (2014) investigated FL anxiety and enjoyment in a large sample from different education levels, and females scored higher means in both scales, suggesting that they experienced more negative and positive emotions in the FL classroom. In a follow-up study by Dewaele et al. (2016), while female participants reported less confidence in using the FL and worried more about their mistakes than males, they simultaneously had significantly more fun, pride in their FL performance and interest in the FL class.

In contrast, empirical data has indicated that females can also feel less anxiety than male learners (Campbell & Shaw, 1994; Kitano, 2001; MacIntyre et al., 2002). Higher anxiety levels of secondary males in a French immersion program (MacIntyre

et al., 2002) and postsecondary males in an FL intensive course in the USA (Campbell & Shaw, 1994) were related to a combination of different variables such as a lower self-consciousness and willingness to communicate, attitudes towards FL learning, learning styles, fear of academic failure, teacher differences and teaching methodology. In a study carried out by Kitano (2001), male students enrolled in Japanese language courses exhibited higher anxiety levels than females given their lower perceived competence in spoken Japanese.

In spite of the significant gender-based divergence identified in research, effect sizes are generally small. Besides, other researchers did not spot gender differences at all (Aida, 1994; Dewaele & Al-Saraj, 2015). Interestingly, some researchers found that gender-differences tend to diminish in terms of motivation in CLIL settings (Fernández Fontecha & Canga Alonso, 2014; Heras Aizpurua, 2016; Heras & Lasagabaster, 2015). A possible explanation is that female and male learners appear to be equally motivated to learn the target language in the CLIL classroom since there might have been an increase in males' motivation to learn both the language and the subject matter, owing to the richer and more meaningful environment created by CLIL-type exposure.

As for the interface of anxiety, gender and CLIL, there is a scarcity of research, hence the present study will compare the degrees of language anxiety experienced by female and male learners in the Basque learning context. Given the vanishing effects of CLIL on gender differences in motivation, it may be the case that this approach helps to reduce the gender-based divergence in our participants' anxiety levels. Yet, Simons et al. (2019) observed that female CLIL learners were more inclined to FLA than their male peers.

The goal of this study is, thus, to provide a profound insight into the anxiety experienced by primary education students when receiving CLIL, given the dearth of research on this affective variable in CLIL contexts at primary education. In this respect, whether levels of FLA fluctuate depending on language of instruction, achievement, linguistic repertoire and gender will be analysed. Specifically, the setting in which this investigation takes place is linguistically diverse, and the differences among the languages present in the BAC could be decisive in the experience of anxiety. Within this framework, the research questions of this paper are the following ones:

1. Do primary education students experience language anxiety, and if so, is it higher in EMI or BMI?

2. How do learners' language anxiety levels differ in terms of gender, home languages and achievement?

4. Methodology

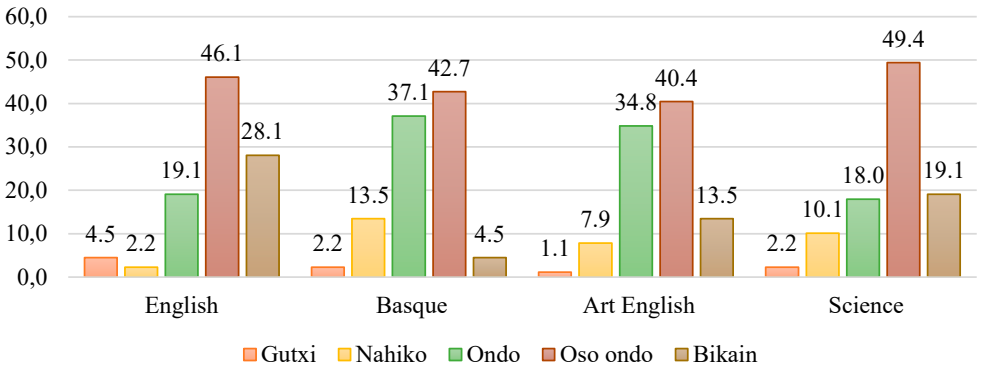
4.1. Participants

The current study investigates the impact of language anxiety in a group of 89 fifth-grade (48 females, 37 males and 3 non-binaries) FL learners of English in a primary, charter school in Bilbao. The mean age of the sample was $\bar{x}$ =10.35. Although all the students were born in the BAC, their home languages varied considerably: the sample consisted mostly of Spanish speakers, exclusively (66.3%) or in addition to Basque (15.7%), but others used only Basque at home (12.4%) and a combination of other languages such as Spanish and English (1.1%), Basque and English (1.1%) or Basque and Polish (1.1%).

By the time of data collection, students were enrolled in model D, and they were all receiving CLIL classes in English. They started learning English at the age of 3, devoting three sessions per week to the FL teaching in the school curriculum. In the case of CLIL subjects, they were attending art English lessons since the first year of primary education and science since third grade, for two hours each, making a total of four CLIL lessons during the week, and seven (plus three hours of EFL classes) are the sum of all the weekly hours that participants were exposed to the FL in compulsory education.

To measure Basque learners' achievement in the current study, their last term's course grades in the subjects of English, Basque, art English and science were taken into account, as displayed in figure 1. In Basque primary education systems, according to the BOPV (Boletín Oficial del País Vasco, 2016), the grading system is based on the following terms: *Gutxi* ('insufficient'), *Nahiko* ('sufficient'), *Ondo* ('good'), *Oso ondo* ('very good') and *Bikain* ('outstanding'). While 'very good' and 'excellent' encompass two numerical grades, i.e. 7-8 and 9-10 respectively, 'sufficient' and 'good' only refer to one grade, i.e. 5 and 6 respectively. For that reason, students scoring 'sufficient' and 'good' in their grades were merged into the same group to conduct the statistical analyses.

Figure 1: Profile of the participants – Grades



External exposure to English was also observed. Overall, 60.7% of the participants were attending English lessons outside school for an average of 4.5 years and 1.8 hours per week. Moreover, 27% of the participants indicated that they had travelled abroad: 10.1% to an English-speaking country, 13.5% to a non-English-speaking one, and 3.4% to both. Those who had visited foreign countries spent 65.1 days in total on average, namely, 21.3 days in an English-speaking destination and 93 days in a non-English-speaking one.

4.2. Instruments and procedure

This study used a mixed method approach to measure students' FLA in EMI and BMI, and is cross-sectional in its design, as data were gathered at one, albeit separate, point in time. Quantitative data were collected through an anxiety questionnaire, developed on the basis of the FLCAS, which was originally created by Horwitz et al. (1986) and widely used in previous studies. It consisted of 38 items presented on a five-point Likert type scale ranging from *strongly disagree* (1) to *strongly agree* (5), encompassing anxiety episodes in both EMI and BMI classes (a shortened version of the instrument in English is attached in the Appendix section). Participants completed the questionnaire in Basque, and it was administered in mid-May in paper-based format within school hours. Ten items were phrased to indicate low anxiety in an attempt to make learners pay careful attention when filling out the questionnaire. These items were later reverse-coded so that high scores in all the items indicate high anxiety in the scale. Likewise, the items included in each instruction type were equivalents, albeit differently formulated, representing thus the same theme under study. Also, four items referred to possible anxiety experiences linked with COVID-19 measures imposed by the school. The purpose of devoting a subsection to the latter is to observe whether the current pandemic situation and health protocols could affect learners' anxiety levels in the classroom.

To collect qualitative data semi-structured interviews were carried out in June. In this phase, 10 participants (7 females and 3 males) were selected by the school (they were asked to choose heterogeneous students in terms of gender, linguistic background and proficiency) and grouped in pairs (see table 1) so that they could feel comfortable, without teacher or peer pressure to express their perspectives and classroom experiences openly; the aim was to capture relevant thoughts, comments and reflections, with which learners may not arise in individual interviews. Learners responded to several questions, concerning their anxiety in the process of Basque and English learning, either in Basque or Spanish, as preferred by participants. The length of each interview varied from 40 minutes to almost 2 hours, depending on how forthcoming participants were to share their learning experiences and thoughts.

Table 1: Profile of the sample selected for the interview

Pairs	Students	Gender	L1	Home language	Grades	
					English	Basque
Pair 1	ST1	Female	B & S[a]	S	Good	Very good
	ST2	Female	B & S	B & S	Excellent	Excellent
Pair 2	ST3	Male	B[b]	B	Very good	Very good
	ST4	Female	S[c]	S	Very good	Excellent
Pair 3	ST5	Female	B	B & S	Excellent	Very good
	ST6	Male	S	S	Excellent	Good
Pair 4	ST7	Female	B & S	S	Very good	Sufficient
	ST8	Female	B & S	S	Excellent	Very good
Pair 5	ST9	Female	B & S	S	Very good	Good
	ST10	Male	B & S	S	Very good	Good

a. Basque & Spanish

b. Basque

c. Spanish

Prior to the main data collection, a pilot study was conducted so as to test and maximise the efficacy of the research instruments employed. Three similarly-aged students, not included in the sample, were arbitrarily selected to complete this preliminary phase in April, in which they filled out the anxiety questionnaire in Basque and took part in the interview in Spanish. After the pilot test, some items from the questionnaire were reformulated and highlighted to facilitate the reading

and comprehension of the text. Additionally, the question related to learners' L1s created great confusion among participants, and therefore, their home languages were used as a parallel variable to the L1s to represent participants' linguistic repertoire. Furthermore, the parents of each participant, including both the pilot test and main study, were asked to sign a parental consent form, allowing their children to participate in the study. Due to the pandemic situation, the teachers of the school took charge of distributing the questionnaires, and the interviews were conducted by the researcher via an online platform.

4.3. Data analysis

Empirical data was examined through two distinct procedures. On the one hand, the Statistical Package for Social Sciences (SPSS) 25 was used to analyse quantitative results. More specifically, reliability analyses were first carried out to validate the internal consistency of the present scale, followed by factor analyses in order to identify common underlying themes that may explain the variability of our set of observed variables (Dörnyei, 2010; Larson-Hall, 2010). Before conducting any statistical test, a Kolmogorov-Smirnov test was computed in order to know whether the distribution of our data set was normal (Larson-Hall, 2010). In addition, mean scores and standard deviations were calculated for the different anxiety categories in both languages of instruction. Wilcoxon signed-ranks tests were performed to compare means of anxiety in EMI and BMI (Larson-Hall, 2010). Mann Whitney U and Krukall-Wallis tests were also computed to observe whether there were statistically significant differences between groups in two of our independent variables (gender and home languages) in relation to anxiety. These variables' effect sizes were measured as well to establish the magnitude of their influence on anxiety. Finally, the Spearman's correlation coefficient between participants' grades and anxiety scores were calculated to test the strength of their relationship.

On the other hand, interview transcriptions were examined by means of the systematic thematic analysis (TA) to identify patterns across our data set and make sense of these commonalities. Developed by Braun and Clarke (2012), a six-phase approach to TA consisted in: (1) familiarising with the data set's content, (2) generating initial descriptive codes, (3) searching for patterns associated with the extracts, (4) reviewing potential themes to capture the most important elements, (5) defining themes relevant to the research questions, and (6) producing the report by connecting the themes coherently and including the interpretation of the data. The selection of the themes is in accordance with the items included in the quantitative analysis, and further relevant features mentioned by the participants were also taken into account. These selected themes will be depicted in participants' quotations in the following section, and will simultaneously reinforce the results obtained from the questionnaires.

5. Results

Prior to any type of analysis, the Cronbach's alpha coefficient was measured in order to check the reliability of the instrument used in the study. At this phase of the process, 6 items were excluded due to their incompatibility with the rest of the target items. The present scale showed a satisfactory internal consistency in the entire set of target items with $\alpha = .754$ ($n = 32$), in the anxiety items related to EMI with $\alpha = .712$ ($n = 14$) and in those related to BMI with $\alpha = .693$ ($n = 14$). Furthermore, items from the scale were classified into two main factors in each language of instruction: (1) Classroom performance and atmosphere, and (2) Knowledge and difficulties in the language and subjects, along with an additional factor (3), alluding to COVID-19 measures. Reliability coefficients for the two main categories were adequate in both languages: Factor 1 with $\alpha = .608$ ($n = 9$) and Factor 2 with $\alpha = .641$ ($n = 5$) in EMI, and Factor 1 with $\alpha = .629$ ($n = 9$) and Factor 2 with $\alpha = .602$ ($n = 5$) in BMI; but it was not for the items composing the COVID-19 category ($\alpha = .491$, $n = 4$).

A Kolmogorov-Smirnov test was performed, revealing that the distribution of the scores was not normal ($KS = 0.082$, $p < 0.05$), and as a result, non-parametric statistics were computed to answer the research questions.

RQ1. Do primary education students experience language anxiety, and if so, is it higher in EMI or BMI?

To address the first research question, descriptive statistics for global anxiety scores plus the three factors under study are presented in table 2. These students experienced a low level of language anxiety on average in fifth-grade primary education ($\bar{x} = 2.60$). However, anxiety linked with COVID-19 was moderate ($\bar{x} = 3.22$), hence it was higher than global anxiety. The Wilcoxon signed ranks test showed that the anxiety experienced in the classroom significantly differed from one language to the other. They appeared to be statistically more anxious in their subjects taught in English ($\bar{x} = 2.63$) than in Basque ($\bar{x} = 2.39$) (p-value = .003), and the effect of the language of instruction was moderate ($r = .31$). The difference between the two main factors, that is, the anxiety associated with classroom performance and the complexity of the subjects in EMI ($\bar{x} = 2.71$ (F1), $\bar{x} = 2.48$ (F2)) and BMI ($\bar{x} = 2.49$ (F1), $\bar{x} = 2.22$ (F2)) was statistically significant (p-value = .014 (F1), p-value = .027 (F2)). Even though both factors' mean scores were higher in EMI, the effect of the language of instruction was small in the two cases ($r = .26$ (F1), $r = .23$ (F2)).

Table 2: Descriptive statistics and Wilcoxon signed ranks test on global anxiety and factors in EMI and BMI

		Mean	SD	Z	Asymp. Sig. (2-tailed)	Effect size (r)
Anxiety	EMI	2.63	.64	-2.948 [a]	**.003**	.31
	BMI	2.39	.62			
	Global	2.60	.48			
F1: Performance & atmosphere	EMI	2.71	.69	-2.468 [a]	**.014**	.26
	BMI	2.49	.73			
	Global	2.60	.58			
F2: Knowledge & difficulties	EMI	2.48	.87	-2.208 [a]	**.027**	23
	BMI	2.22	.79			
	Global	2.35	.61			
F3: COVID-19		3.22	.98			

a. Based on positive ranks.

To understand these results in more detail learners' contributions in the interviews were closely examined. The following words supported the fact that these students experienced global anxiety to different degrees: "I don't feel very nervous, but a bit more than intermediate" (ST1), "For me it's very strong" (ST8), "Sometimes I can't avoid it, but other times I can control it" (ST9), "It's not a big deal" (ST10). But, they were aware of the variability of these feelings, as most of them expressed having a decline in anxiety since they started CLIL and a possible increase in the subsequent academic courses. Also, learning subjects in English was perceived as beneficial for several reasons:

> (ST2): "When I started, I was nervous because I didn't know very well what science was, and now I'm not so much, but I'm when we do individual activities. [...] I'm fine with these subjects because that way we master English more and if we go abroad, it's easier that way. [...] I'm nervous because of DBH ('Compulsory Secondary Education') because it seems to be more difficult".
>
> (ST5): It's always good for travelling and for communicating with people from there. [...] I have less anxiety now because we've been learning them since third grade. [...] (In two years' time), it'll be higher (in English), because it

> seems to me that it's even more difficult but apart from this, you have more things to learn".
>
> (ST8): "It'd be useful for travelling to places and as English is one of the most spoken languages, so to understand people from there. [...] Now I have less anxiety because I've become used to it, and this makes me feel better. [...] It'll be higher (in Basque), because you learn new things, and the teacher has already told us that what we're learning here will be like harder in two years' time".

With regard to COVID-19 anxiety, qualitative data disagreed with quantitative results, since most participants denied suffering from anxiety with current health measures, but commonly described feeling "discouraged", "listless" and "frustrated" instead.

> (ST4): "The issue of friends doesn't make me feel so nervous, I don't care about it. It discourages me, but it doesn't bother me so much".
>
> (ST8): "It makes me feel a bit more listless, because this year's school trips were all cancelled due to coronavirus".
>
> (ST10): "I get frustrated. It discourages me a little bit because when we're in the break time we can't meet students from other classes, and for example, in the building below, they can meet other students, but we can't".

When it comes to EMI and BMI classes, students' comments in the interviews seemed to support quantitative data in that most of them felt more anxious in English subjects; however, three participants reported experiencing the opposite. The two main factors in our set of questionnaire items, classroom performance (F1) and knowledge and difficulties in the language or subject matter (F2) were mentioned as possible sources of feelings of anxiety together with other factors, including frequency of use and personal interests in the language.

> (ST1): "In English, I'm a little bit more nervous than in Basque, because in English it's more difficult. [...] Not daily, but I always speak Basque, and it's my language. I don't speak English every day, and thus, I find it more difficult".
>
> (ST8): "I'd prefer to study the subjects in English or Spanish because I find it easier to speak in other languages than Basque because I don't like Basque. [...] (CLIL) is a way of learning faster. It's like science and in English at

> the same time, so it's better. When I have to speak aloud in Basque, I feel uncomfortable because sometimes I get tongue-tied, and I don't know how to continue. I like English subjects a lot. I feel comfortable. I hardly ever get tongue-tied".
>
> (ST10): "In English, I feel with a bit more like anguish because there are words that I don't understand, and it's like it embarrasses me to ask."

As for the individual items, intergroup comparisons of equivalent pairs in EMI and BMI were made, and results in table 3 indicated that there were statistically five significant differences out of 14 pairs. Apparently, students suffered from anxiety when handling their classmates' laughter in their subjects taught in Basque ($\bar{x}$ = 2.73) more than in English classes ($\bar{x}$ = 2.21) (p-value = .020), but the effect size was small (r = .25). Nevertheless, participants exhibited substantially higher levels of anxiety in EMI due to their embarrassment to volunteer ($\bar{x}$ = 2.43 (EMI), $\bar{x}$ = 1.93 (BMI)) (p-value = .019), when speaking without preparation ($\bar{x}$ = 3.15 (EMI), $\bar{x}$ = 2.46 (BMI)) (p-value = .000) and giving oral presentations ($\bar{x}$ = 3.89 (EMI), $\bar{x}$ = 3.20 (BMI)) (p-value = ,000), plus because of the difficulty of learning subjects in that language ($\bar{x}$ = 2.37 (EMI), $\bar{x}$ = 1.63 (BMI)) (p-value = ,000). The effect of the language of instruction was moderate when the teacher asked them questions without prior preparation (r = .39), they gave presentations in front of the class (r = .39) and when they struggled learning the subjects because of the difficulty of the language (r = .42). In both cohorts, the item in which students experienced the highest level of anxiety was when they were required to deliver an oral presentation in the classroom, especially in English as the degree of anxiety was nearly large.

Table 3: Descriptive statistics and Wilcoxon signed ranks test on independent items in EMI and BMI

		Mean	SD	Z	Asymp. Sig. (2-tailed)	Effect size (r)
Item 20 -	BMI	1.93	1.25	-2.344[a]	**.019**	.25
Item 1	EMI	2.43	1.38			
RCItem 21 -	BMI	1.78	1.41	-.012[a]	.991	
Item 2	EMI	1.76	1.23			
Item 18 -	BMI	2.46	1.42	-3.697[a]	**.000**	.39
Item 3	EMI	3.15	1.50			
Item 28 -	BMI	3.02	1.49	-1.297[b]	.195	
Item 4	EMI	2.80	1.32			
Item 30 -	BMI	2.87	1.47	-.527[a]	.599	
Item 6	EMI	2.99	1.60			
RCItem 26 -	BMI	2.11	1.32	-1.823[a]	.068	
RCItem 7	EMI	2.54	1.53			
RCItem 27 -	BMI	2.31	1.33	-.101[a]	.919	
Item 8	EMI	2.33	1.35			
RCItem 32 -	BMI	2.73	1.64	-2.322[b]	**.020**	.25
Item 9	EMI	2.21	1.58			
RCItem 19 -	BMI	3.20	1.46	-3.713[a]	**.000**	.39
Item 10	EMI	3.89	1.27			
Item 17 -	BMI	2.64	1.46	-1.560[a]	.119	
Item 11	EMI	2.96	1.36			
Item 22 -	BMI	2.56	1.53	-.351[a]	.725	
Item 12	EMI	2.63	1.49			
Item 23 -	BMI	1.63	1.26	-3.949[a]	**.000**	.42
Item 15	EMI	2.37	1.43			
Item 31 -	BMI	1.78	1.19	-1.557[a]	.119	
Item 16	EMI	2.09	1.35			
Item 38	BMI	2.49	1.05	-1.519[a]	.129	
Item 37	EMI	2.63	.95			

a. Based on positive ranks.

b. Based on negative ranks.

Table 4 includes the items in which higher percentages of anxiety (4 or 5 on the scale) were expressed. It is worth noting that a very high percentage of students (68.5%) found presentations in English anxiety-provoking (a lower but still high 46% in Basque). High percentages were also found especially in EMI classes when students were asked questions they had not prepared in advance (46.1% in EMI versus 29.2% in BMI), because they considered that other students were better at English/Basque (40.5% vs. 32.6%), or when they felt afraid of not understanding the content (40.5% vs. 31.5%).

Table 4: Items with the highest percentages of anxiety in EMI and BMI

	Frequency (%)					
	0	**1**	**2**	**3**	**4**	**5**
Item3	5.6%	12.4%	12.4%	23.6%	23.6%	22.5%
Item18	4.5%	24.7%	30.3%	11.2%	19.1%	10.1%
Item6	7.9%	14.6%	12.4%	24.7%	16.9%	23.6%
Item30	6.7%	12.4%	19.1%	29.2%	14.6%	18%
Item10	1.1%	6.7%	5.6%	18%	25.8%	42.7%
Item19	3.4%	12.4%	15.7%	22.5%	21.3%	24.7%
Item11	3.4%	14.6%	18%	23.6%	28.1%	12.4%
Item17	4.5%	22.5%	22.5%	19.1%	18%	13.5%

Data obtained from interviews equated to questionnaire results in that most students reported experiencing anxiety when speaking or making oral presentations, especially in English. Other instances of anxiety were related to their classmates and specific activities in which they worried about making an unsuccessful performance.

> (ST1): "In Basque, I'm nervous most of the times, when we have a worksheet, but when we have regular classes I'm well. In English, I'm more nervous in presentations because I get tongue-tied many times. All of them are staring at you, and if you get tongue-tied I feel more nervous and then I do it worse. If I say it wrong, I feel nervous. Some classmates start to laugh. [...] I think that if you're with more people, you feel nervous because there are a lot of people with you".

> (ST6): "In oral presentations in Basque, when there is something that doesn't come up, I get very nervous there. In English, nothing causes me nervousness.

> [...] You may feel nervous in oral presentations or exams or when they ask you something that you aren't studying because you don't know what to say, like a surprise question. [...] When you're with people who are disturbing, this can cause you anxiety. You aren't comfortable. In the presentation of this term instead of listening to me they were talking or playing. I felt really bad and nervous".

RQ2. How do learners' language anxiety levels differ in terms of gender, home languages and achievement?

In view of the few non-binary participants in the sample (n = 3), this group was excluded from the statistical analyses, and Mann Whitney U tests were carried out to compare anxiety levels between female and male students. Table 5 offers the gender-based differences in global and COVID-19 anxieties plus in the anxiety categories in EMI and BMI. As can be observed, there was only one statistically significant difference between the two gender samples, namely in the COVID-19 category (p-value = .036), being females more anxious ($\bar{x}$ = 3.37) than males ($\bar{x}$ = 2.99). Yet, the effect of gender on COVID-19 anxiety was small (r = .23). In the contrasts between EMI and BMI, both genders experienced more anxiety in English subjects, and females appeared to exhibit higher anxiety levels, although without statistical significance.

Table 5: Descriptive statistics and Mann Whitney U test on gender-based differences on global anxiety, COVID-19 anxiety and factors in EMI and BMI

		EMI				BMI				Global				
		M	**SD**	**Z**	**Sig.**	**M**	**SD**	**Z**	**Sig.**	**M**	**SD**	**Z**	**Sig.**	**Effect size (r)**
Anxiety	F[a]	2.61	.62	-.328	.743	2.45	.71	-.808	.419	2.64	.54	-1.508	-132	
	M[b]	2.56	.62			2.31	.51			2.50	.37			
F1	F	2.77	.65	-1.416	.157	2.57	.79	-.937	.349					
	M	2.55	.69			2.41	.58							
F2	F	2.34	.86	-1.326	.185	2.22	.83	-.009	.993					
	M	2.57	.83			2.12	.65							
COVID-19	F									3.37	.99	-2.100	**.036**	.23
	M									2.99	.93			

a. Female
b. Male

Although females turned out to be more anxious due to pandemic measures than males in the questionnaires, gender differences did not arise in the interviews, as participants described feeling anxious to a greater or lesser extent, independently of their gender:

> (ST5) (F): "It makes me mad because I get totally depressed. A friend that I haven't seen for a long time and that I can't hug, it discourages me a lot".
>
> (ST6) (M): "It discourages you a lot".

The second variable under research deals with students' linguistic repertoire, and the influence of speaking a certain home language in the anxiety experienced in each language of instruction was examined. Those speaking languages other than Basque, Spanish or both of them at home were discarded, as only one participant was included in each of these groups. Table 6 revealed that users of Spanish had significantly more anxiety in EMI classes ($\bar{x}$ = 2.72) than users of Basque ($\bar{x}$ = 2.59) or of both Basque and Spanish ($\bar{x}$ = 2.32) (p-value = .033), and the effect of the languages spoken at home was moderate (η^2 = .06). Also, those speaking Spanish at home were the ones facing the greatest difficulties in both language classes (p-value = .031 (EMI), p-value = .005 (BMI)), and the effect size of the variable was moderate (η^2 = .06 (EMI), η^2 = .11 (BMI)).

Table 6: Descriptive statistics and Kruskal-Wallis test on home languages-based differences on global anxiety and factors in EMI and BMI

		EMI					BMI					Global			
		M	**SD**	**H**	**Sig.**	**Ef. size (η^2)**	**M**	**SD**	**H**	**Sig.**	**Ef. size (η^2)**	**M**	**SD**	**H**	**Sig.**
Anxiety	B[a]	2.59	.57	6.836	**.033**	.06	2.17	.48	2.550	.279	2.42	2.48	.33	4.501	.105
	B & S[b]	2.32	.45				2.25	.56				.33			
	S[c]	2.72	.65				2.48	.66			2.68	.52			
F1	B	2.64	.58	2.067	.356		2.39	.65	.018	.991					
	B & S	2.54	.49				2.48	.65							
	S	2.79	.74				2.53	.78							
F2	B	2.51	.98	6.971	**.031**	.06	1.76	.46	10.725	**.005**	.11				
	B & S	1.91	.67				1.84	.52							
	S	2.58	.81				2.39	.85							

a. Basque
b. Basque & Spanish
c. Spanish

Qualitative data included one user of Basque (ST3), seven of Spanish (ST1, ST4, ST6, ST7, ST8, ST9 & ST10) and two of both languages (ST2 & ST5). Even though most of the interviewed participants spoke Spanish at home, they appeared to feel more anxiety than the other groups in both language classes:

> (ST2): "I feel well in Basque subjects, but when I speak and something doesn't come up, I feel a bit nervous, but not because Basque classes are taught in Basque. [...] In English, in worksheets or when I have to express something and I don't know, I feel nervous, but not because we're taught in English".
>
> (ST5): "I don't feel nervous because I know English since I was 5 years old, and Basque as well".
>
> (ST8): "I find Basque more difficult to learn. I feel more nervous in Basque because I feel like I don't master it, like more people speak it much better".
>
> (ST10): "I feel more nervous in English but less often in Basque or Spanish because I master them a bit more than English".

To analyse the influence of achievement, a series of Spearman correlation coefficients were calculated to establish to what extent global anxiety and factors in EMI and BMI were associated with course grades in subjects taught in English (i.e. English, art English and science) and Basque respectively, as displayed in table 7. Results showed a significant negative and moderate correlation between anxiety in EMI and grades in English (r_s = -.438, p-value = .000), suggesting that students who reported higher global anxiety in EMI classes had lower grades in English than those reporting lower anxiety. The correlation between factor 2 in EMI and grades in English was significantly negative as well, but the strength of the relationship was weak (r_s = -.314, p-value = .003), which indicated that those who experienced more anxiety when facing difficulties in the English subjects were slightly inclined to receive lower course grades in the FL.

With regard to art English grades, there was a statistically significant, albeit weak, negative correlation between factor 2 in EMI and achievement (r_s = -.248, p-value = .021). In other words, students receiving a lower grade in art English may have tended to feel more anxious when they perceived difficulties in their subjects taught in English.

Along with art English, science is another subject that our participants learnt in the FL, and statistically significant differences were observed in two categories. Global anxiety (r_s = -.314, p-value = .003) and factor 2 (r_s = -.355, p-value = .001) in EMI were negatively correlated with grades in science, that is, those being generally more anxious

and becoming anxious when they face difficulties in the classroom may be expected to score a lower grade in science, but the strength of the relationship was weak in the two cases.

To determine the impact of Basque grades on anxiety in BMI, the same procedure was followed. Anxiety in factor 2 was negatively correlated with Basque grades, despite sharing a weak relationship (r_s = -.285, p-value = .007). That is to say, learners who achieved a lower grade in Basque were potentially more prone to feel anxiety when they had difficulties in learning Basque subjects.

Table 7: Spearman correlations between grades in English and CLIL subjects and categories in EMI, plus between Basque grades and categories in BMI

			English grade	Art English grade	Science grade	Basque grade
Anxiety	EMI	r_s	-.438**	-.095	-.314**	
		Sig.	.000	.381	.003	
	BMI	r_s				-.202
		Sig.				.058
F1	EMI	r_s	-.203	.007	-.212	
		Sig.	.056	.945	.047	
	BMI	r_s				-.131
		Sig.				.221
F2	EMI	r_s	-.314**	-.248*	-.355**	
		Sig.	.003	.021	.001	
	BMI	r_s				-.285**
		Sig.				.007

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Several students' comments supported quantitative data, since they exhibited higher levels of anxiety in the language of instruction in which they performed worse, but others did not follow the same pattern:

> (ST4) (English: 'very good'; Basque: 'excellent'): "I find English subjects more difficult because the language changes and mastering it is difficult".

(ST5) (English: 'excellent'; Basque: 'very good'): "I think that science in English is much easier for me than in Basque because we've learned them more in English than in Basque".

(ST9) (English: 'very good'; Basque: 'good'): "I find science in English a little bit more difficult, because it's taught in English. In English I find words more difficult because I've never heard them before. After all, what we learn is Basque and Spanish".

6. Discussion

The purpose of this study was twofold. First, to explore the impact of CLIL on primary education students' anxiety, while levels of anxiety in EMI were contrasted with those in BMI, and second, to establish to what extent anxiety fluctuates in each language of instruction on account of gender, home languages and course grades.

Regarding the first of the research questions, students' anxiety episodes were not excessive; in fact, they experienced low levels of anxiety overall. Feeling a small degree of anxiety in primary education may be linked to the lower academic demands in comparison to other educational stages such as secondary or tertiary education settings. Particularly, the transition from primary to secondary education could be recognised as a shift into a stronger literacy orientation, as well as a period of stress and worry, which may affect learners' academic achievement and affective stance (De Smet et al., 2018; Muñoz, 2017; Somers & Llinares, 2021). In the interviews, most of them were conscious of the gradually increasing workload and complexity of their subjects, which may lead to an increase of anxiety as they climb the educational ladder. In addition to academic demands, fear of ambiguity has been previously identified in CLIL contexts (Papaja, 2019; Somers & Linares, 2021; Thompson & Sylvén, 2015), and is understood as the way individuals process ambiguous information and confront unfamiliar contexts, which may contribute to feeling nervous and tense. The present qualitative data underpin this observation, as participants explained that their level of anxiety would increase in the subsequent academic courses because they would be dealing with new content that had not been learnt previously.

Nonetheless, a contradictory result was found in participants' reports, as most of them agreed that their anxiety level in English subjects decreased since the onset of CLIL-type instruction. While some expressed their familiarity with the language or subject contents, a few others also mentioned overcoming their fear of ambiguity in CLIL classes, which may have played a facilitative role in reducing their anxiety levels over time. Doiz et al. (2014) explained that after a few years of CLIL-type exposure students' anxiety may be alleviated in view of their increased familiarity with the FL.

When students get familiar with CLIL, Papaja (2019) argued that their initial fear of ambiguity dispels and thus, their anxiety levels tend to decrease. This inconsistency in the direction in which their anxiety did and would fluctuate may also be in line with the recently contemplated Dynamic Approach, which advocates the dynamic conceptualisation of language anxiety. Indeed, anxiety is an emotion that varies from person-to-person and within the individual over time, as it is in constant interaction with a multitude of ever-changing learner and situational variables (Daubney et al., 2017). Therefore, change is inherent in language anxiety.

Furthermore, some of our participants reported having a positive experience in CLIL subjects at the moment and were perfectly aware of its benefits; namely, for increasing the frequency of FL use, improving FL linguistic skills and for travelling purposes among others. In agreement with Lasagabaster (2009), it seems that CLIL fosters positive attitudes, and recognising the international status of English, as pointed out by Teimouri et al. (2019), may have a facilitative role in reducing learners' anxiety in the FL.

When it comes to the language of instruction, despite their low level of global anxiety in both classes, students felt significantly more anxious in their subjects taught in English than in Basque. Given that the effect size was moderate, it can be stated that the language of instruction plays a decisive role in students' anxiety levels. In addition, mean scores of both anxiety factors were statistically higher in EMI, that is, students were more prone to anxiety when performing in the FL and facing difficulties in the EMI classroom. These results concur with two other studies conducted by Dewaele (2007) and Dewaele and Al-Saraj (2015); the former perceived that multilingual speakers' anxiety gradually increased over the languages learnt later in life, whereas the later researchers associated higher language anxiety levels with the lower frequency of use of that language. An equivalent finding was reported by Santos Berrondo (2017) in the BAC specifically, where L1-Spanish and L1-Basque university students showed having higher levels of communicative anxiety in their L3 (English) than in their L1 or L2 (Spanish or Basque). According to these authors, the exposure to the FL and frequency of FL use generally determine how anxious students may feel when using that language; in fact, given the limited opportunities to use English outside the classroom in the BAC, anxiety was significantly higher in EMI than in BMI subjects. In regard to the CLIL approach, a few researchers (Lasagabaster & Doiz, 2017; Papaja, 2019; Somers & Llinares, 2021) concluded that the process of learning subjects in the FL may be academically more demanding, resulting in an increase of anxiety among students. In fact, our participants attributed their higher levels of anxiety in English classes to their lower competence and poorer performance in the FL, greater complexity of the FL and lower frequency of FL use, whereas their mastery and frequency of Basque use was higher.

Contrary to expectations, however, some participants recognised feeling more anxious in BMI than in EMI. Dewaele (2007) argued that in stressful situations learners may tend to experience higher anxiety levels even in their L1. Some of our participants explained that their command and interests in English were greater, while in Basque subjects, they felt more uncomfortable and perceived Basque as a more difficult language. Probably, the meaningful exposure to the FL provided by CLIL and the out-of-school exposure to English may have alleviated their anxiety, since more than half of the sample were learning English in a language school, and some had travelled abroad as well. The facts that primary school students started learning English at the age of 3 at school and had been learning CLIL subjects since first grade in compulsory education seem to have increased their familiarity with English, as explained by a few participants and evidenced by Doiz et al. (2014).

According to our participants, the most common anxiety-provoking situation was when giving presentations, especially in English. This finding aligns with experts (Horwitz et al., 1986; Papaja, 2019; Young, 1990), who sustain that performing orally in a language in which learners' mastery is limited, is probably one of the most frequently cited concerns of anxious FL learners. Our participants' interviews reinforced the quantitative results and further confirmed prior research, in that students pointed out experiencing anxiety in oral presentations or when speaking in both language classes, since they lacked the knowledge to answer or express themselves, feared making mistakes, and were reluctant to volunteer. However, instances of anxiety were not exclusive to speaking. They also felt anxious when performing certain tasks in class, such as individual assignments and exams, and were worried about a bad performance, since learners tend to experience anxiety when performing in highly evaluative situations in the classroom (Horwitz et al., 1986).

Further sources of anxiety were connected with peers, as participants acknowledged feeling anxious due to classmates' stares, laughter, distractive noises and negative comments. In fact, peers have been argued to play a determinant role in language learning (Horwitz et al., 1986; Young, 1990), and many learners seem to be highly sensitive to their reactions, especially when their mistakes are in the spotlight. Considering that some participants also described feeling anxious due to the number of peers in the classroom and preferred learning in smaller groups, MacIntyre and Dewaele (2014) held that having fewer students in the FL class may be conducive to closer social bonds and a more positive learning atmosphere, and this could, in turn, contribute to reducing anxiety levels.

With respect to COVID-19, the pandemic has exponentially impacted peoples' mental well-being worldwide and sparked adverse psychological consequences, including stress, anxiety and fear (Oducado et al., 2021; Pizarro-Ruiz & Ordóñez-Camblor, 2021).

In this primary school, learners experienced moderate levels of anxiety in relation to wearing masks, having limited activities, not sharing material and keeping safety distance with friends in class. Such anxiety levels simulate results in prior research among graduate and university students (Oducado et al., 2021) as well as among children and adolescents (Pizarro-Ruiz & Ordóñez-Camblor, 2021), albeit during the confinement. To our knowledge, no empirical data have been reported indicating the influence of post-confinement school measures in students' levels of anxiety, which should be targeted in future research. Nonetheless, the target items composing the COVID-19 category in the questionnaire did not show internal consistency, and this may be associated with the fact that most students in the interviews pointed out that the current health protocols did not affect their degree of anxiety, but made them feel discouraged and frustrated instead.

To address the second research question, the degree of anxiety appeared to vary significantly on the basis of the independent variables under study. The differential role of gender reached no statistical significance, excepting in COVID-19 anxiety, but the effect of gender was small, and nor did learners differ in terms of gender in the interviews. To some extent females felt more anxiety than males when dealing with the health measures imposed by the school. The small effect size of gender in the anxiety caused by COVID-19 was also observed in a study conducted among Spanish teenagers (Pizarro-Ruiz & Ordóñez-Camblor, 2021), albeit during the quarantine. The lack of significant gender-based differences in the other cohorts parallels previous data (Aida, 1994; Dewaele & Al-Saraj, 2015; Elkhafaifi, 2005). For that reason, the vanishing effects of CLIL on gender-related differences found in motivation in the Spanish context (Fernández Fontecha & Canga Alonso, 2014; Heras Aizpurua, 2016; Heras & Lasagabaster, 2015) concur with our results concerning anxiety, although there is a scarcity of research in this field.

As for the anxiety differences attributed to the linguistic repertoire of learners, statistical tests indicated that students speaking Spanish at home experienced low, but significantly higher levels of anxiety than the users of Basque or of both Basque and Spanish in EMI exclusively. In addition, Spanish speakers were the ones coping with the greatest difficulties in both EMI and BMI classes. In the qualitative analyses, users of Spanish also expressed feeling more anxious than their peers from the other groups in both language classes. The moderate effect of learners' home languages suggests that the languages spoken in their daily settings may bear relevance in the anxiety experienced in the EMI classroom and when facing difficulties in the two languages of instruction.

These findings partially agree with Santos Berrondo's (2017) study, in which L1-Spanish undergraduates experienced more anxiety in Basque than their L1-Basque counterparts in Spanish, probably resulting from the minority status of Basque and

the fewer opportunities to use Basque outside school. In our sample, Spanish speakers' lower mastery of Basque and higher difficulties in BMI may be thus linked to their lower exposure to the language in comparison with those using Basque or both languages at home. In the contrast between EMI and BMI, more anxiety was observed in the subjects taught in the FL, consistent with Santos Berrondo (2017). This author did not find significant differences in relation to their L1s in English; participants were nonetheless mainly university students, who had received same exposure to the FL in the school curriculum and came from similar academic backgrounds. On the contrary, in our sample, those speaking Spanish at home exhibited significantly higher anxiety levels than their peers in EMI as well. This finding agrees with Dewaele and Al-Saraj (2015) as knowing more languages may reduce learners' anxiety across languages, and also with the study conducted by De Smet et al. (2018), in which lower anxiety levels were found among bilingual speakers compared to monolingual speakers. This could be explained by the advantages of bilingualism in L3 learning (Cenoz, 2011); indeed, those who speak only Basque or Basque and Spanish at home have more opportunities to use Spanish outside the school in the BAC than speakers of Spanish to use Basque, considering the dominant position of Spanish and the minority status of Basque. According to Cenoz (2011) and Lasagabaster (1998), having a broader linguistic repertoire at learners' disposal may be beneficial when acquiring an additional language in terms of language proficiency, metalinguistic awareness and learning strategies among others, although no positive effects of bilingualism on L3 anxiety were mentioned. This finding should be thus addressed in future research.

When examining the degree of association between anxiety and course grades, results should be interpreted with caution as our sample includes few participants achieving very low or high grades in both instruction types. In BMI classes, scoring lower grades in Basque was somehow associated with facing difficulties in the Basque subjects, and the same trend was found in EMI. Moreover, students' grades in English and the two CLIL subjects were negatively correlated with the complexity of the English classes, that is, students with lower grades in the three English classes were in some way more prone to encounter difficulties in these subjects.

As for qualitative data, the relation between grades and anxiety levels was substantially evidenced, because many students pointed out feeling more anxious in the language in which they achieved a lower grade. Accordingly, numerous SLA scholars (Aida, 1994; Elkhafaifi, 2005; Horwitz et al., 1986; MacIntyre & Gardner, 1991; Philips, 1992; Saito et al., 1999) commonly agreed on the negative impact of anxiety on achievement, and the magnitude of the negative correlations ranged from weak to moderate, as shown by prior research (Teimouri et al., 2019). Therefore, instead of cause-effects relationships between anxiety and achievement, it can be assumed that

low-achieving students revealed certain tendencies towards higher anxiety levels. In contrast to art English, the stronger correlations between English and science grades and anxiety categories may be ascribed to the more demanding learning and heavier academic load of these subjects. In fact, many students considered learning science in the FL to be more challenging, since they struggled to understand and learn specific terminology and lessons, whereas difficulties in art English were scarcely reported, because it was mainly focused on painting.

7. Conclusion

In sum, this research project encompasses the interaction between anxiety and language of instruction in the BAC. After exploring the case of 89 CLIL learners, it can be concluded that primary school students experience some episodes of anxiety at the initial stages of language learning. However, there was a great variability, and personal anxiety experiences differed considerably from learner to learner in view of the dynamic nature of anxiety and of the multiple sources that may lead to this negative emotion.

In a context in which the FL is hardly used beyond the school walls, the meaningful environment provided by CLIL may have increased learners' familiarity with English, and thus reduced their FL anxiety levels. Nevertheless, learners still felt substantially more anxiety in their subjects taught in English than in Basque, and specifically when giving oral presentations.

As far as the targeted variables are concerned, there were not significant differences based on gender, except for COVID-19 anxiety, in which females turned out to be substantially more anxious. Regarding participants' home languages, users of Spanish exhibited the highest levels of anxiety in EMI, and faced more difficulties in both language classes. Lastly, the relation between anxiety and achievement was significant in several categories in EMI. General anxiety in EMI was linked to scoring lower grades in English and science, and those students facing difficulties in EMI were associated with achieving lower grades in English and both CLIL subjects.

Any research suffers from specific limitations, which should be acknowledged. First, the anxiety category related to COVID-19 did not completely succeed in illustrating participants' emotions provoked by the pandemic situation, due to the low internal consistency of the questionnaire items, as well as the lack of participants' reports about feeling anxious with the current protocols in the interviews. The second limitation alludes to the data collection procedure, as participants' questionnaire responses were gathered by the school's teachers, and interviews were carried out through an online platform. Probably, attending in person would have been convenient for coming close

to the students and clearing up immediate confusions. Thirdly, while the achievement in the FL and CLIL subjects has been targeted, achievement in the L1 or L2 has been represented exclusively by means of grades in Basque, disregarding the difficulties and anxious feelings with which learners may deal in other subjects taught in Basque.

Further research on FLA should consider a longitudinal perspective, since it would offer a meaningful insight into anxiety fluctuations from the initial stages of language learning to a more advanced phase of the process. Another field worth exploring in future research within learners' linguistic repertoire will include the potentially positive effects of bilingualism on L3 anxiety. Also, considering the dearth of empirical data on the interface between gender, anxiety and CLIL, this is an issue that needs to be further addressed. Finally, not only learners but also teachers could be interviewed in future investigations, as teachers may identify possible signs of anxiety and share their strategies employed to reduce students' anxiety levels. Their contributions could be brought together with students' reports in a scheme of strategies from which the learners with a higher propensity to anxiety could benefit.

A number of pedagogical implications can be drawn from these findings. It is essential that teachers become aware of the role of anxiety and help students cope with anxiety-provoking situations. Particularly in tasks involving evaluation, they should provide positive reinforcement and allay students' fears of making mistakes or performing badly so that they understand that mistakes are part of the learning process.

Classmates appeared to be a key element in the study, since their negative comments, stares, laughter and distractive noises, among others, caused discomfort among some participants. In this regard, organising small groups may be helpful because students could perform exclusively in front of the members of their group, who will provide constructive feedback rather than negative comments. As a result, group dynamics may create an adequate learning environment, and learners may perceive peer support and gain confidence to engage in classroom interactions.

If oral tasks are made in front of the whole class, the teacher needs to carry out previous work to avoid any comment or gesture that could lead to anxiety-provoking situations. Students must be aware that empathy and respect are key values and that some of their classmates may feel a high degree of anxiety due to lower self-confidence and personality traits, which is why teachers should be very strict when signs of lack of respect are found in class. This is not always taken into consideration by some teachers, but our results clearly indicate that this question should be paid heed to.

With reference to the methods employed, developing dynamic and engaging activities (such as 'spelling bee', i.e. a spelling competition mentioned by two participants) may not only attract their attention and strengthen their motivation, but it may also contribute to reducing their anxiety experienced in the language classroom.

Acknowledgements

We wish to acknowledge the collaboration of Begoñazpi school and John's School academy in this research project as well as the cooperation of the teachers and counsellors who participated in the data collection process. This work was supported by the Spanish Ministry of Science and Innovation [grant number PID2020-117882GB-I00] and the Department of Education of the Basque Government (IT1426-22).

References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in English. *Educational Research and Evaluation*, *12*(1), 75-93.

Aida, Y. (1994). Examination of Horwitz, Horwitz and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*, *78*(2), 155-168.

Arnold, J. & Brown, H.D. (1999). A map of the terrain. In J. Arnold (ed.), *Affect in Language Learning* (pp. 1-24). Cambridge, UK: Cambridge University Press.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (ed.), *APA handbook of research methods in psychology: Vol. 2. Research designs* (pp. 57-91). Washington, DC: American Psychological Association.

Campbell, C.M., & Shaw, V.M. (1994). Language anxiety and gender differences in adult second language learners: Exploring the relationship. In C.A. Klee (ed.), *Faces in a crowd: The individual learner in multisection courses* (pp. 215-243). Boston: Heinle & Heinle Publishers.

Cenoz, J. (2009). *Towards multilingual education: Basque educational research from an international perspective*. Bristol: Multilingual Matters.

Cenoz, J. (2011). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching, 46*(1), 71-86.

Coyle, D., Hood, P., & Marsh, D. (2010). *Content and language integrated learning*. Cambridge, UK: Cambridge University Press.

Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, *31*, 182-204.

Daubney, M., Dewaele, J.M., & Gkonou, C. (2017). Introduction. In C. Gkonou, M. Daubney, & J.M. Dewaele (eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 1-7). Bristol: Multilingual Matters.

De Smet, A., Mettewie, L., Galand, B., Hiligsmann, P., & Van Mensel, L. (2018). Classroom anxiety and enjoyment in CLIL and non-CLIL: Does the target language matter? *Studies in Second Language Learning and Teaching*, *8*(1), 47-71.

Decreto 236/2015, de 22 de diciembre, por el que se establece el currículo de Educación Básica y se implanta en la Comunidad Autónoma del País Vasco. *Boletín Oficial del País Vasco, 9*, 15 January 2016, pp. 1-279. Retrieved from https://www.euskadi.eus/bopv2/datos/2016/01/1600141a.pdf.

Dewaele, J.M. (2007). The effect of multilingualism, sociobiographical, and situational factors on communicative anxiety and foreign language anxiety of mature language learners. *International Journal of Bilingualism*, *11*(4), 391-409.

Dewaele, J.M. (2017). Are perfectionists more anxious foreign language learners and users? In C. Gkonou, M. Daubney, & J.M. Dewaele (eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 70-90). Bristol: Multilingual Matters.

Dewaele, J.M., & Al-Saraj, T.M. (2015). Foreign language classroom anxiety of Arab learners of English: The effect of personality, linguistic and sociobiographical variables. *Studies in Second Language Learning and Teaching*, *5*(2), 205-228.

Dewaele, J.M., & MacIntyre, P.D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*(2), 237-274.

Dewaele, J.M., MacIntyre, P.D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition*, *2*(1), 41-63.

Dewaele, J.M., & Shan Ip, T. (2013). The link between Foreign Language Classroom Anxiety, Second Language Tolerance of Ambiguity and Self-rated English proficiency among Chinese learners. *Studies in Second Language Learning and Teaching*, *3*(1), 47-66.

Dewaele, J.M., & Thirtle, H. (2009).Why do some young learners drop Foreign Languages? A focus on learner-internal variables. *International Journal of Bilingual Education and Bilingualism*, *12*(6), 635-649.

Dörnyei, Z. (2010). *Questionnaires in Second Language Research. Construction, Administration, and Processing* (2nd ed.). New York: Routledge.

Doiz, A., Lasagabaster, D & Sierra, J.M. (2014). CLIL and motivation: the effect of individual and contextual variables. *The Language Learning Journal*, *42*(2), 209-224.

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal*, *89*(2), 206-220.

European Commission. (2012). *Europeans and their languages. Special Eurobarometer 386*. Report. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.

Eurydice. (2006). *Content and Language Integrated Learning (CLIL) at School in Europe*. Brussels: European Commission.

Fernández Fontecha, A., & Canga Alonso, A. (2014). A preliminary study on motivation and gender in CLIL and non- CLIL types of instruct. *IJES*, *14*(1), 21-36.

García Mayo, M.P. (2003). Age, Length of Exposure and Grammaticality Judgements in the Acquisition of English as a Foreign Language. In M.P. García Mayo, & M.L. Lecumberri (eds.), *Age and the Acquisition of English as a Foreign Language.* (pp. 94-114) Clevedon, UK: Multilingual Matters.

Heras Aizpurua, A. (2016). *The impact of CLIL: Affective factors, content-related vocabulary & gender differences*. Unpublished dissertation. University of the Basque Country.

Heras, A. & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, *19*(1), 70-88.

Horwitz, E.K., Horwitz, M.B., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal*, *70*(2), 125-132.

Kitano, K. (2001). Anxiety in the College Japanese Language Classroom. *The Modern Language Journal*, *85*(4), 549-566.

Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York: Routledge.

Lasagabaster, D. (1998). The threshold hypothesis applied to three languages in contact at school. *International Journal of Bilingual Education and Bilingualism*, *39*, 119-133.

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *The Open Applied Linguistics Journal*, *1*(1), 30-41.

Lasagabaster, D. (2009). The implementation of CLIL and attitudes towards trilingualism. *ITL - International Journal of Applied Linguistics*, *157*(1), 23-43.

Lasagabaster, D. & Doiz, A. (2017). A Longitudinal Study on the Impact of CLIL on Affective Factors. *Applied Linguistics, 38*(5), 688-712.

MacIntyre, P.D., Baker, S.C., Clément, R., & Donovan, L.A. (2002). Sex and age effects on willingness to communicate, anxiety, perceived competence, and L2

motivation among junior high school French immersion students. *Language Learning*, *52*(3), 537-564.

MacIntyre, P.D., & Gardner, R.C. (1991). Investigating language class anxiety using the focused essay technique. *The Modern Language Journal*, *75*(3), 296-304.

Muñoz, C. (2002). Relevance and potential of CLIL. In D. Marsh (ed) *CLIL/EMILE - The European dimension: Actions, trends and foresight potential.* (pp. 35-36) Jyväskylä: University of Jyväskylä, Continuing Education Centre.

Muñoz, C. (2017). Tracing Trajectories of Young Learners: Ten Years of School English Learning. *Annual Review of Applied Linguistics*, *37*, 168-184.

Odlin, T. (1989). *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

Oducado, R.M.F., Parreño-Lachica, G.M., & Rabacal, J.S. (2021). Personal resilience and its influence on COVID-19 stress, anxiety and fear among graduate students in the Philippines. *International Journal of Educational Research and Innovation*, *15*, 431-443.

Onwuelgbuzie, A.J., Bailey, P., & Daley, C.E. (1999). Factors associated with foreign language anxiety. *Applied Psycholinguistics*, 20(2), 217-239.

Papaja, K. (2019). To fear or not to fear CLIL: Some remarks on the role of anxiety in a CLIL classroom. *Konińskie Studia Językowe*, 7(2), 171-196.

Park, G.P., & French, B.F. (2013). Gender differences in the foreign language classroom anxiety scale. *System*, *41*, 462-471.

Pérez Cañado, M.L. (2018a). CLIL and Educational Level: A Longitudinal Study on the Impact of CLIL on Language Outcomes. *Porta Linguarum*, *29*, 51-70.

Pérez Cañado, M.L. (2018b). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction*, *57*, 18-33.

Philips, E.M. (1992). The Effects of Language Anxiety on Students' Oral Test Performance and Attitudes. *The Modern Language Journal*, *76*(1), 14-26.

Pizarro-Ruiz, J.P., & Ordóñez-Camblor, N. (2021). Effects of COVID-19 confinement on the mental health of children and adolescents in Spain. *Scientific Reports*, *11*(11713), 1-10.

Pladevall-Ballester, E. (2019). A longitudinal study of primary school EFL learning motivation in CLIL and non-CLIL settings. *Language Teaching Research*, *23*(6), 765-786.

Ruiz de Zarobe, Y. (2008). CLIL and Foreign Language Learning: A Longitudinal Study in the Basque Country. *International CLIL Research Journal*, *1*(1), 60-73.

Ruiz de Zarobe, Y. & Lasagabaster, D. (2010). CLIL in a Bilingual Community: The Basque Autonomous Community. In D. Lasagabaster & Y. Ruiz de Zarobe (eds.), *CLIL in Spain: Implementation, results and teacher* training (pp. 12-29). Newcastle: Cambridge Scholars Publishing.

Saito, Y., Horwitz, E.K., & Garza, T.J. (1999). Foreign language reading anxiety. *The Modern Language Journal*, *83*(2), 202-218.

Santos Berrondo, A. (2017). *Anxiety in second and third languages: The case of adult multilinguals from the Basque Autonomous Community*. Unpublished dissertation. University of the Basque Country.

Seikkula-Leino, J. (2007). CLIL Learning: Achievement Levels and Affective Factors. *Language and Education*, *21*(4), 328-341.

Simons, M., Vanhees, C., Smits, T., & Van De Putte, K. (2019). Remedying Foreign Language Anxiety through CLIL? A mixed-methods study with pupils, teachers and parents. *Revista de lingüística y lenguas aplicadas*, *14*, 153-172.

Somers, T., & Llinares, A. (2021). Students' motivation for content and language integrated learning and the role of programme intensity. *International Journal of Bilingual Education and Bilingualism*, *24*(6), 839-854.

Teimouri, Y., Goetze, J., & Plonsky L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387.

Thompson, A.S., & Sylvén, L.K. (2015). "Does English make you nervous?" Anxiety profiles of CLIL and non-CLIL students in Sweden. *Apples - Journal of Applied Language Studies*, *9*(2), 1-23.

Young, D.J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals*, *23*(6), 539-553.

Appendix: FLA Questionnaire

In the following section, I would like you to ***circle*** the degree to which you agree or disagree with the following statements:

1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree, 0 = Not sure

Anxiety in the English-medium classes						
1. It embarrasses me to volunteer in my subjects given in English because I feel that there are too many eyes observing me.	1	2	3	4	5	0
2. I worry when there are too many students in my subjects learnt in English.	1	2	3	4	5	0
3. I feel nervous when the teacher asks me questions which I haven't prepared in advance in my English-medium classes.	1	2	3	4	5	0
4. I feel nervous when I can't write or express myself in English.	1	2	3	4	5	0
5. It **wouldn't** bother me at all to take more classes in English.	1	2	3	4	5	0
6. I keep thinking that the other students are better at English than I am.	1	2	3	4	5	0
7. I **don't** feel anxious when I speak English in front of the class.	1	2	3	4	5	0
8. I worry about how demanding the subjects taught in English are.	1	2	3	4	5	0
9. I am afraid that the other students will laugh at me when I speak English.	1	2	3	4	5	0
10. I feel nervous when standing to give a presentation in English in front of the class.	1	2	3	4	5	0
11. It worries me not understanding the contents of the English subjects.	1	2	3	4	5	0
12. Although I am well prepared for the English subjects, I feel anxious about it.	1	2	3	4	5	0
13. I **don't** feel comfortable with my classmates in the subjects given in English.	1	2	3	4	5	0
14. I **don't** worry about making mistakes in the English-medium classes.	1	2	3	4	5	0
15. I find it hard to study these subjects in English, because they are taught in English.	1	2	3	4	5	0
16. I am usually relaxed during my lessons taught in English.	1	2	3	4	5	0

Anxiety in the Basque-medium classes						
17. I get nervous when I don't understand the contents in the Basque-medium subjects.	1	2	3	4	5	0
18. I start to panic when the teacher asks me to speak without preparation in my Basque-medium classes.	1	2	3	4	5	0
19. I **don't** get nervous when I give oral presentations in Basque.	1	2	3	4	5	0
20. I **don't** like volunteering in my subjects learnt in Basque because I am a shy person.	1	2	3	4	5	0
21. I **don't** worry about the group size in my subjects given in Basque.	1	2	3	4	5	0
22. Even if I prepare the Basque-medium classes in advance, I feel nervous.	1	2	3	4	5	0
23. I have difficulties in learning Basque subjects, because they are taught in Basque.	1	2	3	4	5	0
24. It **doesn't** embarrass me to make mistakes in my subjects learnt in Basque.	1	2	3	4	5	0
25. I feel uncomfortable with my classmates in my subjects given in Basque.	1	2	3	4	5	0
26. I feel very confident when I have to speak in my Basque-medium classes.	1	2	3	4	5	0
27. I feel that my subjects taught in Basque **aren't** very demanding.	1	2	3	4	5	0
28. I get nervous when I can't express what I want in Basque.	1	2	3	4	5	0
29. It worries me giving less lessons in Basque.	1	2	3	4	5	0
30. I always feel that the other students know Basque better than I do.	1	2	3	4	5	0
31. I usually feel nervous in my subjects learnt in Basque.	1	2	3	4	5	0
32. I don't care if my classmates laugh at me when I speak in my Basque classes.	1	2	3	4	5	0

Anxiety related to COVID-19 measures						
33. I get upset when they don't understand me while I am speaking because of the mask.	1	2	3	4	5	0
34. I get upset because the current health protocol does not allow us to do many activities or play certain games.	1	2	3	4	5	0
35. It bothers me that we can't share material in class.	1	2	3	4	5	0
36. Not sitting next to my friend, due to the safety distance, makes me feel anxious.	1	2	3	4	5	0

37. I would rate my level of overall anxiety in my English-medium classes:

extremely high high moderate low extremely low

38. I would rate my level of overall anxiety in my Basque-medium classes:

extremely high high moderate low extremely low

Thank you very much for your participation!! 😊

Collaborative writing and patterns of interaction in young learners: The interplay between pair dynamics and pairing method in LRE production

María Basterrechea
Departamento de Filología Inglesa
Universidad del País Vasco UPV/EHU
maria.basterrechea@ehu.eus

Francisco Gallardo-del-Puerto
Departamento de Filología
Universidad de Cantabria
francisco.gallardo@unican.es

Abstract

A considerable body of research within the Socio-cultural theory (Lantolf & Appel, 1994) examines how learners express their linguistic gaps verbally, or question their own or others' language use when writing collaboratively, i.e., produce Language-related episodes (LREs; Swain & Lapkin, 1998). Several studies have also explored the effect that different *patterns of interaction* (Storch, 2002) have on the production of LREs with adult learners (e.g., Mozaffari, 2017; Storch & Aldosari 2013), but little research has compared the effect of these patterns of interaction and pair formation method (i.e., student-selected and proficiency-matched) on young EFL learners' ability to attend to language, and much less on the *type* of grammatical features they focus on in LREs.

This study examines young EFL learners' (aged 10-12) production of LREs and pair dynamics in student-selected vs. proficiency-matched groups while completing a collaborative writing task. It was found that young EFL learners mainly exhibit a collaborative type of dynamics and resolved more LREs accurately, together with expert-novice groups. Matched proficiency was more beneficial, as these groups produced more target-like LREs. As per the type of form-focused LREs produced, these young learners focused primarily on spelling issues and less on grammatical knowledge-induced ones.

Keywords: Language-related episodes (LRE), patterns of interaction, pair formation method, collaborative writing, form-focused LREs.

Resumen

Un número considerable de estudios enmarcados en la Teoría Sociocultural (Lantolf y Appel, 1994) investiga el modo en el que los/as aprendices expresan verbalmente sus lagunas lingüísticas, o cuestionan el uso propio o ajeno del lenguaje durante la escritura colaborativa, es decir, producen Episodios Relacionados con el Lenguaje (ERLs; Swain y Lapkin, 1998).Varios estudios han examinado, además, el efecto que los *patrones de interacción* (Storch, 2002) ejercen en la producción de los ERLs con aprendices adultos (véase Mozaffari, 2017; Storch y Aldosari, 2013), aunque pocos estudios han comparado el efecto de dichos patrones de interacción y el método de emparejamiento (es decir, formación de parejas en base a las preferencias de los/as propios/as aprendices y en base a la proficiencia) en la capacidad de los/las aprendices jóvenes de inglés como lengua extranjera (ILE) para atender a la lengua, y menos aún en el *tipo* de rasgos gramaticales en los que éstos/as se enfocan en la producción de ERLs.

Este estudio examina la producción de los ERLs y los patrones de interacción en parejas formadas en base a sus preferencias o su proficiencia en jóvenes aprendices de ILE (10-12 años) durante una tarea colaborativa escrita. Los resultados demostraron que los jóvenes aprendices de ILE mantuvieron generalmente un tipo de dinámica colaborativa, quienes también resolvieron una mayor cantidad de ERLs de forma correcta, junto con las parejas experto/a-novato/a. En cuanto al tipo de emparejamiento, las parejas establecidas en base a su proficiencia obtuvieron resultados más beneficiosos en tanto en cuanto produjeron un mayor número de ERLs con resolución correcta. En cuanto al tipo de LREs enfocados a la forma, los/las participantes se enfocaron mayormente en aspectos relacionados con la ortografía y en menor medida en aspectos derivados de un conocimiento gramatical.

Palabras clave: Episodios Relacionados con el Lenguaje (ERLs), patrones de interacción, método de emparejamiento, escritura colaborativa, ERLs enfocados en la forma.

1. Introduction

The field of Second Language pedagogy is witnessing an increasing interest in young learners' development in second language (Mackey, 1994; Mackey & Oliver, 2002; Mackey & Silver, 2005; Oliver, 1998; Oliver, Philp & Duchesne, 2017; Roehr-Brackin & Tellier, 2019) and, more scarcely, in foreign language (Azkarai & Kopinska, 2020; Coyle & Roca de Larios, 2014; García Mayo & Hidalgo, 2017; García Mayo & Imaz Aguirre, 2019; García Mayo & Lázaro-Ibarrola, 2015; Pinter, 2006, 2007; Pladevall-Ballester & Vraciu, 2020) contexts. Particularly, there has been a growing

interest in the potential contribution that collaborative writing makes in fostering learners' reflection on language form (López-Serrano, Roca de Larios & Manchón, 2019). When writing collaboratively, learners express their linguistic gaps verbally, question their own or others' language use and resort to their internal linguistic knowledge about form and meaning (Swain & Lapkin, 2001). These processes were originally labelled as Language-related episodes (LREs) by Swain & Lapkin, (1998: 70), who defined them as "[...] any part of the dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct."

According to the Socio-cultural framework (Lantolf & Appel, 1994), the relationship that learners build when working collaboratively and their *patterns of interaction* (Storch, 2002) have been signalled as one of the factors that affect the volume of languaging (i.e., the production of LREs) in collaborative writing (see overview by Storch, 2016) with adult learners (e.g., Mozaffari, 2017; Storch & Aldosari 2013), but little research has compared the effect of these patterns of interaction on young EFL learners' ability to attend to language. Thus, the present study sought to analyse young EFL learners' production of LREs and how the nature of pair dynamics in task-based interaction affects these in their (i) incidence (quantity), (ii) nature (form- or meaning-focused), and (iii) resolution (target-like or non-target-like or non-resolved). It also explores the extent to which LREs are influenced by how learner pairs are established (matched proficiency or self-selection). Additionally, the study further analyses the type of language aspects involved in form-focused LREs comparing each type of pattern of interaction and of pairing method, as merely giving account of the amount and resolution of LREs falls short in the analysis of the extent to which young EFL learners explicitly attend to specific language features.

2. Literature Review

An increasing number of investigations in the Socio-cultural theory of learning examine how collaborative writing contributes to language development and knowledge building (Swain, 2000). The distinguishing traits of collaborative writing identified pertain to (i) *process -* as a collaborative construction of the authors in the composition process -, (ii) *product -* as a final result of a unique text. Thirdly, (iii) the concept of *text ownership* refers to the decision-making process of text production (Ede & Lunsford, 1990, as cited in Storch, 2016). Writing most typically occurs under planned conditions, which allows learners to address both content and form (Williams, 2008). In this respect, the off-line nature of the writing process is believed to facilitate learners' explicit reflection on language and to provide peer feedback, both of which may contribute to advance in the form-meaning relationship of the target language (Manchón & Williams, 2016). As Williams (2012: 328) put it, "[..] learners have a richer opportunity to test their hypotheses when they write than when they speak".

LREs are identified as the unit of analysis to code those instances in the data where learners deliberate about morphosyntax, lexis or mechanics (Storch, 2016), both in the Socio-cultural framework and the Interaction framework (Long, 1996). From a socio-cultural perspective, production of language is seen as a communicative and a cognitive activity, and LREs are viewed as instances where learners verbalize their thinking or their deliberations, processes termed as *languaging* (Swain, 2006, 2010).

Numerous studies on LREs have focused on adults in English as a second language (ESL; Benson, Pavitt, J.,& Jenkins 2005), immersion (Kowal & Swain, 1994; Swain, 1998; Swain & Lapkin, 1998), content-based instruction (Leeser, 2004), or foreign language settings (Basterrechea & García Mayo, 2013; Basterrechea & Leeser, 2019; Kim & McDonough, 2008; García Mayo, 2002; García Mayo & Azkarai, 2016; Malmqvist, 2005; Storch & Aldosari, 2013). Studies examining LRE production by young learners have grown steadily in the past few decades, analysing negotiation strategies (Oliver, 1998), task effect on attention to form (Mackey, 1994; Plonsky & Kim, 2016), and, to a lesser extent, LREs on written production, focussing in particular on task repetition (Hidalgo & García Mayo 2021), and on feedback (Coyle & Roca de Larios, 2014).

These studies examine not only the (i) incidence (quantity), (ii) nature (form- or meaning-focused), and (iii) resolution (target-like, non-target-like or non-resolved) of the LREs produced by the learners, but also the impact that the patterns of interaction exhibited by adult learners in collaborative tasks (e.g., Donato, 1998; Kim & McDonough, 2008, Storch, 2002; Storch & Aldosari, 2013), and, more recently, by young learners in ESL (Azkarai, García Mayo & Oliver, 2020; Oliver & Azkarai, 2019) and EFL (Azkarai & Kopinska, 2020; Butler & Zeng, 2015; García Mayo & Imaz Aguirre, 2019; Martinez-Adrián & Gutiérrez-Mangado, 2022) settings have on the production of LREs. Overall, research on patterns of interaction point to a moderating effect of pair behaviour on the incidence of LRE production in collaborative interaction. In that regard, in Storch's (2002) pioneering research on the effect of pair behaviour on languaging in a collaborative writing task with adult ESL learners, 4 patterns of pair relations were identified, based on the learners' level of engagement with the other member of the dyad (i.e., mutuality) and the level of contribution to the task (i.e., equality), as follows:

a) Collaborative pattern: both members of the dyad contribute to the task actively, by pooling their resources and incorporating and/or repeating each other's utterances and extending on them. Thus, learners' level of contribution to both the task and with the other member of the pair is high.

b) Dominant/dominant pattern: both participants contribute to the task but do not engage with each other's contributions and hence there is not a

joint contribution of the text. Learners may use disputational talk and show disagreement. This pattern may also include a cooperative (Storch, 2001; Tan, Wigglesworth & Storch, 2010) or passive/parallel (Butler & Zeng, 2015) pattern when there is no engagement among peers but participants do not attempt to take control over the task - also described as division of labour (Storch, 2002).

c) Dominant/passive pattern: Learners do not engage with each other's contribution, and one of the members takes control over the task; little assistance is sought or offered.

d) Expert/novice pattern: The expert or capable peer takes control over the task, but s/he seeks to involve the novice member, by providing assistance. Hence, the novice contributes to the task to a lesser extent, but a high level of engagement exists among the participants.

Results in the study by Storch (2002) showed that collaborative and expert-novice were the patterns that contributed to language gains more effectively, as attested by the higher amount of LREs transferred to subsequent individual tasks by pairs that exhibited these types of patterns. Subsequent studies that have examined the interplay between the relative proficiency of the dyad members (i.e., homogenous vs heterogeneous proficiency) and the patterns of interaction in adult ESL (e.g., Kim & McDonough, 2008) and EFL learners (e.g., Storch & Aldosari, 2013) have shown that homogeneous groups sustain more optimal patterns of interaction (i.e., collaborative, and expert/novice pattern) and produce a larger amount of LREs compared to dominant/passive dyads, although the type of role relationships among mixed proficiency dyads is not so conclusive, with learners exhibiting a wider range of patterns in mixed proficiency dyads (see Storch & Aldosari, 2013), or adopting less beneficial roles in heterogeneous groupings (see Kim & McDonough, 2008). As for young learners, García Mayo and Imaz Aguirre (2019) was the first study that examined pair dynamics in young EFL learners (11-12 years old). All learners exhibited a collaborative type of relationship in task-based interaction. However, in the study by Azkarai and Kopinska (2020) with the same type of population, learners exhibited not only a collaborative pattern, but equally a cooperative one in that a considerable amount of learners did not engage in each other's contribution and did not attempt to control over the task. The authors explained this finding in the light of the task demands employed in the study – a dictogloss task (Wajnryb, 1990) -, where one of the students took the role of the scribe, which may have resulted in a more passive attitude. Pladevall-Ballester (2021) also examined LRE production and patterns of interaction among young EFL learners and whether these change over time (or as proficiency increases), analysing the pair dynamics by 4th and 6th graders (aged 10 and 12 respectively) in an oral task in this

case. As attested in previous studies with EFL schoolchildren, learners were mainly collaborative, at the two testing times. As per type of LREs, the majority had a lexical focus, which also corroborates findings in prior young learner literature (Basterrechea & Gallardo-del-Puerto, 2020; Gallardo-del-Puerto & Basterrechea, 2021; García Mayo & Imaz Aguirre, 2019); similar results were obtained in studies that compared high and low proficiency adult learners in ESL (e.g., Leeser, 2004) and EFL (Basterrechea & Leeser, 2019; Kim & McDonough, 2008; Malmqvist, 2005; Storch & Aldosari, 2013) contexts, with low proficient learners producing mainly meaning-focused LREs. As for resolution, it was the expert-novice dyads who produced the largest amount of resolved LREs in Pladevall-Ballester's (2021), findings that corroborate the benefits of this type of pattern attested in earlier studies (e.g., Oliver & Azkarai, 2019; Storch, 2002).

Apart from proficiency, research with L2 learners has explored the impact of other variables on the patterns of interaction, such as age (Butler & Zeng, 2015) or task type (Ahmadian & Tajabadi, 2017), with collaborative patterns resulting in the most beneficial type of dynamics in promoting a more collaborative construction of knowledge overall. However, there is a dearth of studies exploring the extent to which patterns of interaction are influenced by how pairs are established, and if having learners choose their partners or matched proficiency would have a different impact on learners' ability to attend to language. Basterrechea and Gallardo-del-Puerto (2020) found a wider range of patterns of interaction in a study that examined the interplay between pair dynamics and pair formation method in young learners' LRE production, a variable that had previously been investigated in adult EFL contexts (e.g., Mozaffari, 2017) resulting in proficiency-matched pairs producing more LREs, whereas student-selected pairs talked about matters unrelated to the task more frequently, although Gallardo-del-Puerto and Basterrechea (2021) found that it was also the self-selected interactancts who produced more target-like meaning-focused LREs. In the study by Basterrechea and Gallardo-del-Puerto (2020), patterns of interaction of proficiency-matched versus self-selected pairs showed that the latter exhibited not only a wider range of patterns of interaction, but also the types of dynamics which are believed to have a detrimental effect on language development (i.e., dominant/dominant and dominant/passive). Proficiency-matched pairs, instead, were mainly of a collaborative and, to a lesser extent, an expert-novice pattern. This study, although preliminary, showed that patterns of interaction may be affected by factors other than proficiency level, as having learners choose their partners may play a role in the potential that peer collaboration has in task-based dynamics. Along these lines, it has been suggested that the "[...] learners' ability to profit from peer interaction is greatly affected by the social dynamics of their group or pair" (Sato & Ballinger, 2016, p. 19).

With the aim to fill this gap, the present study examines the potential interrelationship between pair dynamics and pairing method and how the different patterns of interaction

and pair formation procedures affect learners' ability to attend to language. Although some of the studies above have shown that collaborative is the most frequent pattern among young EFL learners, the results are not so clear-cut when adding pair formation method into the analysis. Additionally, in order to gain further insights into the language learning opportunities that arise in collaborative task performance, rather than focusing solely on the profits of producing and solving LREs collaboratively, a detailed analysis of the type of LREs -particularly form-focused - needs to be done. as "[..] merely counting LREs fail to capture the complexity of the interactions" (Storch, 2016: 397).

Previous research has attested that young learners produce LREs with a lexical focus to a larger extent than form-focused LREs. However, to our knowledge, no studies have reported on the *type* of grammatical features learners focus on in child EFL pair dynamics. A closer analysis of the nature of form-focused LREs will help to uncover the extent to which these learners are able to discuss language and/or have metalinguistic awareness, all of which may contribute to language development (Roehr-Brackin & Tellier, 2019). Hence, the present study incorporates the analysis of the specific language aspects involved in the form-focused LREs, in order to provide evidence of young EFL learners' ability to reflect on language use. Our study is a follow-up of Basterrechea and Gallardo-del-Puerto (2020), but now the focus is on the interface between pair dynamics and pair formation method. In order to rule out the task effect, in this study we have now centred on the data coming from a task only - a convergent map task (a writing task) -. Additionally, the present study incorporates the focus of the *type* of LREs (within meaning-focused or form-form focused), an issue which has only been explored in adult learners so far (Niu, 2009), and is yet to be looked into in young learners.

3. Research questions

On the basis of previous research, we entertain four different research questions:

1) Which patterns of interaction do young EFL learners exhibit when they engage in a collaborative writing task?

2) Does pairing method (matched proficiency vs self selection) per type of pattern have an effect on the incidence, type and resolution of LREs?

3) What type of form-focused LREs do these learners produce? What is the relationship between patterns of interaction and form-focused LRE types?

4) What is the relationship between pairing method per type of pattern and form-focused LRE types?

4. Materials and Methods

4.1. Participants

Fifty-seven (57) schoolchildren from 5 intact classrooms in their fifth and sixth year of Primary Education (aged 10 to 12) in the Basque Autonomous Community (northern Spain) participated in the study. The school is located in an area where Spanish is more frequently used than Basque, and the latter is generally learnt in the school context in a total immersion language model, with Spanish and English as school subjects. In order to increase the amount of exposure to the foreign language (English), the school programme incorporates Content and Language Integrated Learning (CLIL) lessons into the curriculum from 3rd grade onwards (age 8-9), whereby *Arts and Crafts, Physical Education* and *Science* are taught in English. Hence, at the time of data gathering, the participants' English exposure amounted to 777 hours in Grade 5 and 962 hours in Grade 6. As for their English proficiency, they are all considered beginner learners according to the Key English Test (KET, Cambridge University Press, 2008) which they took at the outset of the project.

4.2. Instruments and procedure

A general background questionnaire and the English proficiency test enabled us to assess participants' biographical profiles and English proficiency. The results in the latter were used to establish proficiency-matched pairs of half of the participants in the experimental phase, whereas the second half of the participants were asked to choose a partner they would work with. The children were grouped into 24 dyads (2-member groups) and 3 triads (3-member groups) (due to the uneven number of participants in 3 out of the 5 intact classrooms).

In the experimental phase, participants underwent a convergent map task (e.g., Gilabert, Barón & Llanes, 2009) in groups. This was a consensus task (Gass, Mackey & Ross-Feldman, 2005) with only one possible solution. Research has shown that convergent tasks lead to more negotiation of meaning and production of LREs than divergent tasks (e.g., Gilabert et al., 2009). Learners were asked to agree on an itinerary the main character of a story would have to follow and write a short text collaboratively (see Appendix). Participants were first provided with a town map and a picture showing a boy who has found a lost dog in a park. In a second phase, students had to agree on the itinerary the boy had to follow around various landmarks on the map. In a final stage, they had to write collaborately a short note for the boy explaining who the dog's owner is and giving directions from the park to the owner's workplace so that the boy would be able to take the dog back to its owner.

The learner groups were taken to a quiet room in their school premises in turns so that we could video-record their interaction while they accomplished the writing task. No time restrictions were imposed. The investigators stressed the importance of paying attention to language accuracy and encouraged learners to work on their own and to pool their own resources in the task.

4.3. Analysis

The participants' oral production was video-recorded and then transcribed using the CHAT conventions in CHILDES (Child Language Data Exchange System; MacWhinney, 2000). The patterns of interaction exhibited by each of the dyads (and triads) were analysed on the basis of the most representative pattern and classified according to the taxonomy proposed by Storch (2002) (collaborative, expert/novice, dominant/passive or dominant/dominant - see above). These analyses observed the learners' behaviour in requesting and/or providing feedback, explicit peer repairs, the degree of assistance sought or offered, repetition of requests, level of engagement in each other's contributions or disagreement (op. cit.). The following extracts illustrate the dyadic patterns found in our data. In example (1) participants engaged in the task collaboratively. In this excerpt, the students complete each other's utterances and pool their resources by providing suggestions about the preposition that should follow *go*. They recast each other's utterances until they come to an agreement – no preposition follows *go*. They exhibit a collaborative pattern, where they show high degree of participation, and the level of contribution to both the task and with the other member of the dyad is high.

Collaborative (high equality – high mutuality):

(1) *CHI2: ah go after to the park.
*CHI1: and the xxx go after the park. (she starts writing)
*CHI1: go?
*CHI2: out.
*CHI1: xxx. (CHI2 takes the pen and writes)
*CHI2: go out to the park.
*CHI1: go (.) to the park.
*CHI1: <*vale*>@s [OK]
*CHI2: <*vale*>@s [OK]

Extract (2) shows an example of an expert/novice pattern, where Child 1 takes the lead but also seeks to involve the novice by modeling the sentences that Child 2 incorporates into her interaction. Although Child 2 is not passive and asks for repetition, their degree of participation is unequal. However, a high level of engagement with each other's contribution exists among the participants.

Expert/novice (low equality – high mutuality):

(2) *CHI1: the boy have to pass for the main street. (..) the boy (.) have to pass (..) to the main street.

*CHI2: to the?

*CHI1: main street. (she points at the name of the street)

*CHI1: and then is the laboratory. (CHI2 writes)

Extract (3) features an example of a dominant/passive dyad. Child 1 is the dominant peer and Child 2 – the writer – a more passive one, as evinced by the unequal number of turns of Child 1, her long monologues (turns 1 to 3, 5 to 8, and 12) and her dominant attitude over the task (turns 3, 5, 6, 7, 8, and 11), commanding her partner to write what she thinks is correct, or erasing what Child 2 has written, or ignoring Child 2's suggestions (turn 11).

Dominant/passive (low equality – high mutuality):

(3) (1) *CHI1: <*a ver, tú pon* >@s [let's see, write] going to the church. (CHI2 starts writing) (..) to the.

(2) *CHI1: <no, a ver >@s [no, let's see] (.) into the (.)
<*a ver, cómo, a ver, a ver* >@s [no, let's see, how, let's see, let's see]
(CHI2 erases what she has written and CHI1 takes the paper to continue writing)

(3) *CHI1: to the (.) <*espera, espera*>@s [hold on, hold on]

(4) *CHI2: ah <*junto*>@s [next to]. (CHI1 erases something more)

(5) *CHI1: <*es que voy a academia*>@s [I take private lessons, you know].

(6) *CHI1: (writing) go (.) to (.) the (.) church.

(7) *CHI1: to church and right no right?

(8) *CHI1: reat <*algo así era, no*?>@s [it was something like that, right?]

(9) *CHI2: <*algo así era, no*>@s [it was something like that]

(10) *CHI1: and going reat <*igual era recto*>@s [perhaps it was straight]

(11) *CHI2: and go (.) <*y si ponemos* >@s [and if we write]
and go to the <*es que ya hemos puesto, no?*>@s [we have already written that, right?]

(12) *CHI1: going to the church <*o sea ir hasta la (.) hasta la iglesia y luego.* >@s [I mean go to the (.) to the church and then]

Extract (4), the three members of a triad show a desire to dominate the task – described in the situations coded as %sit -, as evinced by the little engagement with each other's contribution. The three of them want to take the initiative to the point that they sound authoritarian, as evinced in turns 3, 6, 8, 9 and 15 by Child 2, or %sit, where Child 3 takes the paper; Child 1 wants to stand out (see turns 4 and 14) as well, but the other members do not seem to pay attention to him.

Dominant/dominant:

(4) %sit: CHI1 takes the paper.

(1) *CHI1: in the park. (he writes)

(2) *CHI2: in the park.

(3) *CHI2: <*escribe; quién es su dueño*. >@s [write; who his owner is]

(4) *CHI1: Jack Smith <*ya lo hemos puesto*>@s [we have already done it] is Jack Smith.

(5) *CHI2: <*y, a ver.* >@s [and, let's see]

(6) *CHI2: <*pon*>@s [write down] Jack Smith is vet. (CHI1 writes)

(7) *CHI1: Smith (.) Jack Smith?

(8) *CHI2: is vet.
%sit: CHI3 takes the paper.

(9) *CHI2: <*ahora pon aquí cómo llega del parque hasta xxx* >@s [now write down here how it gets from the park to the xxx]

(10) *CHI1: <*¿cómo se dice todo recto?* >@s [how do you say straight?]

(11) *CHI2: the park.

(12) *CHI3: <*por aquí* >@s [over here]

(13) *CHI2: in the park eh left.

(14) *CHI1: <*a ver*>@s [look] left <*es izquierda eh.* >@s [is left, you know?]

(15) *CHI2: <*pues eso, si lo vemos así.* >@s [I know, if we see it this way]

Learner collaborative interaction was also analysed for LREs. Following earlier research on LREs (García Mayo & Azkarai, 2016, among others), interactions were coded on whether they had a lexical focus (*meaning-focused LREs*), namely when the interaction involved the meaning or use of a word, or a focus on form (*form-focused LREs*), which included morphosyntax, prepositions, spelling, but also pronunciation, on the grounds that research has shown that incidental focus-on-form episodes include those involving pronunciation and phonological (or phonetic) form (Ellis, Basturkmen & Loewen, 2001). Based on the work of Leeser (2004), the resolution of LREs was coded as *unresolved* when the participants failed to provide a solution to the issue raised in the LRE and *resolved*, when participants reached a correct resolution. Resolved LREs were further classified as *target-like*, when the LRE was solved correctly, or as *non-target-like* when the solution reached was incorrect.

Interactions were first coded by one of the researchers and then independently coded by a different researcher. Only those LREs which both researchers agreed on were included for analysis. The following examples illustrate the different types of LREs encountered in our data. Example (5) features a meaning-focused LRE, where the participants struggle with the word *straight*. After many turns, including those asking the investigator (turn 5), Child 2 comes up with the word (not until turn 14), which is then incorporated in Child 1's turn (turn 15).

(5) Meaning-focused. Resolved: target-like.

(1) *CHI2: <*recto* y *a la derecha.* >@s [straight and to the right]

(2) *INV: you speak louder eh?

(3) *CHI1: <*pero cómo.* >@s [but how]

(4) *CHI2: <*a la derecha.* >@s [to the right]

(5) *CHI2: right (.) <*¿cómo se dice recto?* >@s [how do you say straight?] (to the investigator)

(6) *INV: you have to collaborate.

(7) *CHI1: xxx.

(8) *INV: and maybe you have other ways.

(9) *CHI2: <*no sabemos.* >@s [we don´t know]

(10) *INV: other ways to eh to communicate that particular word right?

(11) *CHI1: <*ah como en gimnasia.* >@s [ah like in P.E.]

(12) *CHI1: <*¿cómo era?* >@s [how was it?]

(13) *CHI1: right right.

(14) *CHI2: straight.

(15) *CHI1: <*eso*>@s [that's it] straight

In example (6), participants discuss a grammatical feature. Child 2 corrects Child 1 in her use of the possessive adjective, which is then incorporated in Child 1's turn. The LRE is resolved in a target-like manner.

(6) Form-focused. Resolved: target-like.

*CHI1: in her t-shirt.

*CHI2: in his.

*CHI1: <*ay es verdad*>@s [yeah, you're right] in his t-shirt <no? >@s [right?]

*CHI2: <*sí* >@s [yes]

In example (7), the grammatical feature involved is a preposition, with a non-target-like resolution. Child 2 suggests the incorrect preposition in *at the right* in turn 3 (instead of on) to complete his companion's utterance. Subsequently, Child 1 recasts it with *at right*, believing that no definite article is needed in that phrase. Both members move on to complete the sentence (turn 5 onwards). This LRE was thus classified as form-focused and non target-like.

(7) Form-focused. Resolved: non target-like.

(1) *CHI2: *to the church.*

(2) *CHI1: to the church and then eh.

(3) *CHI2: at the right.

(4) *CHI1: at right.

(5) *CHI2: is.

(6) *CHI1: eh.

(7) *CHI2: is the.

(8) *CHI1: is the.

(9) *CHI2: vet clinic.

(10) *CHI1: vet clinic.

5. Results

The results shown in this section will be organised according to the four research questions of the study. First, they will shed light on the first research question about the role of the independent variable of the study, that is, on how the patterns of interaction exhibited by young EFL learners may affect the number, the nature and the outcome of the LREs produced (see Table 1). Second, they will offer the information regarding the second research question on the influence that the moderator variable of the study (pairing method: proficiency-matched vs self-selected) might exert on the number, nature and outcome of the LREs produced by interactancts with different patterns of interaction (see Table 2). Then, we will display the data to answer the third research question addressing the various types of form-focused LREs produced by these learners as well as the potential relationship between those types and the patterns of interaction and/or pairing method (see Table 3). Finally, in Table 4 we will present the results pertaining to the relationship between pairing method per type of pattern of interaction and the types of form-focused LREs produced by the learners (research question 4).

Table 1 organises the LRE data according to the different patterns of interaction. Specifically, it displays the number of groups who exhibit those patterns as well as their

behaviour regarding the production of LREs, namely their incidence (operationalized as the total and the mean number of LREs produced), nature (total number of meaning and form-focused LREs) and resolution (total number of target-like LREs, target-like meaning-based LREs and target like form-focused LREs). Additionally, it offers percentages in an attempt to compare those figures better for the sake of readers' convenience. As can be seen, 20 out of 27 groups, that is, nearly three quarters of the learner groups (74.10%) exhibited a collaborative pattern of interaction, the rest of pairs minimally representing the other patterns of interaction. More precisely, the dominant-passive, the expert-novice and the dominant-dominant dynamics amounted to 14.80% (n=4), 7.40% (n=2) and 3.70% (n=1) of the groups, respectively. As for the incidence of LREs, it is by looking at the mean number of LREs per pattern, rather than at the total number of LREs, that we can establish a reliable comparison among different pair dynamics. Data indicated that the groups that obtained the highest number of LREs were of an expert-novice pattern (mean=8) whereas dominant-passive dynamics produced the lowest number (mean=6). Collaborative and dominant-dominant pairs yielded intermediate values (means= 7 and 6.8). As for the distribution of the LREs according to its nature, a comparison between the percentages of meaning- and form-focused LREs will better reveal which type of LRE is more frequent in the data. As shown in the table, meaning-focused LREs were produced more frequently than form-focused LREs in all pair dynamics, the gap between these two LRE categories being wider in the case of the dominant-dominant groups (71.5% vs. 28.5%). Finally, as regards the outcome of the LREs, the ratio of those LREs resolved in a target-like fashion was higher in the expert-novice (81.25%) and collaborative (69.11%) groups than in the dominant-passive (62.50%) and dominant-dominant (42.86%) groups. When target-likeness was analysed according to the nature of the LREs, a different pattern emerged for meaning- and form-focused LREs. While the former replicated the tendency abovementioned, expert-novice (77.70%) and collaborative (70.60%) dynamics being more accurate than dominant-passive (46.15%) and dominant-dominant (20%) groups in target-like meaning-focused LREs, the proportions of accurate resolutions in the form-focused LREs produced by dominant-dominant (100%), dominant-passive (81.81%) and expert-novice (85.70%) patterns were higher than in case of collaborative groups (67.21%).

Table 1: LRE distribution per pattern of interaction

		All		Meaning		Form	
Patterns	*Number of groups*	*Total/Mean*	*Target-like*	*Total*	*Target-like*	*Total*	*Target-like*
Coll.	20 (74.10%)	136 / 6.8	94 (69.11%)	75 (55.15%)	53 (70.60%)	61 (44.85%)	41 (67.21%)
Dom-Dom.	1 (3.70%)	7 / 7	3 (42.86%)	5 (71.50%)	1 (20%)	2 (28.50%)	2 (100%)
Dom-Pas.	4 (14.80%)	24 / 6	15 (62.50%)	13 (54.16%)	6 (46.15%)	11 (45.84%)	9 (81.81%)
Exp-Nov.	2 (7.40%)	16 / 8	13 (81.25%)	9 (56.25%)	7 (77.70%)	7 (43.75%)	6 (85.70%)

Research question 2 sought to explore the interplay between pairing method and patterns of interaction in the incidence, type and resolution of LREs. The data displayed in Table 2 is structured as in the previous table but according to the pairing method variable (proficiency-matched vs self-selected) in this case. The influence of this variable will only be explored in terms of the data obtained from collaborative groups, following Azkarai & Kopinska's (2020) study, who, due to the imbalanced number of dyads within each pattern, selected a subset of dyads for further analyses. In our study, in the rest of the patterns of interaction one of the two pairing methods is not represented in the data, as is the case of proficiency-matched students in dominant-dominant dynamics and of self-selected students in expert-novice patterns, or there is a lack of one of the two types of LREs in terms of nature, as is the case of form-focused LREs in the proficiency-matched groups with dominant-passive dynamics. Hence, following from the collaborative patterns solely, Table 2 data revealed that 70% of the groups (n=14) had been paired according to their English proficiency whereas 30% of them (n=6) had been chosen by the students themselves. With respect to the whole sample, the number of proficiency-matched collaborative groups represented 82.35% of the total number of proficiency-matched groups (n=~~20~~ 17). Self-selected collaborative groups represented 60% of all self-selected groups (n=10), though. As for the incidence of LREs in either pairing method, Table 2 indicates that the mean number of LREs in self-selected groups (7.33) was slightly higher than the mean number of LREs in proficiency-matched groups (6.57). Regarding the nature of the LREs produced by these groups, proficiency-matched learners' ratio of form-focused LREs (52.17%) was slightly higher than that of meaning-focused LREs (47.83%). On the contrary, self-selected students produced a far greater proportion of meaning-focused LREs (70.45%) than of

form-focused LREs (29.55%). With regard to the outcome of LREs, it was proficiency-matched students who produced a larger proportion of target-like resolutions than self-selected groups (72.82% vs. 61.36%). This very same tendency was observed for both meaning-focused (77.27% vs. 61.29%) and form-focused (68.75% vs. 61.53%) LREs.

Table 2: LRE distribution per pairing method in collaborative-pattern groups

		All		Meaning		Form	
Pairing	*Number of groups*	*Total/Mean*	*Target-like*	*Total*	*Target-like*	*Total*	*Target-like*
PM	14 (70%)	92 / 6.57	67 (72.82%)	44 (47.83%)	34 (77.27%)	48 (52.17%)	33 (68.75%)
SS	6 (30%)	44 / 7.33	27 (61.36%)	31 (70.45%)	19 (61.29%)	13 (29.55%)	8 (61.53%)

The third research question enquired about the interplay between form-focused LREs in particular and the variable patterns of interaction (see Table 3). Results indicate that the average number of this type of LREs per group is higher in expert-novice (3.50) and collaborative (3.05) patterns than in dominant-passive (2.75) and dominant-dominant (2) dynamics. With regard to the different types of form-focused LREs, the collaborative groups produced the highest proportions of focus on language form during their interaction on spelling (47.54%) and morphosyntax (29.51%), followed by prepositions (18.03%) and, minimally, pronunciation (4.92%). Expert-novice groups behaved quite alike, but they excelled in morphosyntax (57.14%) instead of spelling (28.57%), followed by prepositions (14.29%) and exhibiting an absolute lack of pronunciation LREs. Dominant-passive groups' distribution of the various form categories was more homogeneous. These students focused mainly on spelling (36.36%) and pronunciation (27.28%), and then on morphosyntax (18.18%) and prepositions (18.18%). Finally, the dominant-dominant pattern yielded LREs centred on spelling only.

Table 3: Form-focused LRE distribution per pattern of interaction

	All Form	Morphosyntax	Prepositions	Spelling	Pronunciation
Patterns	*Total/Mean*	*Total/Mean*	*Total/Mean*	*Total/Mean*	*Total/Mean*
Coll. (n=20)	61 / 3.05	18 / 0.90 (29.51%)	11 / 0.55 (18.03%)	29 / 1.35 (47.54%)	3 / 0.15 (4.92%)
Dom-Dom. (n=1)	2 / 2	0 / 0 (0%)	0 / 0 (0%)	2 / 2 (100%)	0 / 0 (0%)
Dom-Pas. (n=4)	11 / 2.75	2 / 0.50 (18.18%)	2 / 0.50 (18.18%)	4 / 1 (36.36%)	3 / 0.75 (27.28%)
Exp-Nov. (n=2)	7 / 3.50	4 / 2 (57.14%)	1/0.5 (14.29%)	2 / 1 (28.57%)	0 / 0 (0%)

As for the last research question on the differential behaviour of proficiency-matched and self-selected groups regarding the production of form-focused LRE types, Table 4 offers the data on form-focused LREs for these two different types of pairings in collaborative groups. As can be seen, the proficiency-matched learners produced a higher average number of form-focused LREs than self-selected students (3.43 vs. 2.17). Regarding the occurrence of the different types of form-focused LREs, both pairing groups behaved quite alike, spelling LREs representing nearly half of the LREs produced, and morphosyntax one third. The rest of the LREs centred on prepositions and, to an even lesser extent, on pronunciation.

Table 4: Form-focused LRE distribution per pairing method in collaborative-pattern groups

Form LREs	All	Morphosyntax	Prepositions	Spelling	Pronunciation
Pairing	*Total/Mean*	*Total/Mean*	*Total/Mean*	*Total/Mean*	*Total/Mean*
PM (n=14)	*48 / 3.43*	14 / 1.00 (29.17%)	9 / 0.64 (18.75%)	23 / 1.64 (47.91%)	2 / 0.14 (4.17%)
SS (n=6)	*13 / 2.17*	4 / 0.67 (30.77%)	2 / 0.34 (15.38%)	6 / 1.00 (46.16%)	1 / 0.17 (7.69%)

6. Discussion

Our first research question examined the types of pair dynamics used by young EFL learners in a collaborative writing task. The results showed that, although the

number of the participants in the study is low, these learners exhibited all four types of pair dynamics defined by Storch (2002), of which a high proportion exhibited a collaborative type of dynamics, which is in line with previous research on child pair dynamics in ESL (Oliver & Azkarai, 2019) and EFL (Azkarai & Kopinska, 2020; Basterrechea & Gallardo-del-Puerto, 2020; García Mayo & Imaz Aguirre, 2019) settings. A closer look at the data revealed that proportionately it was the interactancts with an expert-novice pattern the ones that, although marginally, produced the highest number of LREs, groups who also exhibited a higher rate of resolution, together with collaborative groups. The findings in the study are in line with Azkarai and Kopinksa's (2020) study in that task features seem to interplay with the patterns of interaction. In their study, the task chosen was a dictogloss task (Wajnryb, 1990), and the role that prevailed was that of the passive-parallel, with one member acting as the scribe (the passive member) and the other acting as a facilitator. Our findings seem to indicate that the map task, a convergent writing task, enhanced collaboration from the first phase (here learners had to decide who the owner of the dog is), which is observed by Azkarai and Kopinska (2020) as key in fostering optimal types of dynamics (including expert-novice).

As regards resolution, the percentage of meaning-focused LREs which were resolved accurately was much higher in groups who exhibited dynamics which are more conducive to language learning (collaborative or expert/novice), particularly in the case of the expert-novice pattern. The higher number of resolved LREs by expert-novice interactancts, in spite of the limited number of groups, supports previous work with young EFL learners (Pladevall-Ballester, 2021), findings that seem to support the benefits of this type of pattern. However, as for form-focused LREs, it is the dominant/dominant type of dynamics who reached more target-like resolutions. One could speculate that the low mutuality and high competitiveness among the group members and the desire to control over the task typical of this type of dynamics seem to favour target-like resolutions and hence increase learning opportunities, contradicting previous research finding as regards this type of deleterious type of dynamics. Nevertheless, we must be extremely cautious with this assertion, as it is based on the data coming from the interaction of just a single triad. Patterns of interaction other than the collaborative one are particularly scarce in the young learner interaction literature (as attested in e.g., García Mayo & Imaz Aguirre, 2019, or Oliver & Azkarai, 2019). The present study is no exception in that regard, dominant-dominant, dominant-passive and expert-novice dynamics being underrepresented. Thus, we make a call for future research with larger samples of participants, as it will be the only way to increase the chance of observing young learners' behaviour when they engage in these three types of interactional dynamics.

The second research question purported to examine whether the pairing method variable exerted any influence in how learner groups differing in their patterns of interactions attended to language during collaborative writing. As indicated in the previous section, the study only managed to examine the pairing method effect on the production of LREs (in terms of type and accuracy) in the collaborative groups subsample. The pair formation effect was observed not only in the nature, but also in the accuracy of LREs (target-likeness) in favour of proficiency-matched groups. As already attested in previous studies with adult (Mozaffari, 2017; Storch & Aldosari, 2013) and young (Basterrechea & Gallardo-del-Puerto, 2020) EFL learners as well as Spanish-as-a-FL learners (Leeser, 2004), matched proficiency has proven to favour target-like resolutions to a larger extent than self-selection. Mozaffari (2017) interpreted this finding as student-selected pairs being less task-oriented due to their pre-existing friendship, which hindered a greater language focus and acted as a distractor. A closer look at our data revealed that these young learners did not engage in talk unrelated to the task extensively (4 groups out of 17), although the few instances of off-task talk encountered concentrated mostly in self-selected groups (3 groups). Additionally, the comparison of the time needed for task completion showed that on average, self-selected groups needed more time (an average of 13'50'') than proficiency-matched groups (an average of 9'22''). This qualitative inspection of our data confirms Mozaffari's (2017) observation that self-selected pairs are less task-oriented than proficiency-matched ones. However, the participants in our study were not surveyed on the reasons to select a partner and, hence we cannot attribute these findings to pre-existing friendship as the reason behind this off-task behaviour. Future studies that triangulate a questionnaire about the learners' motives, the pair formation variable and the production of LREs (including incidence, type and resolution) would help us reinforce those arguments on the effects of differing grouping procedures. This would also allow us to determine the extent to which the relationship between peers, the attitudes of the learners and focus on language work in interdependent ways in peer interaction (Philp, Walter & Basturkmen, 2010).

The study also sought to explore how pair dynamics would interact with the type of form-focused LREs, an underresearched area with young EFL learners. In an attempt to answer the third research question, it was discovered that all types of patterns focused mainly on spelling, except expert-novice groups, who focused more on morphosyntactic features, a potential indication of expert-novice interactancts' greater explicit grammatical knowledge. These results, again, should be taken with caution due to the low number of groups under this label. Previous studies that look into LREs involving spelling have attested a high amount of these in writing tasks. For instance, Niu (2009) examined the production of LREs by adult EFL learners in oral vs. written tasks. It was found that learners focused heavily on spelling issues in the written task.

Calzada & García Mayo (2021) also reported a high amount of spelling-related LREs in child EFL learners in a written task as well, results that authors attribute to the intrinsic characteristics of the written mode, as has already been observed in studies on writing-to-learn (Manchón, 2011). Muñoz (2014) particularly notes that spelling is a concern at early ages, whereas the focus of morphosyntactic features develops in later maturational stages.

The fourth research question further intended to explore whether pairing method interacted with pair dynamics in these schoolchildren's attention to language form. The data obtained enabled us to examine the variety of form-focused LREs produced by the collaborative cohort only, with proficiency-matched interactancts comparatively producing a larger amount of form-focused LREs than self-selected groups which, again, supports prior research (Basterrechea & Gallardo-del-Puerto, 2020; Leeser, 2004; Mozaffari, 2017; Storch & Aldosari, 2013). As for type, no effect of pair formation method was found in these, as both proficiency-matched and self-selected interactancts focused largely on spelling, followed by morphosyntax. It must be noted that all participants were both young and low-proficient learners of English, facts which could account for their greater awareness of surface-level spelling alterations than of grammatical knowledge-induced morphological or syntactic deviations. A further analysis of the level of elaboration or the various LREs found could be enlightening in that respect (see Pladevall & Vraciu, 2020).

7. Conclusion

Our study attemped to contribute to the field on young EFL learners' task-based attention to form by exploring how patterns of interaction and pairing method affect LRE production in collaborative writing. Particularly, it has also analysed which language features these young learners draw their attention to while completing a convergent map task, in order to provide evidence of young EFL learners' ability to reflect on language use and discuss language, all of which may contribute to further gains in language development. Our data has provided further evidence that young EFL learners exhibit a collaborative type of dynamics to a larger extent, as evinced in previous studies on young learner literature (Basterrechea & Gallardo-del-Puerto, 2020; García Mayo & Imaz Agirre, 2019). The learners with this type of dynamics resolved more LREs accurately, together with expert-novice groups (Pladevall-Ballester, 2021). It was also found that whereas collaborative or expert/novice interactancts excelled in resolving meaning-focused LREs, the limited number of groups exhibiting a dominant/dominant pattern reached more target-like resolutions in form-focused LREs.These preliminary finding corroborates the observation that pair dynamics in young learners involve a complex array of factors (Oliver & Azkarai, 2019), and calls

for future studies with a larger amount of participants that may allow to interpret results more rigurously.

The study also supported previous research findings on the benefits of matched proficiency (e.g., Basterrechea & Gallardo-del-Puerto, 2020), as a pair formation effect was observed in the nature and the accuracy (target-likeness) of LREs in favour of proficiency-matched groups, who, contrary to self-selected interactancts, exhibited less off-task talk and concentrated on task completion more efficently, suggesting that teachers should consider exploring different pairings in order to maximise young EFL learners' learning opportunities. Finally, a closer look at the type of form-focused LREs produced suggested that these young learners focused primarily on spelling issues and less on grammatical knowledge-induced ones, supporting previous findings that language learning depends on the learners' developmental readiness (Leeser, 2004; Pladevall-Ballester, 2021). It is also worth noting that the participants of the study were 10-12-year-old children. At the age of 10, children begin to think in an organised and logical fashion and can reflect about their own thinking and their own language use (Pinter, 2006). Nonetheless, whether these learners receive any kind of explicit attention to grammatical form in their regular classes was not controlled for in the study, an issue worth investigating in future research.

Another aspect of further enquiry would derive from a limitation of our study concerning the size of the groups, a factor that has preliminarily shown some influence on the production of LRE and their focus (meaning or form) in adults (Fernández Dobao, 2012), but its incidence in young leraners is yet to be explored.

Acknowledgements

This work was supported by grants FFI2016-74950-P (Spanish Ministry of Economy and Competitiveness, National Research Agency and European Regional Development Fund- AEI/FEDER/EU) and IT904-16 (Basque Government).

8. References

Ahmadian, M., & Tajabadi, A. (2017). Patterns of interaction in young EFL learners' pair work: the relationship between pair dynamics and vocabulary acquisition. *3L: The Southeast Asian Journal of English Language Studies, 22*(3), 98–114.

Azkarai, A., & Kopinska, M. (2020). Young EFL learners and collaborative writing: A study on patterns of interaction, engagement in LREs and task motivation. *System, 9*, Article 102338. https://doi.org/10.1016/j.system.2020.102338

Azkarai, A., García Mayo, M.P., & Oliver, R. (2020). The effect of task repetition on the patterns of interaction of ESL children. *International Journal of Applied Linguistics (ITL), 171*(1), 90-112.

Basterrechea, M., & García Mayo, M.P. (2013). Language-related episodes (LREs) during collaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough, & A. Mackey (eds.), *Second language interaction in diverse educational contexts* (pp.25-43). Amsterdam: John Benjamins.

Basterrecha, M. & Gallardo-del-Puerto, F. (2020). Language-related episodes and pair dynamics in primary school CLIL learners: A comparison between proficiency-matched and student-selected pairs. *Studies in Second Language Learning and Teaching, 10*(3), 423-447.

Basterrechea, M. & Leeser, M. (2019). Language-related episodes and learner proficiency during collaborative dialogue in CLIL. *Language Awareness, 28*(2), 97-113.

Benson, C., Pavitt, J., & Jenkins, M. (2005). The use of dictogloss to encourage discussion of language use. *Edinburgh Working Papers, 14,* 1–17.

Butler, Y. G., & Zeng, W. (2015). Young foreign language learners' interactional development in task-based paired assessment in their first and foreign languages: A case of English learners in China. *Education, 44*(3), 292-321.

Calzada, A., & García Mayo, M.P. (2021). Child learners' reflections about EFL grammar in a collaborative writing task: When form is not at odds with communication. *Language Awareness, 30*(1), 1-16.

Cambridge University Press (2008). *Key English Test*. Cambridge: Cambridge University Press.

Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition, 36*(4), 451-485.

Donato, R. (1994). Collective scaffolding in second language learning. In J.P. Lantolf, & G. Appel (eds.), *Vygotskian Approaches to Second Language Research* (pp. 33-56).Westport (CT): Ablex Publishing.

Ede, L., & Lunsford, A. (1990). *Singular texts/plural authors.* Carbondale: Southern Illinois University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly, 35*(3), 407–432.

Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group. Pair, and individual work. *Journal of Second Language Writing, 21,* 40-58.

Gallardo-del-Puerto, F., & Basterrechea, M. (2021, online). The role of proficiency and pair formation method in language-related episodes: A study of young CLIL learners' interaction. *Language Teaching Research*, 1-19.

García Mayo, M.P. (2002). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research, 37,* 323-341.

García Mayo, M.P. & Azkarai, A. (2016). EFL task-based interaction: does task modality impact on language-related episodes? In M. Sato, & S. Ballinger (eds.), *Peer interaction and Second Language Learning: Pedagogical Potential and Research Agenda* (pp. 241-266). Amsterdam: John Benjamins.

García Mayo, M.P. & Hidalgo, M.A. (2017). L1 use among young EFL mainstream and CLIL learners in task-supported interaction. *System, 67*, 132-145.

García Mayo, M.P. & Imaz Aguirre, A. (2019). Task modality and pair formation method: Their impact on patterns of interaction and LREs among EFL primary school children. *System, 80,* 165-175.

García Mayo, M.P. & Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System, 54*, 40-54.

Gass, S.M., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning, 55*(4), 575–611.

Gilabert, R., Barón, J., & Llanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics, 47*(3–4), 367–395.

Hidalgo, M.A., & García Mayo, M.P. (2021). The influence of task repetition type on young EFL learners' attention to form. *Language Teaching Research, 25*(4), 565–586.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research, 12*, 211–234.

Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3*(2), 73-93.

Lantolf J.P., & Appel, G. (1994). *Vygotskian Approaches to Second Language Research.* Norwood, NJ: Ablex.

Leeser, M. (2004). Learner proficiency and focus on form during collaborative Dialogue. *Language Teaching Research, 8,* 55-81.

Long, M.H. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie, & T.K. Bhatia (eds.), *Handbook of Language Acquisition: Vol. 2. Second Language Acquisition* (pp. 413-468). New York: Academic Press.

López-Serrano, S., Roca de Larios, J., & Manchón, R. (2019). Language reflection fostered by individual L2 writing tasks: developing a theoretically motivated and empirically based coding system. *Studies in Second Language Acquisition, 41*(Special Issue 3), 503-527.

Mackey, A. (1994). Targeting morpho-syntax in children's ESL: An empirical study of the use of interactive goal-based tasks. *Working Papers in Educational Linguistics, 10*, 67-88.

Mackey, A. & Oliver, R. (2002). Interactional feedback and children's L2 development. *System, 30*(4), 459-477.

Mackey, A. & Silver, R. E. (2005). Interactional tasks and English L2 learning by immigrant children in Singapore. *System, 33*(2), 239-260.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk (third edition). Vol. I. Transcription format and programs, Volume II: The database.* Mahwah, NJ: Lawrence Erlbaum.

Malmqvist, A. (2005). How does group discussion in reconstruction tasks affect written language output? *Language Awareness, 14*(2–3), 128–141.

Manchón, R.M. (2011). Writing to learn the language: Issues in theory and research. In R.M. Manchón (ed.), *Learning-to-Write and Writing-to-Learn in an Additional Language.* (pp. 61–82). John Benjamins.

Manchón, R.M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón, & P.K. Matsuda (eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin, Germany, and Boston, MA: De Gruyter Mouton.

Martínez-Adrián, M., & Gutiérrez-Mangado, M.J. (2022). Gender pairings in EFL task-based interaction. *Vigo International Journal of Applied Linguistics, 19*, 103-128

Mozaffari, S.H. (2017). Comparing student-selected and teacher-assigned pairs on collaborative writing. *Language Teaching Research*, *21*(4), 496-516.

Muñoz, C. (2014). The development of language awareness in the transition from primary to secondary school. In M.P. García Mayo (ed.), *Learning Foreign Languages in Primary School: Research Insights* (p. 49-68). Bristol: Multilingual Matters.

Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness, 18*(3-4), 384-402.

Oliver, R. (1998). Negotiation of meaning in child interactions. *The Modern Language Journal, 82*(3), 372-386.

Oliver, R. & Azkarai, A. (2019). Patterns of interaction and young ESL learners: What is the impact of proficiency and task type? *Language Teaching for Young Learners, 1*(1), 82-102.

Oliver, R., Philp, J., & Duchesne, S. (2017). Children working it out together: A comparison of younger and older learners collaborating in task based interaction. *System, 69*, 1-14.

Philp, J. Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: what factors foster a focus on form? *Language Awareness, 19*, 261-279.

Pinter, A. (2006). Verbal evidence of task related strategies: Child versus adult interactions. *System, 34*(4), 615-630.

Pinter, A. (2007). Some benefits of peer-peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research, 11*(2), 189-207.

Pladevall-Ballester, E. (2021). Pair dynamics and language-related episodes in child EFL task-based peer interaction. *Language Teaching for Young Learners, 3*(2), 189 – 213.

Pladevall-Ballester, E. & Vraciu, A. (2020). EFL child peer interaction: Measuring the effect of time, proficiency pairing and language of interaction. *Studies in Second Language Learning and Teaching, 10*(3), 449-472.

Plonsky, L., & Y. Kim. (2016). Task-based learner production: A substantive and methodological review. *Annual Review of Applied Linguistics, 36*, 73–97.

Roehr-Brackin, K., & Tellier, A., (2019). The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition, 41*(5), 1111-1131..

Sato, M., & Ballinger, S. (2016). Understanding peer interaction: Research synthesis and directions. In M. Sato, & S. Ballinger (eds.), *Peer Interaction and Second Language Learning. Pedagogical Potential and Research Agenda* (pp. 1-30).Amsterdam: John Benjamins.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research, 5*, 29-53.

Storch, N. (2002). Patterns of interaction in ESL pairwork. *Language Learning, 52*, 119–158

Storch, N. (2016). Collaborative Writing. In R. M. Manchón, & P. K. Matsuda (eds.), *Handbook of Second and Foreign Language Writing* (pp. 387-406). Berlin, Germany, and Boston, MA: De Gruyter Mouton.

Storch, N. & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research, 17*, 31–48.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty, & J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 64-81). Cambridge: Cambridge University Press.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). London: Continuum.

Swain, M. (2010). "Talking it through": languaging as a source of learning. In R. Batstone (ed.), *Sociocognitive perspectives on language use/learning* (pp. 112-130). Oxford: Oxford University Press.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J.P. Lantolf (ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.

Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320-337.

Swain, M. & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (eds.), *Researching Pedagogic tasks: Second Language Learning, Teaching, and Testing* (pp. 99-118). New York: Longman.

Tan, L., Wigglesworth, G., & Storch, N. (2010). Pair interaction and mode of communication: Comparing face-to-face and computer mediated communication. *Australian Review of Applied Linguistics, 33,* 1-27.

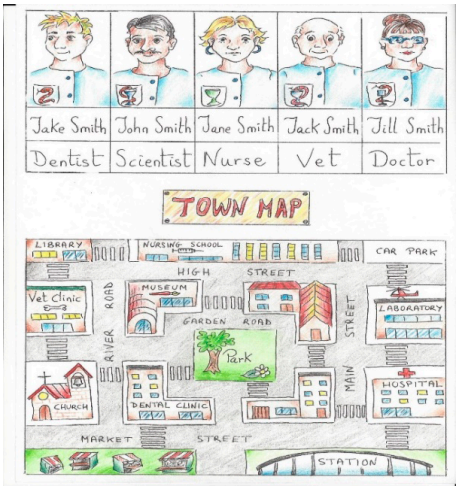
Wajnryb, R. (1990). *Grammar Dictation*. Oxford: Oxford University Press.

Williams, J. (2008). The speaking-writing connection in second language and academic literacy development. In D. Belcher, & A. Hirvela (eds.), *The oral-literate connection: perspectives on L2 speaking, writing and other media interactions* (pp. 10-25). Ann Arbor: The University of Michigan Press.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21,* 321-331.

Appendix

WHOSE DOG IS IT?



Alberto San Emeterio Bolado©

Look at the pictures on the two pages provided by the researcher. In the first one you can see a young boy and a dog in a park. The boy has found a dog in the park. There is something written in the collar and a picture where you can see the dog in the hands of its owner. But you cannot see the owners face. On the second page you can also see some possible owners and where each of them works: you have a dentist, a scientist, a nurse, a vet and a doctor. You can also see a map where you can see all the most important places in the town, including all the places where the possible owners work.

Your task is to work together to do two things:

1) Examine the pictures and decide who the owner of the dog is and discuss why you think so.

2) Write down a short note for the boy explaining the following:

   a. who the owner is

   b. why you think so

   c. how to take the dog to its owner.

Metaphors set in motion in the context of L2 academic spoken discourse

Irene Castellano-Risco
Department of English Philology
University of Extremadura
ircastellano@unex.es

Marta Martín-Gilete
Department of English Philology
University of Extremadura
mmgilete@unex.es

Alberto Hijazo-Gascón
Department of Linguistics and Spanish Literature
University of Zaragoza
ahijazo@unizar.es

Iraide Ibarretxe-Antuñano
Department of Linguistics and Spanish Literature
University of Zaragoza
iraide@unizar.es

Abstract

Theoretical and applied research in physical motion events (MEs) (Talmy, 2000; Ibarretxe-Antuñano, 2017) has shown that (1) speakers tend to organize MEs around a specific semantic component resulting in at least two different lexicalization patterns (satellite-framed and verb-framed languages), with different narrative styles (Slobin, 1996, 2004); (2) the use of specific patterns may also depend on genre and discourse (Caballero, 2017); and (3) the adoption of the narrative style of an L2 is challenging (Cadierno, 2008). These findings seem to also apply to metaphorical motion events (MMEs) (Özçalışkan, 2005; Ibarretxe-Antuñano & Caballero, 2014).

This study analyses the linguistic realizations and discourse functions of MMEs in the specific context of oral EMI seminars to examine ME prevalence and the impact inter-typological differences may have on the production of MMEs. Data from three seminars from METCLIL corpus were selected to compare MME verb production by English L1 speakers (n=4) and L2 English speakers (satellite-framed: n=4; verb-framed: n=9). Results show that (i) individual variables (speakers' L1 typology or lecturer's

input) do not seem to have a determining impact on MME employment, while (ii) contextual factors (to-be-performed activities) have been shown to have a more salient role in MME usage and performance of discursive functions.

Keywords: metaphor, motion, spoken discourse, academic seminars, English as a Medium of Instruction (EMI).

Resumen

La investigación sobre eventos de movimiento (EMs) físico (Talmy, 2000; Ibarretxe-Antuñano, 2017) ha demostrado que (1) los hablantes tienden a organizar los EMs según un componente semántico específico que produce al menos dos patrones de lexicalización (lenguas de marco satélite y verbal) con diferentes estilos narrativos (Slobin, 1996, 2004); (2) el uso de patrones específicos puede depender del género y del discurso (Caballero, 2017); y (3) la adopción del estilo narrativo de una L2 es complicada (Cadierno, 2008). Esto parece aplicarse también a los eventos de movimiento metafórico (EMMs) (Özçalı kan, 2005; Ibarretxe-Antuñano y Caballero, 2014).

Este estudio analiza el empleo de EMMs en seminarios académicos examinando la frecuencia de EMs y el impacto de las diferencias inter-tipológicas en la producción de EMMs, en tres seminarios del corpus METCLIL, comparando el uso de verbos EMM en hablantes de inglés L1 (n=4) y L2 (lengua satélite: n=4; marco verbal: n=9). Los resultados indican que variables individuales como la tipología de L1 no parecen tener un impacto decisivo en el uso de EMMs, mientras que factores contextuales, como la actividad a desarrollar, son más determinantes en el uso de EMMs y su ejecución de funciones discursivas.

Palabras clave: metáfora, movimiento, discurso oral, seminarios académicos, Inglés como Medio de Instrucción (EMI, según sus siglas en inglés).

1. Introduction

English as Medium of Instruction (EMI) programs have undergone a significant increase worldwide in the last couple of decades (Dafouz & Smit, 2020; Macaro, 2018). These educational programs offer content courses in which English is used as a vehicular language, mostly in Higher Education institutions where English does not have official language status. In this regard, one of the main benefits of the programs is that students, primarily L2 English speakers, are provided with opportunities to prepare themselves to meet the demands of a globalized economic world, in a context where development of English language skills is not a primary aim of the course

but is nevertheless the result of its instructional application in the class (Pecorari & Malmström, 2018).

EMI programs have developed in parallel to the implementation of other content-based practices, such as the Content and Language Integrated Learning (CLIL) approach, at different educational levels. Both content-based practices share the objective of helping students develop a second language (L2) in an educational context where the additional language is used meaningfully. However, research in both areas has taken different, although complementary, paths despite their convergent goals. On the one hand, research on EMI has heavily concentrated on characterizing the phenomenon theoretically (Macaro, 2018; Pecorari & Malmström, 2018) and identifying the self-perceived needs of main stakeholders (Aguilar & Rodríguez, 2012; O'Dowd, 2018). In contrast, CLIL research has primarily focused on analyzing the linguistic impact of its implementation and characterizing the approach linguistically (Alejo-González, 2018). In this regard, research has explored language use in CLIL focusing on elements such as interpersonal discourse management (Dalton-Puffer, 2003; Dalton-Puffer & Nikula, 2006), the identification of classroom registers (Llinares et al., 2012) or academic genres (Lorenzo, 2013) and their implications for language learning, and the analysis of the impact of the first language (L1) and the L2 on classroom interaction (Nikula, 2005).

In our view, the linguistic-centered approach adopted in CLIL research should be incorporated into the EMI context. Exploring the linguistic setting in which EMI takes place and identifying the potential linguistic difficulties students may encounter in real EMI environments might represent a step forward in helping students overcome the linguistic challenges EMI courses may pose when attending to and interacting in lectures delivered in English.

This study seeks to provide some insights into EMI learners' metaphor use in the source domain of motion, by means of exploring the employment of these linguistic elements and their performance of discourse functions. As will be seen, section 2 first reviews the use of metaphor in academic contexts and the realization of motion events (MEs) in discourse. Among the different linguistic realizations to be explored, metaphorical motion events (MMEs) have been selected. Given the current literature that identifies metaphor as one of the missing linguistic characterizations in CLIL and EMI research (Alejo-González & García-Bermejo, 2019), this study seeks to provide some insights into EMI learners' metaphor use in the source domain of motion by exploring the function of the target linguistic expressions. Among the different EMI academic registers (i.e., lectures, written reports, essays, or seminars), the present study has concentrated on academic seminars to explore L2 learners' genuine oral production in a highly participative instructional context where students have a central

role. The focus of the analysis of students' oral production of motion expressions is twofold: (i) metaphor production, and (ii) discursive-pragmatic functions for which metaphorical motion is employed. Section 2 first reviews metaphor usage in academic contexts and the realization of MEs in discourse. Section 3 details the methodology employed. Results are presented in section 4 and discussed in section 5. Finally, some conclusions derived from our findings are drawn in section 6.

2. Metaphor, academic contexts, and motion

Before undertaking this study, however, it is necessary to justify the importance of metaphor in academic environments and the key role of motion events in discourse. The following sections present a brief literature review of some selected works which have explored these aspects and served as the theoretical and methodological basis for our research.

2.1. Metaphor in academic discourse

Metaphor, the understanding of one idea (source domain) in terms of another (target domain), is not only ubiquitous in everyday language (Lakoff & Johnson, 1980, 1999) but is also of importance to academic discourse (Herrmann, 2013). In the case of EMI contexts, as in any other academic environment, the high degree of abstract language employed involving linguistic realizations of broader generalizations (conceptual metaphors [CMs]) and some of the difficulties these pose for L2 students (Littlemore, 2001; Littlemore & Low, 2006; MacArthur, 2016a; Littlemore et al., 2014; Littlemore & Low, 2016a) suggest that metaphor in university contexts deserves some attention.

Much of the current research on academic events has largely focused on written discourse (see Semino, 2008, for a review), and scant attention has been paid to the presence of metaphor and its role in academic talk, which may be mostly limited to the analysis of specific linguistic aspects in the academic mentoring of L2 students in English as a Lingua Franca (ELF) contexts (e.g., MacArthur & Littlemore, 2011; MacArthur, 2016a, 2016b; Alejo-González, 2021). As has been observed (Alejo-González, 2022), unevenly distributed metaphor activity in office-hour consultations accounts for metaphor bursts or clusters (cf. Cameron & Stelma, 2004) –bounded peaks of metaphor density– occurring with discursive functions, characteristic of the nature of this dialogic interaction. Previous studies (Low et al., 2008; Beger, 2011) also investigated lecturers' use of metaphor to organize their discourse, framing problems, changing topic, or for evaluative purposes. However, little is known about L2 learner metaphor use in other spoken academic environments like seminars, where metaphor is encountered in diverse interactions.

There is a need to describe the metaphors used by L2 learners when talking about the learning experience in academic discourse, with special attention to frequently constructed source domains resulting in varied conventional expressions (Kövecses, 2002; Semino, 2005). To the best of our knowledge, visual perception (UNDERSTANDING IS SEEING) has been a source domain explored in relation to metaphor use in spoken academic discourse (see MacArthur et al., 2015). For the pedagogical purposes of EMI instruction, the exploration of how some key source domains are constructed in academic discourse may contribute not only to a broader characterization of the use of metaphor in the talk about different topics and in different academic activities but also to facilitate students' achievement in their academic work (Boers, 2000).

The source domain of MOTION has been found to be scarcely used by L2 English speakers in written academic language (Moghadam & Samar, 2020) yet frequently present in the academic mentoring of ELF students (Alejo-González, 2021). In this regard, exploring the construction of MOTION in other academic spoken discourse contexts like seminars may shed light on the characterization of educational talk among L2 students.

2.2. Motion events

The study of MEs from a Cognitive Linguistics perspective stems from Talmy's (1991, 2000) work on semantic typology. According to Talmy, languages can be grouped into two typological groups, depending on the encoding of the semantic component of Path, which includes the trajectory of the movement. In verb-framed languages, Path tends to be encoded in the main verb, e.g., *salir* 'go out' in Spanish. In contrast, in satellite-framed languages Path is normally encoded outside the main verb of the event, e.g., *up* in *run up* in English. Slobin (1996, 1997, 2004) found in ME typology an excellent case to develop his Thinking for Speaking hypothesis, i.e., how our semantic and conceptual encodings in our L1 may influence some cognitive abilities such as attention and memory. Slobin (1996, 1997, 2004) found that speakers of satellite-framed languages usually express the components of Manner and Path of motion with more details in their narratives, because of their typological characteristics –Manner is easily encoded in the main verb, Path is not normally inferred, etc. His findings were corroborated with further research on MEs in many languages –see the studies in Berman & Slobin (1994); Strömqvist & Verhoeven (2004) and Ibarretxe-Antuñano (2017). The study of motion events has generated heated debates, but has also been successfully applied to several fields, including translation and L2 acquisition –see Filipović & Ibarretxe-Antuñano (2015); Cadierno (2017) for an overview. Current views on motion events argue for the consideration of motion events typology as a cline, with high- and low-Manner-salient languages (Slobin, 2004) and high- and low-Path-salient languages (Ibarretxe-Antuñano, 2009).

The scope of the study of motion is not limited to translational motion, as it also reaches fictive and metaphorical motion. Fictive motion refers to MEs that are used to describe dynamic situations, using an entity that cannot move (e.g., *a road*) and a motion verb that highlights the special configuration (Matlock, 2004). For instance, *This road zigzags along the coast*. Metaphorical motion involves the use of motion as a source domain to express the non-literal movement of an abstract entity (Özçalışkan, 2005; Ibarretxe-Antuñano & Caballero, 2014; Caballero, 2017). For example, *Prices are going up after the crisis*.

This paper focuses on the latter. More concretely, it examines whether learners use MMEs and for which purpose. To this aim, we follow Ibarretxe-Antuñano & Caballero's (2014; Caballero & Ibarretxe-Antuñano, 2015) approach to metaphorical motion. These authors classify verbs in MMEs into two categories in agreement with two criteria. First, the semantic information of verb involved. MMEs could be Motion1 (M1), i.e., motion verbs *stricto sensu*, when the verb includes motion information in its semantic description as illustrated in example (1), and Motion2 (M2), i.e., when the verb, despite not being a motion verb per se, can be reinterpreted as such due to the construction it is used in as in example (2). Second, MMEs were classified based on the motion semantic elements they encoded. For instance, both examples (1) and (2) lexicalize Manner, that is, the way in which motion is performed.

(1) *Excellent serving saw Federer leapfrog his opponent to reach match point*

(2) *Rafa rampages into Fourth Straight French Open Final*
[Ibarretxe-Antuñano & Caballero, 2014: 148]

The typological differences found in non-figurative motion seem also present in metaphorical motion. For example, Özçalışkan (2005) examined MMEs in written texts (newspapers and novels) and elicitations with questionnaires, focusing on the semantic component of Manner. She found that the typological differences between English and Turkish are extended in metaphorical motion, since Turkish speakers, as speakers of a verb-framed language, paid less attention to Manner, expressing it less in the verb or other elements of the MMEs, and made fewer distinctions of Manner expression.

The restructuring of MEs in an L2 presents numerous challenges for learners (Cadierno, 2004; Negueruela et al., 2004; Navarro & Nicoladis, 2005; Larrañaga et al., 2011; Hijazo-Gascón, 2021). Cadierno (2004) shows how learners are faced with the complex task of acquiring the L2 Thinking for Speaking, in a process called Re-Thinking for Speaking (Robinson & Ellis, 2008). Cross-linguistic influence and conceptual transfer (Jarvis & Pavlenko, 2008) play a crucial role in this process, and results in experimental studies show different degrees of influence of the L1

lexicalization patterns in the expression of motion in the L2 (see Cadierno, 2017). The study of MMEs in the acquisition of an L2 is still scarce and, to our knowledge, virtually non-existent in the case of MMEs in the classroom.

In coherence with previous research in non-metaphorical motion events, it is expected that the typology of the L1 will influence the use of MMEs in L2 English. Consequently, learners whose L1 is satellite-framed, like the target language, should produce more varied and more frequent MMEs than learners whose L1 is verb-framed. An open question that this study aims to address is when and how MMEs are used in the classroom and with what function. Most research in MEs and L2 acquisition focuses on data elicited in experimental settings –with very few exceptions, like Li et al. (2014). Looking into how MMEs are used in an EMI context will expand our understanding of how learners use the MOTION source domain and to what extent typological differences influence this use.

3. Methodology

3.1. Objectives and research questions

This article intends to investigate motion in the specific context of EMI interactive seminars. In particular, this study seeks to characterize seminar talk by identifying the source domain of MOTION in the learning experience of EMI students –both L1 and L2 English speakers. The main aim of this study is, thus, twofold:

a. To explore inter-typological differences in the MMEs used in EMI spoken academic discourse. With this objective, we aim to address the following research questions:

   RQ1. Do advanced L2 English learners use MMEs when talking about marketing and business in spoken academic interaction? If so, what are their forms, frequency and density?

   RQ2. If used, does participants' L1 typology (satellite- or verb-framed) influence their use of MMEs? In what way?

b. To determine when and with which discursive-pragmatic purposes, MMEs are employed by EMI learners. This aim will be developed by considering these two research questions:

   RQ3. In which context, how and for what purposes is the source domain of MOTION constructed metaphorically in EMI academic seminars?

RQ4. Does the L1 typology of participants have any impact on the pragmatic-discursive goals of the activities and the extent to which MMEs are used in the different seminar activities?

3.2. Corpus background information: METCLIL

Data were extracted from the open access METCLIL (Corpus of Metaphor in Academic Talk), the main outcome of the Spanish national-funded project METCLIL (Alejo et al., 2021). METCLIL contains transcribed oral production –110,496 tokens– from the interaction of EMI instructors and students (mostly L2 English speakers) in nine academic seminars on Marketing and Business administration topics at six European Higher Institutions in six different countries (Spain, Portugal, Italy, the Netherlands, Sweden, and Norway).

Among the different seminars available in the corpus, the three seminars taught in the Spanish Higher Education Institution have been selected. The choice corresponds to the diversity of mother tongues, and the similarity in the structure and balanced combination of activities carried out in the seminars.

3.2.1. Linguistic setting: type of interaction and activities

The three 90-minute seminars are similar in nature of the interaction, duration, and type of activities: 50 minutes devoted to individual presentations (namely, business pitch) and constructive feedback discussion, and 40 minutes for the performance of the same activities but in small-group dynamics.

Seminars are organized around four complementary activities. The first activity consisted of an initial 30-second business pitch (henceforth, pitch) delivered by each student. In the pitch, required to be prepared in advance, students were asked to set themselves apart by providing an informative overview of their own expertise, interests, and prospective career while emphasizing their strengths. The pitch was followed by a second activity –constructive feedback discussion– in which the lecturer and the rest of students reviewed the performance. Subsequently, students were split into two groups to carry out the third activity, where they delivered a further 90-second pitch on the same topic, receiving feedback exclusively from their fellow students. Finally, each group agreed on the most outstanding pitch to be delivered as part of an in-class competition, in which this fourth activity required the target representatives of each group to deliver the 90-second pitch version.

3.2.2. Participants

Concerning participants in the Spanish seminars, only a convenience sample has been explored. After a preliminary analysis, lecturer's speech was not included in the final study given its significant differences in comparison to students' speech in terms of the number of tokens uttered, role in the classroom, and activities and actions performed. In addition, from the total of students (N = 39), the final population comprised exclusively the spoken discourse of 17 students: 4 native speakers (NSs) of English; 9 Spanish L2 English learners, and 4 L2 English participants with a satellite-framed language as their mother tongue (Swedish, German, and Dutch). Given the specific nature of our data (EMI discursively and typologically constrained speakers), the availability of suitable informants is restricted. Therefore, despite the limited data employed in our study, results reported below seem robust. Future studies may shed some light on the scope of the findings later discussed in this paper. Table 1 offers an overview of the participants' speaker-related and linguistic characteristics according to their L1 typology.

Table 1: Participants' overview by L1 typology: speaker-related and linguistic characteristics

L1 typology	N	L1	Self-rated L2 level	Gender
Satellite-framed (NSs)	4	English	C2: 4	Males: 1 Females: 3
Satellite-framed L1 languages	4	Swedish: 1 Dutch: 2 German: 1	C1: 3 C2: 1	Males: 2 Females: 2
Verb-framed L1 language	9	Spanish	B2: 1 C1: 6 C2: 2	Males: 5 Females: 4

3.3. Procedure

For the purposes of this study, the exploration of MMEs concerned primarily the analysis of verb use as being the major MEs realized in the seminars. Therefore, further references to MMEs will address the metaphorical occurrence of motion verbs in discourse.

3.3.1. Corpus analysis

Three different analyses were applied in this study: (i) a motion analysis, to separate the use of MMEs; (ii) a metaphor analysis, to examine the use of MMEs; and

(iii) a pragmatic-function analysis, to explore the context in which, how and for which purpose, MMEs were employed. The first and last analyses were carried out by the authors of this paper, whereas for the metaphor analysis, METCLIL corpus (Alejo-González et al., forthcoming), available at Sketch Engine, was consulted. This corpus provides information about metaphor use. Further information about the different analyses will be given in the following sections.

3.3.1.1. Metaphor

The metaphor analysis was obtained from METCLIL corpus v2 (in process), which includes a metaphor-annotated version following an adapted version of the Metaphor Identification Procedure VU University Amsterdam procedure (MIPVU; Steen et al., 2010). This protocol consists of four consecutive phases: (i) general understanding of the text; (ii) identification of the lexical units; (iii) determination of the basic sense and the contextual meaning of each unit; and (iv) decision on whether the item has a metaphorical use: *not-metaphor*, when the linguistic unit is not considered metaphorical; *indirect*, when the linguistic unit is considered metaphorical but there is not a direct reference to the metaphor, or *direct*, when there is a direct reference to metaphor. To illustrate the procedure, consider the linguistic unit (LU) *reach* and *creating* in example (3) extracted from one of the seminars explored:

(3) *I will try to explain better how we* ***reached*** *that idea of* ***creating*** *this app what was exactly the payment solving* <ELO; L1=Spanish; L2 level=C1>

The MIPVU procedure employs Macmillan Dictionary as the primary source to determine the contextual and basic meanings of LUs. In the specific case of *reach*, its contextual meaning is "MM3. to achieve something after discussing it or thinking about it for a long time", whereas the basic sense of the term is "MM2. to move you hand, arm, leg, etc. towards something that you are trying to touch or pick up". There is no contrast between its contextual meaning and basic sense. Accordingly, *reach* is identified as an *indirect* metaphor-related word (MRW). On the other hand, the LU *creating* is considered *non-metaphor* because its contextual meaning ("MM1. to make something new or original that did not exist before") coincides with the basic sense of the term

3.3.1.2. Motion

The exploration of MEs was dependent on the systematic exploration of the contextual meaning of each verb utterance. First, all the verbs were listed using Sketch Engine and arranged by lemma. Second, the definition of each verb was looked up in

the Macmillan dictionary online to examine the semantic content of each verb and its reference to (or lack of) motion.

Third, the sense in the different contexts in which each verb was produced was examined. In this respect, following Ibarretxe-Antuñano & Caballero's (2014) procedure (see Section 2.2), two main types of motion expressions (MEs) were identified: (i) If the contextual meaning implied motion, and the dictionary contained a definition directly related to motion, this item was labeled as M1; (ii) whereas when in the specific context in which the word was uttered the verb entailed motion –but the dictionary did not include any motion-related meaning– this verb utterance was tagged as M2. Finally, once this first analysis of MEs was concluded, each ME was semantically classified into four types following previous motion verbs categorizations (Cifuentes-Pérez, 2010; Hijazo-Gascón et al., 2013; Hijazo-Gascón, 2021): Path (e.g., *come*), Manner (e.g., *bounce*), Caused-motion (e.g., *bring*), and Neutral (e.g., *move*).

This procedure has led to the inclusion of verbs that other authors may have classified differently. For instance, among the different verbs we have identified, in the literature, *go* has been considered Neutral or Path and *get*, being such a general verb, as motion or non-motion. Following the criteria employed in this study, *go* is a Neutral verb since it means "move from one place to another" and *get* a (Neutral) motion verb since one of its meanings is "move in order to pick up or bring (something)" (definitions taken from the online Oxford Dictionary at https://www.lexico.com/).

3.3.1.3. Pragmatic function

To explore the main discursive-pragmatic functions for which the MMEs were employed, these were identified in the contexts of interaction occurring in the seminars. The two main types of activities performed (pitch-delivery and constructive feedback discussion) served as a basis for determining the specific context in which MMEs were used in the seminar, and for what purposes they were employed by looking at the target domains they relate to.

Besides, we drew on Christie's (2002) classification of L2 classroom registers (applied to CLIL contexts in Llinares et al., 2012) to explore potential pragmatic functions. Christie (2002) distinguishes between two "sets of language choice in classroom texts" (p.15), also known as registers: a 'regulative' register, dealing with the type of language employed to regulate behaviors in the classroom, and an 'instructional' register, referring to the language employed when teaching and learning of the content of the classroom.

3.3.2. Quantitative analysis

Two corpus programs were applied to perform the quantitative analysis. Sketch Engine, the website corpus tool where METCLIL is hosted, was employed to calculate the number of tokens, lemmas and MRWs as well as to look at participants' speakers-related and linguistic features. Additionally, the application of TAALED (Kyle et al., 2021) complemented the analysis by providing some lexical diversity indexes, such as the number of content words or the lexical density, i.e., the number of content words divided by the number of tokens.

4. Results

Data will be presented and explored following the objectives set in section 3.1.

4.1. Objective a: To explore inter-typological differences in the MMEs used when talking about marketing and business in EMI spoken academic seminars

A preliminary description of the sample shows that over the number of items uttered in the seminar, the 17 speakers produced a total of 9,871 lexical items (see Table 2 for a more detailed data description). Nearly a third of them were content words, that is, words with lexical meaning, resulting in a lexical density of 0.37.

Table 2: A summary of the analysis of lexical and verb usage

	Measures	TOTAL
Lexical analysis	Tokens	9,871
	Content words	3,668
	Lexical density (content words/tokens)	0.37
Verb-use analysis	ME tokens	74
	ME lemmas	19
	Total number of verbs	1,803
	ME density (ME tokens/verb tokens) (in %)	4.10
	ME diversity (ME lemmas/ tokens)	0.66

Regarding MEs, their employment seems to be relatively low in comparison to the total number of verbs. Focusing exclusively on lexical verbal items with a motion sense,

as shown in Table 2, a total of 74 occurrences out of 1,803 verbs have been pinpointed. This results in an ME density of 4.10; that is, approximately one out of every twenty-five verbs in the corpus is employed with a motion meaning. Besides, when examining the typology of MEs employed (M1 vs M2), none of the ME occurrences have been identified as M2.

This consistency detected in motion typology contrasts with the variability observed in the frequency analysis. As can be inferred from Table 3, MEs are not used uniformly, but there is a notable preference for some of them –mainly, *go*, *put* and *come*– resulting in virtually half of the total of occurrences. The use of these three MEs lemmas contrasts sharply with the little occurrences of eight MEs (about half of the verb lemmas), appearing only once.

Table 3: Motion verbs distribution

Verb	Element of motion	No. of occurrences	Relative frequency over the total number of verbs (in %)
Add	Caused-motion	4	0.22
bounce	Manner	1	0.06
Bring	Caused-motion	6	0.33
Come	Path	8	0.44
continue	Path	1	0.06
follow	Path	3	0.17
Get	Neutral	2	0.11
Go	Neutral	23	1.27
Hurry	Manner	1	0.06
Lead	Path	3	0.17
Move	Neutral	4	0.22
pursue	Path	1	0.06
Put	Caused-motion	8	0.44
Reach	Path	2	0.11
Run	Manner	1	0.06
Send	Caused-motion	3	0.17
Skip	Manner	1	0.06
Slow	Manner	1	0.06
Take	Caused-motion	1	0.11

Concerning their MOTION semantic roles, the classification into four categories (Path, Manner, Caused-motion, and Neutral) allows the observation of apparent pattern differences regarding the frequency of appearance of various MOTION elements. In general, as shown in Table 4, Manner verbs are barely used, with a total of 5 occurrences, and their frequency contrasts with Neutral verbs (*go, move*, and *get*), which prevail with 29 cases. However, this difference is only perceived when examining the number of occurrences, as in the number of lemmas per category, Path, Manner, and Caused-motion verb lemmas share a similar distribution, albeit Neutral verbs are less varied in number.

To conclude the preliminary analysis of the sample, the metaphor evidence on MEs was calculated (see Table 4). Out of the 74 MEs identified, 94.59% are used figuratively. This is quite revealing in the sense that the use of MMEs is significantly more frequent than non-metaphorical MEs.

Table 4: Number of MEs and their metaphor-related use concerning their motion semantic role

	Path		Manner		Caused-motion		Neutral	
	Overall	MRWs	Overall	MRWs	Overall	MRWs	Overall	MRWs
Tokens	18	18	5	4	22	21	29	27
Lemmas	6	6	5	4	5	4	3	3

Moving on to the exploration of the impact of L1 typology, this variable has been examined considering participants' lexical diversity and production of MEs. In this sense, as shown in Table 5, the three cohorts present differences regarding the number of lexical and ME tokens, lemmas, diversity, and density concerning metaphorical and non-metaphorical uses.

Regarding their lexical diversity, L1 satellite-framed English L2 speakers (i.e., L1 speakers of Swedish, German, and Dutch) present the highest ratio per speaker of tokens –i.e., the total number of utterances–, and MME density– i.e., the number of MMEs divided by the number of ME tokens– in comparison to the rest of groups, who share similar results.

These speakers also showed a higher ratio of ME occurrences (both tokens and types) per speaker. Their average number of tokens per speaker is roughly double the production of the other cohorts. Besides, the average number of lemmas per speaker nearly doubles that of NSs and practically triples the output of the other L1 verb-

framed group. Concerning the other two groups, despite their more homogeneous production in tokens, the NS group's production rate practically doubles the number of ME lemmas employed by L1 verb-framed speakers.

Finally, considerable variations are also observed when examining ME diversity, i.e., the number of ME types divided by the number of ME tokens; and density –the number of ME tokens divided by the total number of tokens produced. As can be observed in Table 5, L1 English speakers' ME diversity surpasses that of the other two groups as opposed to L1 satellite-frame speakers' prevalence in ME density. Looking at inner differences in the above-mentioned measures across the three participant groups, L1 verb-framed speakers' proximity in their figures is remarkable.

In the case of MMEs, the group results yield a nuanced picture at first sight. However, some salient differences are detected among the three groups regarding the proportion of MEs tokens and lemmas per speaker. Although the three participant groups' metaphor density points to a pervasive role of metaphor in their speech, the greatest abundance is detected in the L1 satellite-framed speakers.

Table 5 shows an overview of the differences concerning learners' L1 typology. This includes measures of lexical production and MV realizations presented in raw data and by speaker (i.e., raw data divided by the total number of speakers).

Table 5: A summary of the differences regarding learners' L1 typology

			Satellite-framed (NSs)	L1 satellite-framed	L1 verb-framed
Lexical production	Tokens	Raw data	1,976	3,314	5,228
		By speaker	494	828.5	580.90
	MRWs	Tokens	271	510	697
		Density (in %)	13.71	15.39	13.33
MV realizations	Tokens	Raw data	14	34	26
		By speaker	3.5	8.5	2.89
	Lemmas	Raw data	8	11	12
		By speaker	2	2.75	1.33
	Diversity	Raw data	0.57	0.32	0.46
	Density	In %	0.71	1.02	0.50
	MMEs	Tokens	13	33	24
		Density (in%)	92.85	97.05	92.31

4.2. Objective b: To determine when, and with what discursive-pragmatic purposes, MMEs are employed by NSs and L2 learner speakers of English

To address this objective, the two activities that were performed in the seminars (pitch and discussion) were used to frame the context of MMEs production. Besides, the MME production has been classified according to the main Conceptual Metaphors identified in participants' speech. Thirdly, an analysis exploring the discursive-pragmatic functions that MMEs entailed was carried out following Christie's (2002) classification of L2 classroom registers. In general, it seems that participants resort to MMEs mainly to describe their learning experience in terms of MOTION. At first, five Conceptual Metaphors (CM) have been identified, although the use of two of them, EMOTION IS MOTION and LIFE IS A JOURNEY, are residual and determined directly by the immediate context in which they are uttered. Therefore, this analysis will focus exclusively on the three CMs more systematically found in the seminars:

LEARNING/COGNITION IS MOTION (knowledge as space): the target domain is used to typically profile the source domain of MOTION in academic English, where the learning process is constructed metaphorically as a path. The following example shows how thinking up ideas in the learning experience is constructed as arriving at the end of a path.

(3) *I will try to explain better how we* ***reached*** *that idea of creating this app what was exactly the payment* <ELO; L1=Spanish; L2 level=C1>

CAREER IS MOTION (professional life is a journey): the target domain of *career* is characteristically found in the pitches, where speakers talk about their professional experience. The following excerpt illustrates how *career* is understood as a horizontal path, a trajectory that one must take, and in which one must advance.

(4) *when they finish university students are usually told to* ***follow*** *a bunch of determined career (.) you should go either for consulting or banking they say my name is Marc and I consider myself as one as one of those students who decide to* ***get off*** *that road and choose by themselves* <EMH; L1=Spanish; L2 level=C2>

SPEECH ACTIVITY IS MOTION (speech as movement): speech activity is constructed metaphorically in terms of movement along a path that can be traveled back and forward. This CM likens *speech activity* to movement towards or away from the goal of communication (example 5), or to movement towards a speech act of the interaction in example 6. In this study, this CM is usually found when students provide feedback to their peers on their presentations realized:

(5) *you did right maybe you can **put** that at the beginning [of your presentation] and make it a bit more flowing you know* <ENV; L1=Dutch; L2 level=C2>

(6) *that's why when I was five I thought I could make an association but my sister **brought** me bad news I sucked at it* <EMS; L1=Spanish; L2 level=C2>

In figures, the sample consists of 66 MME utterances, employed in the two main activities carried out in the seminars in a balanced way: 30 MMEs occur in the pitch-delivery activity while 36 occurrences are found in constructive feedback discussion. Table 6 provides detailed information regarding the relationship between conceptual frames and the tasks performed.

Table 6: A classification of CM constructions regarding the nature of the task.

Conceptual Metaphor	Element of movement	PITCH	DISCUSSION		Total
			Instructional	Regulative	
LEARNING/COGNITION IS MOTION (n = 10)	Path	2	4	-	6
	Manner	-	1	-	1
	Caused-motion	2	1	-	3
	Neutral	-	-	-	-
CAREER IS MOTION (n = 27)	Path	9	1	-	10
	Manner	-	-	-	-
	Caused-motion	5	1	-	6
	Neutral	9	1	1	11
SPEECH ACTIVITY IS MOTION (n = 30)	Path	-	1	-	1
	Manner	-	3	-	3
	Caused-motion	3	8	-	11
	Neutral	-	3	11	14
Total		**30**	**24**	**12**	**66**

However, a closer look at the data soon reveals noteworthy differences among the prevailing CMs in both activities. For instance, in pitches, most occurrences are Neutral or Path verbs structured via CAREER IS MOTION. In turn, discussions –where speakers give feedback on the pitches– are dominated by CM realizations where the conceptualization of the learning experience is described as SPEECH ACTIVITY IS MOTION and characterized by a prevalence of Neutral verbs. The third CM (LEARNING/COGNITION IS MOTION), with a predominance of Path verbs, is distributed somehow uniformly across both activities.

Besides, within the discussion activity, MMEs seem to be employed with two distinct purposes that correspond to the two main registers detected in the L2 (Christie, 2002) and CLIL (Llinares et al., 2012) classroom settings: to provide feedback or construct knowledge (i.e., instructional register) and to manage group work (i.e., regulative register). In general, as shown in Table 7, the dialogic interaction is characterized by unbalanced register usage, where instructional register surpasses regulative functions. In this regard, it is quite salient that most Neutral verbs are devoted to controlling the group, whereas Caused-motion, Manner, and Path verbs are restricted to the instructional register.

Table 7: An overview of conceptual frames concerning learners' L1 typology

	L1 typology			
Conceptual Metaphor	Satellite-framed (NSs)	satellite-framed	verb-framed	N
LEARNING/COGNITION IS MOTION	2	5	3	**10**
CAREER IS MOTION	8	6	13	**27**
SPEECH ACTIVITY IS MOTION	2	22	5	**29**
Total	**12**	**33**	**21**	**66**

Moreover, L1 verb-framed speakers' prevailing use of MMEs occurs in the pitches, where CAREER IS MOTION is primarily constructed given the activity instructions –i.e., presenting their career path. In contrast, there is a notable production of MMEs under the construction of SPEECH ACTIVITY IS MOTION by L1 satellite-framed L2 speakers of English in constructive feedback discussion contexts, where, again, the task typology is expected to drive students to somewhat structure their MMEs via this CM.

Table 8: Types of interaction and register of use regarding speakers' L1 typology

	Pitch	Discussion		
L1 typology		Instructional register	Regulative register	N
Satellite-framed (NSs)	6	6	-	6
Satellite-framed	10	14	9	23
Verb-framed	14	4	3	7
Total	**30**	**24**	**12**	**36**

Finally, Table 8 provides further information on how speaker groups govern these interactions. As is shown, MME occurrences seem to be somehow equally distributed across pitch-delivery and feedback-discussion interactions. However, L1 typology differences are perceived when considering the activity in which and the function with which MMEs are employed. Regarding activity typology, whereas MMEs are extensively employed in pitch-delivery interaction by the verb-framed speakers, satellite-framed speakers abound in their use of MMEs in constructive feedback discussion interaction. As for their purpose of use, within the discussion activity, while the three groups employ MMEs with an instructional aim, L1 English speakers do not seem to resort to MEs to organize group-work dynamics (i.e., with a regulative function).

5. Discussion: motion verbs in context

The present study was designed to analyze MME occurrence in a specific academic setting: EMI seminars on marketing and business. As shown in the literature review, metaphor is pervasive in academic language, where it plays a key role in the description of abstract objects and processes. MMEs are particularly important, as MOTION is a frequent source domain (Ibarretxe-Antuñano & Caballero, 2014). As previously studied (Özçalışkan, 2005), typological differences persist in MMEs and, therefore, metaphorical motion can vary greatly across different languages and genres (Caballero, 2017). However, despite the abundant literature on MEs and the demonstrated impact of figurative language in contexts of English as Lingua Franca (EFL) (MacArthur & Littlemore, 2011; MacArthur et al., 2015; MacArthur, 2016b or Alejo-González, 2021) or English as a Foreign Language contexts (see MacArthur, 2010, or Nacey, 2017, for a review), in the particular case of EMI, to our knowledge, this phenomenon has not received much attention.

The first objective of this paper was to examine the prevalence of MEs in academic EMI seminars and the impact inter-typological differences may have on the production of MMEs. This objective was explored via two research questions. Research question 1 dealt with the identification of MME employment in EMI learner discourse. In light of the results, it can be stated that MEs are employed at a low rate across the seminars.

First, speakers' discourse is characterized by a small, shifting variety of MEs and a clear predominance of M1 verb usage. Given the purpose of the seminars, i.e., to deliver a business pitch with which students need to differentiate themselves from their peers, it would have been expected that students would resort to some linguistic strategies such as the use of figurative language, to convey abstract ideas, and/or M2, to spark creativity in their task performance. However, this assumption has only been partially confirmed: although metaphorical language and MEs are somewhat present in speakers' speech, no evidence of M2 realizations has been found.

There may be several possible explanations for the low use of MEs, in general, and the complete absence of M2, in particular. First, these findings support the idea that speakers, primarily L2 learners with varying levels of English proficiency communicating in an EMI environment, need to make themselves understood in the seminars. This may result in students with the highest proficiency level adapting their speech, avoiding the inclusion of more abstract or creative language that might hinder their peers' comprehension. Similarly, NSs seem to abandon some of their rhetorical style characteristics in the use of MMEs. The less varied and frequent MME usage by English NSs may be an accommodation strategy to increase efficiency in communication with non-native speakers (NNSs), which is in line with Communication Accommodation Theory (CAT), as speakers can use accommodative strategies and converge with their interlocutors in order to be clearly understood, to maintain face and relationships, especially in intercultural groups –see Gallois et al. (2005) for an overview on CAT. In fact, when CAT has been applied to interactions between NSs and NNSs, NSs seem to use convergence strategies and engage in the so-called 'foreigner talk', with shorter and simpler sentences, slower pace and articulated pronunciation (Zuengler, 1991). Therefore, a less varied and frequent use of MMEs by English NSs may be considered an accommodative strategy to increase efficiency in communication with non-native fellows.

A second rationale may be related to the lecturer's role in carrying out the seminars. The instructor's feedback on every student's pitch and monitoring of the correct development of the seminars is one of the primary sources of input in the classroom. At first, it seemed plausible that students would mirror their instructor's speech, especially when providing feedback and regulating groupwork. Lecturer's use of MMEs has been explored to shed light on this question. In comparison to students' production, lecturer's speech presents substantial differences: (i) a higher ratio of ME usage (11% vs 4.10%), (ii) a more varied usage of MEs (33 vs 19 lemmas), (iii) some evidence of Motion2 utterances, and (iv) lower employment of MMEs (66% vs 93.57%). Therefore, these data would allow us to reject this second explanation. However, further research in this regard is needed to explore this potential explanation.

Finally, a third factor that could explain the low ME occurrence is that MOTION might not be a predominant source domain in the specific context under study. The findings of the current investigation contribute to the need highlighted in other studies (Kövecses, 2002; Semino, 2005) to describe the metaphors used in the academic context. In this respect, it would be relevant to explore not only the extent to which other source domains are present in the corpus, but also how factors relating to the specific academic context such as student's linguistic background, topic covered (i.e., *aboutness*), nature of the task, and degree of interaction may determine the inclusion of MEs in EMI spoken discourse.

Research question 2 addressed the analysis of the identification of inter-typological differences in MME employment, and some differences have been recognized in this respect. The literature on motion events remarks that speakers tend to organize these events around a specific semantic component, resulting in two different lexicalization patterns (satellite-framed and verb-framed languages) with varying narrative styles (Slobin, 1996, 2004). Thus, L1 verb-framed speakers' employment of MEs was expected to differ from the other two groups, which shared the same L1 typology frame (Ibarretxe-Antuñano & Caballero, 2014). Furthermore, within the L1 satellite-verb category, it was also expected to find differences between NS and NNSs' speech insofar as L1 English speakers were expected to use a higher number of MMEs. However, minor and non-systematical differences have been found.

There may be several reasons behind this finding. A plausible interpretation is that the strongest typological differences in learners' rhetorical style are sort of 'blurred' in the interaction in the context of the classroom. According to Filipović & Hawkins (2019), bilinguals tend to maximize the use of common grammatical and lexical representations in the two languages in proportion to environmental factors. In this case, students seemed to avoid M2 verbs, presenting similar usage of MMEs. The interaction among peers in the classroom may be fostering to favor this behavior. In this sense, our study presents important implications for the field of motion events typology and a call to more research in which motion events are measured in real interaction, with different types of interlocutors (e.g., bilingual and monolingual). Moreover, the high levels of metaphor density but the lower ME occurrence in students' speech might indicate that students did not exploit metaphor to their full potential in seminars (cf. Alejo-González, 2022).

The second objective concerned the exploration of the pragmatic-discursive purposes that the realizations of MMEs fulfilled, and the factors contributing to their employment, which was specified into two research questions. Regarding research question 3, that is, the exploration of the purposes with which the source domain of MOTION was employed, it can be argued that (i) EMI students mainly use MMEs to talk about the learning experience in terms of three conceptual frames (LEARNING/COGNITION IS MOTION, CAREER IS MOTION and SPEECH ACTIVITY IS MOTION), (ii) the employment of these target domains is dependent on seminar activities, which seem to regulate the local peaks of metaphor density in both monologic and dialogic interaction (cf. Low et al., 2008; Alejo-González, 2021): CAREER IS MOTION conceptualizations predominate in pitches, while MMEs in the discussion context are structured via SPEECH ACT IS MOTION. Finally, in relation to whether the L1 typology of participants has any impact on the pragmatic-discursive goals (research question 4), the results seem to somehow relate the inclination to one target domain to speakers' L1 typology.

At first sight, MOTION conceptual frames seemed somehow related to speakers' L1 typology afresh: L1 English and verb-framed speakers prefer organizing their speech towards CAREER IS MOTION, whereas L1 satellite-framed speakers opt for SPEECH ACTIVITY IS MOTION. However, this disparity may be related to the contexts where MMEs are employed: L1 satellite-framed speakers mainly resort to MMEs in constructive feedback discussion activities, where they are asked to give feedback on their peers' performance. In contrast, L1 verb-framed speakers' MMEs occurrences mostly occur in pitch-delivery, in which students need to summarize their career trajectory. Considering these findings, it can be argued that the activity performed seems to play a more determinant role than speakers' mental lexicalization patterns for metaphorical conceptualizations to be realized.

Overall, findings on the pragmatic-discursive purposes of MMEs illustrate that EMI students use MMEs for evaluative purposes in academic talk, as demonstrated for lecturer's discourse (Low et al., 2008; Beger, 2011). However, regulative functions performed by EMI students are also recognized. The presence of the regulative register, generally unexpected from students, appears to be fostered by the type of activity performed: students are asked to work on their own, and they need a sort of internal group management to reach their goal. This evidence tallies with current research on CLIL classroom registers at the primary education level (Llinares et al., 2012), where students use the L2 to both regulate peer work and learn while creating new knowledge. This finding seems, therefore, to support the synergies between both content-based practice areas. All in all, considering the main findings of this paper, it may be noted that speakers' L1 typology may have some impact on metaphorical and non-metaphorical MEs usage in the specific context of EMI seminars, but the influence of other contextual factors such as task typology and the discourse functions CMs underlined in the topic of the assignment should not be disregarded.

6. Conclusion

This article has examined the verb use of MEs in a spoken academic context of EMI. The use of MMEs by 17 participants has been analyzed, finding some evidence of MMEs realizations and, in general, limited use of the domain of MOTION when talking about marketing and business. The analysis has provided some evidence that in the specific context of EMI seminars (i) individual variables (L1 typology of the different speakers, lecturer's input) do not seem to have a determining impact on the employment of MMEs, while (ii) contextual factors (the activities to be performed) have been shown to have a more salient role in the use of MMEs. Although it is based on a small sample of participants, the current study aims to contribute to extending

our knowledge of how EMI students behave linguistically in seminar talk. In this regard, ME realizations seem to adapt to the context of the seminar in which they are employed. Although ME usage is limited, motion is observed to be used by EMI students with two different discursive aims: for regulative and evaluative functions.

These results suggest that ME typologies, that is, the way speakers organize the salient information they pay attention to in discourse, is a key aspect to be explicitly taught about in EMI contexts. It is not just a question of learning certain verb items in metaphorical contexts but of using them appropriately to capture the speakers' own conceptualizations. Future research may shed some light on how to apply and implement these results in classroom settings.

Several limitations need to be acknowledged. First, the current investigation has only analyzed MME occurrences in students' speech. Therefore, an exploration of the lecturer's speech, with a key role in giving feedback, regulating groupwork and monitoring the activities in the classroom, should also need to be considered. In this sense, it would be interesting to compare the kind of language used by both interlocutors' roles and explore to what extent learners 'recycle' lecturer's linguistic expressions and mental conceptualizations. In addition, the scope of the study is limited by the number of participants. A more comprehensive sample including instructors and learners with different L1s, for instance, may contribute to further research. Moreover, the findings in this study only address verb usage, implying an overlook of the relevance of motion in this spoken discourse. Hence, further research should include other parts of speech (i.e., particles). Finally, these findings may not apply to other EMI contexts dealing with, for example, written discourse, discussion of different topics, distinct activity typology, or varying degrees of interaction. More research is required to better understand the interplay between metaphor and L2 academic discourse.

Acknowledgements

This research has been funded by the Spanish Ministry of Science and Innovation (AEI/FEDER Funds: METCLIL FFI2017-86320-R; CONESSO FFI2017-82460-P; MOTIV PID2021-123302NB-I00), the Spanish Ministry of Universities and the European Union-Next GenerationEU (María Zambrano Program MZ-240621), the Government of Extremadura (HUM022- 2021/00547/001), the Government of Aragon (Psylex H11-17R; MultiMetAR LMP143_21), the Iberus Campus (ICON action group) and the *Instituto Universitario de Investigación y Prospección Educativa de la Universidad de Extremadura.*

7. References

Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism*, *15*(2), 183–197.

Alejo-González, R. (2018). *The Place of Language in English-Medium Instruction* [Paper presentation]. Presented at the Congreso Internacional de Educación Bilingüe (CIEB), held on October 18th-21st, 2018, in Badajoz (Spain).

Alejo-González, R. (2022). Metaphor in the academic mentoring of international undergraduate students: The Erasmus experience. *Metaphor and Symbol*, *37*, 1-20.

Alejo-González, R., & García-Bermejo, V. (2019). "The manage of Two Kingdoms Must": An analysis of metaphor in two CLIL textbooks. In A. Piquer-Píriz, & R. Alejo-González (eds), *Metaphor in Foreign Language Instruction* (pp. 241–262). Berlin: Mouton de Gruyter.

Alejo, R., Piquer-Píriz, A.M., Castellano-Risco, I., Martín-Gilete, M., Fielden-Burns, L., MacArthur, F., Nacey, S., Philips, G, Krennmayr, T., Coelho, M., Littlemore, J., & Ädel, A. (2021). *METCLIL corpus v1*. https://www.sketchengine.eu/metclil-corpus-of-metaphor-in-academic-talk/

Alejo, R., Piquer-Píriz, A.M., Castellano-Risco, I., Martín-Gilete, M., Fielden-Burns, L., MacArthur, F., Nacey, S., Philips, G, Krennmayr, T., Coelho, M., Littlemore, J., & Ädel, A. (forthcoming). *METCLIL corpus v2*.

Beger, A. (2011). Deliberate metaphors? An exploration of the choice and functions of metaphors in US-American college lectures. *Metaphorik.de*, *20*, 39–60.

Berman, R. A. & Slobin, D.I. (eds.) (1994). *Relating Events in Narrative. A Crosslinguistic Developmental Study.* Hillsdale, NJ: Lawrence Erlbaum.

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics*, *21*(4), 553–571.

Caballero, R. (2017). Metaphorical motion constructions across genres. In I. Ibarretxe-Antuñano (ed.), *Motion and Space across Languages: Theory and Applications* (pp. 229–256). Amsterdam: John Benjamins.

Caballero, R., & Ibarretxe-Antuñano, I. (2015). From physical to metaphorical motion: A cross-genre approach. In V. Pirrelli, C. Marzi, & M. Ferro (eds.), *Word Structure and Word Usage. Proceedings of the NetWordS Final Conference, Pisa, March 30-April 1, 2015* (pp. 155–157). Available at http://ceur-ws.org.

Cadierno, T. (2004). Expressing motion events in a second language: A cognitive typological perspective. In M. Achard, & S. Niemeier (eds.), *Cognitive Linguistics, Second*

Language Acquisition, and Foreign Language Teaching (pp. 13–49). Berlin: Mouton de Gruyter.

Cadierno, T. (2008). Motion events in Danish and Spanish: A focus on form pedagogical approach. In D.K. Sabine, & D.R. Teun (eds.), *Cognitive Approaches to Pedagogical Grammar* (pp. 259–294). Berlin: Mouton de Gruyter.

Cadierno, T. (2017). Thinking for speaking about motion in a second language: Looking back and forward. In I. Ibarretxe-Antuñano (ed.), *Motion and Space across Languages: Theory and Applications* (pp. 279–300). Philadelphia: John Benjamins.

Cameron, L., & Stelma, J.H. (2004). Metaphor clusters in discourse. *Journal of Applied Linguistics*, *1*(2), 7–36.

Christie, F. (2002). *Classroom Discourse Analysis: A Functional Perspective*. London/ New York: Continuum.

Cifuentes-Férez, P. (2010). The semantics of the English and the Spanish motion verb lexicons. *Review of Cognitive Linguistics*, *8*(2), 233–271.

Dafouz, E., & Smit, U. (2020). *ROAD-MAPPING English Medium Education in the Internationalised University*. Charm: Springer.

Dalton-Puffer, C. (2003). Telling each other to do things in class: directives in content and language integrated classrooms. *VIEWZ: Vienna English Working Papers*, *12*, 3–24.

Dalton-Puffer, C., & Nikula, T. (2006). Pragmatics of content-based instruction: Teacher and student directives in Finnish and Austrian classrooms. *Applied Linguistics*, *27*, 241–267.

Filipović, L., & Hawkins, J. (2019). The complex adaptative system principles model for bilingualism: Language interactions within and across bilingual minds. *International Journal of Bilingualism*, *23*, 1223–1248.

Filipović, L., & Ibarretxe-Antuñano, I. (2015). Motion. In E. Dąbrowska, & D. Divjak (eds.), *Mouton Handbook of Cognitive Linguistics* (pp. 526–545). Berlin: Mouton de Gruyter.

Gallois, C., Gogay, T., & Giles, H. (2005). Communication accommodation theory: A look back and a look ahead. In W. B. Gudykunst (ed.), *Theorizing about Intercultural Communication* (pp. 121–148). Thousand Oaks, CA: Sage.

Hijazo-Gascón, A. (2021). *Moving Across Languages. Motion Events in Spanish as a Second Language*. Berlin: Mouton de Gruyter.

Hijazo-Gascón, A., Ibarretxe-Antuñano, I., & Guenbelzu, J. (2013). Clasificando los verbos de movimiento. ¿Qué piensan los hablantes?. In J.F. Val Álvaro, J.L. Mendívil

Giró, M.C. Horno Chéliz, I. Ibarretxe-Antuñano, A. Hijazo-Gascón, J. Simón, & I. Solano (eds.), *De la Unidad del Lenguaje a la Diversidad de las Lenguas* (pp. 361–368). Zaragoza: Prensas Universitarias de Zaragoza.

Herrmann, J.B. (2013). *Metaphor in Academic Discourse: Linguistic Forms, Conceptual Structures, Communicative Functions and Cognitive Representations*. Utrecht: LOT dissertation series.

Ibarretxe-Antuñano, I. (2009). Path salience in motion events. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, N. Nakamura, & Ş. Özçalışkan (eds.), *Crosslinguistic Approach to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin* (pp. 403–414). New York: Psychology Press.

Ibarretxe-Antuñano, I. (2017). *Motion and Space across Languages: Theory and applications*. Philadelphia: John Benjamins.

Ibarretxe-Antuñano, I., & Caballero, R. (2014). Una aproximación al estudio de los eventos de movimiento metafórico desde la tipología semántica y el género. *Anuari de filologia. Estudis de lingüística*, *4*, 139–155.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. London: Routledge.

Kövecses, Z. (2002). Cognitive-linguistic comments on metaphor identification. *Language and Literature*, *11*, 74–78.

Kyle, K., Crossley, S. A., & Jarvis, S. (2021). Assessing the validity of lexical diversity using direct judgements. *Language Assessment Quarterly*, *18*, 154–170.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

Lakoff, G. (1999). *Philosophy In The Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Larrañaga, P., Treffers-Daller, J., Tidball, F., & Gil-Ortega, M.C. (2011). L1 transfer in the acquisition of manner and path in Spanish by native speakers of English. *International Journal of Bilingualism*, *16*, 117–138.

Li, P., Eskildsen, S., & Cadierno, T. (2014). Tracing an L2 learner's motion constructions over time: A usage-based classroom investigation. *The Modern Language Journal*, *98*, 612–628.

Littlemore, J. (2001). The use of metaphor in university lectures and the problems that it causes for overseas students. *Teaching Higher Education*, *6*, 333–349.

Littlemore, J., & Low, G. (2006). *Figurative Thinking and Foreign Language Learning*. New York: Palgrave/ Macmillan.

Littlemore, J., Krennmayr, T., Turner, J., & Turner, S. (2014). An investigation into use at different levels of second language writing. *Applied Linguistics*, *35*, 117–144.

Llinares, A., Morton, T., & Whittaker, R. (2012). *The Roles of Language in CLIL*. Cambridge: Cambridge University Press.

Lorenzo, F. (2013). Genre-based curricula: Multilingual academic literacy in content and language integrated learning. *International Journal of Bilingual Education and Bilingualism*, *16*, 275–388.

Low, G., Littlemore, J., & Koester, A. (2008). Metaphor use in three UK university lectures. *Applied Linguistics*, *29*, 428–455.

Macaro, E. (2018). *English Medium Instruction.* Oxford: Oxford University Press.

MacArthur, F. (2010). Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review*, *23*, 155–173.

MacArthur, F. (2016a). When languages and cultures meet: Mixed metaphors in the discourse of Spanish speakers of English. In R.W. Jr. Gibbs, (ed.), *Mixing Metaphor* (pp. 133–154). Amsterdam: John Benjamins.

MacArthur, F. (2016b). Overt and covert uses of metaphor in the academic mentoring in English of Spanish undergraduate students at five European universities. *Review of Cognitive Linguistics*, *14*(1), 23–50.

MacArthur, F., Krennmayr, T., & Littlemore, J. (2015). How basic is UNDERSTANDING IS SEEING when reasoning about knowledge? Asymmetric uses of SIGHT metaphors in office hours' consultations in English as academic lingua franca. *Metaphor and Symbol*, *30*(3), 184–217.

MacArthur, F., & Littlemore, J. (2011). On the repetition of words with the potential for metaphoric extension in conversations between native and non-native speakers of English. *Metaphor and the Social World*, *1*(2), 201–238.

Matlock, T. (2004). The conceptual motivation of fictive motion. In G. Radden, & K.U. Panther (eds.), *Motivation in Grammar* (pp. 221–248). Berlin: Mouton de Gruyter.

Moghadam, M.S., & Samar, R.G. (2020). Metaphor in second language academic writing. *Language awareness*, *29*, 255–271.

Nacey, S. (2017). Metaphor comprehension and production in a second language. In E. Semino, & Z. Demjén (eds.), *The Routledge Handbook on Metaphor and Language* (pp. 503–515). London: Routledge.

Navarro, S., & Nicoladis, E. (2005). Describing motion events in adult L2 Spanish narratives. In D. Eddington (ed.), *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 102–107). Somerville, MA: Cascadilla Proceedings Project.

Negueruela, E., Lantolf, J.P., Rehn Jordan, S., & Gelabert, J. (2004). The "private function" of gesture in second language speaking activity: A study of motion verbs and gesturing in English and Spanish. *International Journal of Applied Linguistics*, *14*, 113–147.

Nikula, T. (2005). English as an object and tool of study in classrooms: Interactional effects and pragmatic implications. *Linguistics and Education*, *16*, 27–58.

Özçalışkan, Ş. (2005). Metaphor meets typology: ways of moving metaphorically in English and Turkish. *Cognitive Linguistics*, *16*(1), 207—246.

O'Dowd, R. (2018). The training and accreditation of teachers for English Medium Instruction: An overview of practice in European universities. *International Journal of Bilingual Education and Bilingualism*, *21*, 553–563.

Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and English Medium Instruction. *TESOL Quarterly*, *52*, 497–515.

Robinson, P., & Ellis, N.C. (2008). Conclusion: Cognitive linguistics, second language acquisition and L2 instruction – issues for research. In P. Robinson, & N.C. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 489–545). New York: Routledge.

Semino, E. (2005). The metaphorical construction of complex domains: The case of speech activity in English. *Metaphor and Symbol*, *20*, 35–69.

Semino, E. (2008). *Metaphor in Discourse*. Cambridge: Cambridge University Press.

Slobin, D. (1996). From "thought and language" to "thinking for speaking". In J.J. Gumperz, & S.C. Levinson (eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge University Press.

Slobin, D. (1997). Mind, code, and text. In J. Bybee, J. Haiman, & S. A. Thompson (eds.), *Essays on Language Function and Language Type: Dedicated to T. Givón* (pp. 437–467). Amsterdam: John Benjamins.

Slobin, D. (2004). The many ways to search for a frog: linguistic typology and the expression of motion events. In S. Strömqvist, & L. Verhoeven (eds.), *Relating Events in Narrative: Typological and Contextual Perspectives* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum Associates.

Steen, G.J., Dorst, A.G., Herrmann, B., Kaal, A.A., Krennmayr, T., & Pasma, T. (2010). *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. Amsterdam: John Benjamins.

Strömqvist, S., & Verhoeven, L. (eds.) (2004). *Relating Events in Narrative: Typological and Contextual Perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

Talmy, L. (1991). Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, *17*, 480–519.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge, MA: The MIT Press.

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the 'what' to the 'why' in second-language research. In H. Giles, J. Coupland, & N. Coupland (eds.), *Contexts of Accommodation* (pp. 223–244). Cambridge: Cambridge University Press.

Computational analysis of adjuncts in ASD-STE100 for the NLP parser ARTEMIS

Francisco José Cortés-Rodríguez
"Andrés Bello" Institute of Linguistics
Universidad de La Laguna
fcortes@ull.edu.es

Carolina Rodríguez-Juárez
Research Institute of Text Analysis and Applications
Universidad de Las Palmas de Gran Canaria
carolina.rodriguez@ulpgc.es

Abstract

Texts written in Simplified Technical English present semantic and syntactic restrictions with respect to standard English that should be considered in the implementation of natural language processing systems dealing with controlled natural languages. This paper explores the syntax and semantics of adjuncts in a controlled natural language, namely, the Simplified Technical English (ASD-STE100) used in the Airbus corpus with a view to observing how the peculiarities identified in the behaviour of Airbus adverbials should be reflected in the computational treatment of these constituents. Thus, our main aim is to contribute to the computational implementation of the prototype ARTEMIS by designing the parsing rules and the catalogue of feature-bearing matrixes that encode the grammatical constraints of Airbus adjuncts. The parsing rules reflect the positional preferences of the various peripheral types and capture the semantic variability of adjuncts in the corpus. In addition, they provide a weight factor that predicts the scale of markedness of these constituents with respect to the different positions that they occupy in the clause. On the whole, these properties offer a precise description of the syntactic features of adverbials that will facilitate their automatic processing.

Keywords: adjuncts, Airbus corpus, ARTEMIS, ASD-STE100, parsing rules.

Resumen

Los documentos escritos en Inglés Técnico Simplificado (ASD-STE100) muestran determinadas restricciones semánticas y sintácticas que deben tenerse en cuenta para la creación de sistemas de procesamiento de lenguaje natural aplicados a lenguajes controlados naturales. En este trabajo se exploran los rasgos sintácticos y semánticos

de los adverbiales en uno de estos lenguajes controlados naturales, el ASD-STE100, utilizado en el corpus Airbus. El propósito de este análisis es establecer el modo en que las peculiaridades del comportamiento de los adverbiales en este corpus deben encontrar reflejo en el tratamiento computacional de dichos constituyentes. Con ello pretendemos contribuir a la implementación del prototipo ARTEMIS mediante el diseño de las reglas para el análisis sintáctico (reglas de parseado) y las matrices atributo-valor de este prototipo, las cuales codifican las restricciones gramaticales de los adjuntos del corpus de Airbus. Las reglas de parseado reflejan las preferencias posicionales de los diferentes tipos de periferias adverbiales y capturan la variabilidad semántica de los adjuntos en el corpus. Además, ofrecen un factor de peso que predice la escala de marcado de estos constituyentes con respecto a las diferentes posiciones que ocupan en la cláusula. Estas propiedades ofrecen en conjunto una descripción precisa de los rasgos sintácticos de los adverbiales que facilitará el procesamiento automático de los mismos.

Palabras clave: adjuntos, ASD-STE100, ARTEMIS, corpus Airbus, reglas de parseado.

1. Introduction

The complexity of natural languages presents a problem for computers when trying to process and understand them due to the ambiguity and the implicit meaning that these languages present. Formal languages emerged as a possible solution to this issue but have turned out to be difficult to understand by domain specialists as they "cause a cognitive distance to the application domain that is not inherent in natural language" (Schwitter, 2010: 1113). Thus, an intermediate solution between these two types of languages is the use of controlled natural languages (CNLs) since they have been designed to reduce the ambiguity and complexity of natural languages (Schwitter, 2010: 1113). These constrained natural languages belong to different environments and disciplines, and are constructed in accordance with "a well-defined subset of a language's grammar and lexicon" (Kittredge, 2003, quoted in Kuhn, 2014: 122). Besides, they include the specific technical vocabulary needed in a particular domain.

In 2010, Schwitter claimed that it was "an exciting time to work on controlled natural languages" (p. 1120). In fact, over the past years, numerous studies have focused on the development of machine-oriented controlled natural languages (e.g. Attempto Controlled English (ACE, Fuchs et al., 2008), Processable English (PENG, White & Schwitter, 2009) or Computer Processable Language (CPL, Clark et al., 2010), to name a few) that are aimed to "improve the translatability of technical documents (e.g. machine translation (...)) and the acquisition, representation, and processing of knowledge (e.g. knowledge systems (...)) and in particular for the Semantic Web"

(Schwitter, 2010: 1114). In this paper, however, we focus on another category of CNLs, namely, human-oriented CNLs whose aim is to facilitate the comprehensibility and the readability of technical texts (Schwitter, 2010: 1113), such as the ones contained in the Airbus corpus. The Airbus corpus is made up of maintenance documents written according to the standards and rules specified in the document "ASD Simplified Technical English Specification (ASD-STE100)" (2017). The need to develop this simplified language (STE, Simplified Technical English) emerged in the field of the aerospace and defense industry since the writers of these technical documents had to guarantee that readers, mainly airlines staff (of whom 80% are non-native speakers of English), would be able to understand maintenance and operation documents to guarantee the aircraft availability without putting human lives at risk (https://asd-ste100.org/about.html). Such has been the success of STE that industries from areas as diverse as language services, professional translation and interpreting and the academic world are also using this STE (https://asd-ste100.org/index.html).

The issues of ASD-STE100 are constantly being updated to catch up with the technological evolution,[1] but its structure is stable and consolidated, with a set of Writing Rules (Part 1) covering aspects of grammar and style, and a Dictionary of controlled vocabulary (Part 2) that lists the words that are approved and, as a result, can be used (https://asd-ste100.org/about.html). Just to illustrate why this technical language is regarded as a "simplified" language, let's show a few representative examples taken both from the ASD-STE100 specification document and from the Airbus corpus. ASD-STE100 (2021) restricts the parts of speech of a particular word that can be used. Thus, the word "test" can only be used as a noun and not as a verb:

1. STE: Test B is an alternative to test A.

   Non-STE: Test the system for leaks.

   STE: Do the leak test of the system / Do a test for leaks in the system.

   Airbus corpus: Wait for a minimum of 2 seconds before you launch the test.

   Phrasal verbs with idiomatic/abstract meanings cannot be used:

2. NonSTE: This compound can give off poisonous fumes.

   STE: This compound can release poisonous fumes.

   Airbus corpus: The forward kneeling manifolds open to release hydraulic flow …

1 The last updated version of ASD-STE100 dates back to April 30, 2021, ISSUE 8 (https://asd-ste100.org/).

STE recommends not omitting verbs or subjects, because the reader will not understand what the action is or what you are referring to:

3. Non-STE: Rotary switch to INPUT.

   STE: Set the rotary switch to INPUT.

   Airbus corpus: Put the rotary switch (5) in position on control panel 11VU (3).

4. Non-STE: If installed, remove the shims.

   STE: IF shims are installed, remove them.

In procedural writing, for example, ASD-STE100 recommends writing short sentences with a maximum of 20 words, whereas in descriptive writing a maximum of 25 words are allowed. In instructions, the imperative form must be used and only one instruction should be included in one sentence:

5. Non-STE: Set the TEST switch to the middle position and release the SHORT-CIRCUIT TEST switch.

   STE: A. Set the TEST switch to the middle position.

   B. Release the SHORT-CIRCUIT TEST switch.

6. Airbus corpus:

   Release the MLG 1M electrical-harness (21) from the upper bracket.

   Remove the upper bracket (22) from the kneeling cylinder (14).

Natural Language Processing (NLP) systems need to be adapted to the semantics and syntax of CNLs for a successful automated natural-language processing. Thus, with this research we aim to contribute to the development of an NLP prototype called ARTEMIS (Automatically Representing Text Meaning via an Interlingua-based System) (Periñán-Pascual, 2013a/b; Periñán-Pascual & Arcas-Túnez, 2014) that has been designed to obtain the syntactic and semantic representation of linguistic structures and that has been implemented as a parsing device within the lexico-conceptual knowledge base FunGramKB (Periñán-Pascual, 2012, 2013a/b; Periñán-Pascual & Arcas-Túnez, 2014). At the moment, the ARTEMIS parser is being bench tested for the controlled natural language ASD-STE100 with the idea of achieving the parsing of standard English eventually (Fumero-Pérez & Díaz-Galán, 2019: 152). In fact, this paper complements previous studies focused on the implementation of the computational grammar of ARTEMIS through the development of the production rules

(lexical, syntactic and constructional) that are stored in the Grammar Development Environment (GDE) and that take part in the parsing process of linguistic expressions. In particular, these studies have addressed the formalised description of constructional and non-constructional meaning in ASD-STE100, and the present research on ASD-STE100 adjuncts contributes to the further development of the GDE by exploring the semantic and positional variability of adverbs in this CNL.[2]

In this article, we aim to enrich the investigation carried out so far by designing the Airbus syntactic rules for the internal realization of adverbials as well as for the localization of the peripheries at the different layers of the constituent structure of the clause. These rules will reflect the notable mobility and semantic variability of these components, aspects that have been registered and analysed in a previous quantitative corpus-based study (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming). We also aim to provide the set of Attribute-Value Matrixes (AVMs) for ASD-STE100 adverbial units that are configurated as a list of descriptors (attributes) and values that establish the semantic restrictions and/or selection constraints that cannot be directly retrieved from the modules of the knowledge base FunGramKB in which the parser is implemented; in particular, they cannot be obtained from the Lexicon (*Aktionsart* ascription, macrorole assignment, variables, logical structures, etc.), the Grammaticon (inventory of grammatical constructions) or the Ontology (hierarchy of conceptual units). These AVMs have additionally been enriched with the addition of the attribute "weight" that predicts the scale of markedness of nuclear and core adjuncts per position in the clause, thus offering a precise description of their syntactic representation that will facilitate the automatic processing of Airbus adjuncts in NLP tasks.

The remainder of the paper is structured as follows. Section 2 offers a brief overview of the parser ARTEMIS and of the lexico-conceptual knowledge base FunGramKB in which it is implemented. It additionally presents a general account of the grammatical models on which FunGramKB and ARTEMIS are grounded: Role and Reference Grammar and the Lexical Constructional Model. In section 3, we describe the corpus and the methodology followed in this research. In section 4, we

2 Examples of the investigation carried out so far into the design of the production rules for ARTEMIS and, more specifically, into the formalised description of ASD-STE100 are the following: complex grammatical structures withing the Nucleus layer (Cortés-Rodríguez & Mairal-Usón, 2016), DO-auxiliary insertion (Díaz-Galán & Fumero-Pérez, 2016), *yes-no* interrogative sentences (Martín-Díaz, 2017), simple clauses (Díaz-Galán, 2018), *wh*-interrogative sentences (Martín-Díaz, 2018), phrasal constituents (Cortés-Rodríguez & Rodríguez-Juárez, 2018), function words (Fumero-Pérez & Díaz-Galán, 2019), adverbials (Cortés-Rodríguez & Rodríguez-Juárez, 2019), adverbial complex sentences (Martín-Díaz, 2019), non-propositional meaning (Díaz-Galán & Fumero-Pérez, 2020), subordinate clauses (Martín-Díaz & González-Orta, 2020), non-peripheral complex sentences (González-Orta & Martín-Díaz, 2022).

include our main contributions by propounding the catalogue of AVMs for adjuncts and nuclear and core modifiers, as well as the attributes for "AdjunctRole", "concept" and "weight" (subsection 4.1), and by outlining the rules for the internal realization of peripheries (adjuncts, nuclear peripheries and level-1 peripheries) (subsection 4.2), and for the localization of peripheries in the layers of the layered structure of the clause (subsection 4.3). Section 5 wraps up the general contributions and conclusions of our research.

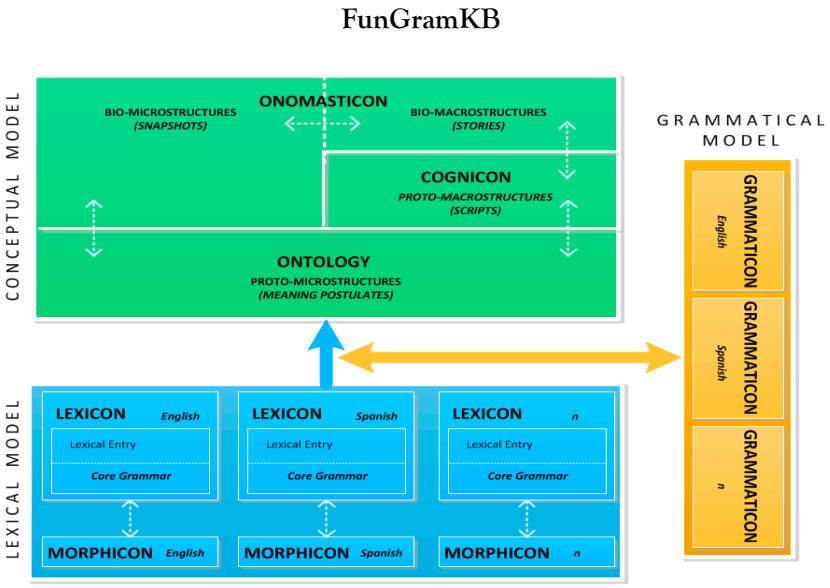
2. Background for the analysis: an overview of ARTEMIS and FunGramKB

Since the final aim of this research is to provide the production rules and Attribute-Value Matrixes (AVMs) that are needed for the computational parsing of Airbus adjuncts (also referred to as peripheral constituents in the remainder of this paper) in ARTEMIS in sentences like "*Lightly* loosen the axle nut (6)" or "*After 15 minutes*, do a leak check of these areas", we need to offer a brief overview of the parsing device ARTEMIS and of the knowledge base FunGramKB from which the NLP prototype extracts the relevant information for the effective semantic and syntactic parsing of sentences. FunGramKB and ARTEMIS are framed within the grammatical theory of Role and Reference Grammar (RRG) (Van Valin & LaPolla, 1997; Van Valin, 2005, 2008) and the Lexical Constructional Model (LCM) (Mairal-Usón & Ruiz de Mendoza, 2008; Ruiz de Mendoza & Mairal-Usón, 2008; Ruiz de Mendoza, 2013; Ruiz de Mendoza & Galera, 2014). RRG is a current theory that accounts for an integrated description of grammar, meaning and function, and whose descriptive potential has been highlighted by scholars such as Kailuweit et al. (2018). The LCM is a construction grammar that presents a strong constructional layered typology that has been adopted by the knowledge base FunGramKB. Thus, in this section we will also outline the most relevant aspects for our research concerning the theoretical tenets of RRG and the LCM.

ARTEMIS (Automatically Representing Text Meaning via an Interlingua-based System) is an NLP prototype that has been implemented as a parsing device within the multipurpose[3] lexico-conceptual knowledge base FunGramKB and that can generate the syntactic and semantic representation of English sentences. FunGramKB stores conceptual, constructional, lexical and morphological information about the English language in independent but interrelated modules (see Figure 1):

3 The knowledge base FunGramKB is said to be multipurpose since it is multilingual in the sense that it can be used with several natural languages, and it can be implemented in different NLP tasks, such as machine translation, information retrieval and automatic summarising, dialogue-based applications such as question-answering, etc. (Periñán-Pascual & Arcas-Túnez, 2014).

Figure 1: The FunGramKB architecture (http://www.fungramkb.com/)



The required linguistic input that FunGramKB needs for the development of the lexical module (Lexicon) has been arranged in the knowledge base following the formal principles of the functional theory of RRG with regard to the lexical representation of predicates. Thus, in the Core Grammar component of the Lexicon in FunGramKB we can find the attributes that are used in the automatic building of the semantic and syntactic representation of sentences in ARTEMIS: *Aktionsart* ascription (verb class), macrorole assignment (Actor/Undergoer), status of variables, inventory of argumental constructions, etc. The development of the grammatical module in FunGramKB (Grammaticon) is mostly grounded on the constructional view of the LCM, whose layered structure of meaning construction (argumental level-1 constructions, implicational level-2 constructions, illocutionary level-3 constructions and discursive level-4 constructions) has allowed the integration of constructional meaning into RRG to deepen semantic processing (Periñán-Pascual, 2013a: 206). Similarly, the LCM notion of construction has been claimed to be more adequate for the computational requirements of the parser ARTEMIS (Periñán-Pascual, 2013a; Luzondo-Oyón & Ruiz de Mendoza, 2015; Fumero-Pérez & Díaz-Galán, 2017). Thus, the term "construction" is reserved for those constructions whose meaning is larger than the meaning of the building blocks conforming it and that are stored in the Grammaticon in FunGramKB. Structures whose meaning is fully compositional are, on the contrary, stored in the Lexicon of FunGramKB and should be referred to as "kernel constructs" as proposed by Luzondo-Oyón & Ruiz de Mendoza (2015).

ARTEMIS, as a syntactic and semantic parser, resorts to FunGramKB to obtain the lexical, semantic, syntactic and constructional information that is needed to transduce fragments of language into their conceptual logical structure (CLS) and syntactico-semantic representation, as represented in Table 1:

Table 1: The process of understanding natural language in ARTEMIS: phases and components (Periñán-Pascual & Arcas-Túnez, 2014)

ARTEMIS ARCHITECTURE			
PHASES	**COMPONENTS/MODULES**		
INPUT TEXT + PRE-PROCESSING	Lemmatisation and tagging of word tokens		
GRAMMAR BUILDING	**GRAMMAR DEVELOPMENT ENVIRONMENT (GDE):** **2 THEORETICAL CONSTRUCTS**		
	1. Representation of feature-based structures as **Attribute-Value Matrixes** (AVMs) for grammatical units		
	2. Feature-based **production rules**		
	Syntactic rules: they build the enhanced model of the Layered Structure of the Clause (LSC).	**Constructional rules:** they embed constructional schemata into the enhanced LSC (FunGramKB L1-Constructicon).	**Lexical rules:** they provide the morphosyntactic and semantic information of tokens (FunGramKB Lexicon and Ontology).
SYNTACTIC PARSING	Generation of parse trees from the given input sentence		
CONCEPTUAL LOGICAL STRUCTURE (CLS) EXTRACTION	**CLS CONSTRUCTOR** Enhanced text meaning representation of RRG logical structures		
COREL (CONCEPTUAL REPRESENTATION LANGUAGE) SCHEME REPRESENTATION	**COREL SCHEME BUILDER** The CLS is modelled into a COREL scheme (formal language that formalises conceptual knowledge in FunGramKB).		

We have used the example provided by Fumero-Pérez and Díaz-Galán (2017: 38-41), "Louise baked a cake for the kids", to illustrate the process that is followed in ARTEMIS in the *understanding* and transformation of natural language input into its equivalent grammatical and semantic structures. The first stage automatically separates the sentence components and assigns an Attribute-Value Matrix to each of them in which their semantic and morphosyntactic features are listed. Table 2 illustrates the output of this process for the word "baked":

Table 2: Attribute-Value Matrix of the word token "baked" (Fumero-Pérez & Díaz-Galán, 2017: 39)

INPUT TEXT	Louise had baked a cake for the kids.
GRAMMAR BUILDING	**1. Attribute-Value Matrix of the word token "baked"** Form baked Lemma bake POS verb Tense past Concept +BAKE_00

The second phase in the Grammar Building process involves the activation of production (syntactic, constructional and lexical) rules. Their task is, firstly, to "ascribe an appropriate syntactic/semantic contour to the sentences" (Fumero-Pérez & Díaz-Galán, 2017: 39) and, secondly, to generate a syntactic tree (syntactic parsing phase). Table 3.1 shows the feature-based syntactic rule that includes the semantic and syntactic information contained in the Core Grammar of the verb "bake". The *Aktionsart* of the predicate "bake" is active accomplishment (ACA) and the verb takes 2 variables, x (theme) and y (referent), that must fulfil the selection restrictions of the theme being +HUMAN_00 and the referent +FOOD_00:

Table 3.1: Feature-based syntactic rule for the predicate "bake" (Fumero-Pérez & Díaz-Galán, 2017: 39)

GRAMMAR BUILDING	Syntactic rule for "bake": Headword bake Core grammar CLS [Aktionsart ACA] Variables Type x Role theme Macrorol Phrase Syntax Prefer +HUMAN_00 Type y Role referent Macrorol Phrase Syntax Prefer +FOOD_00

Table 3.2 illustrates the constructional rule for the verb "bake" when it participates in the For-Benefactive construction (FBEN).[4] The rule accounts for the fact the construction has added the constituent "for the kids", and this requires the inclusion of a new variable "w" with the role of Beneficiary and functioning as an argument adjunct. The construction has also modified the original aspectual value of "bake" (ACA), which is now a causative accomplishment (CACC). The causativity imposed by the FBEN construction triggers another change in the thematic role assigned to the variable "x", which is now an agent and not a theme (as shown in Table 3.1). Besides, the rule includes a realization constraint as regards the type of preposition that can introduce the prepositional phrase ("for") and a selection restriction,

4 The Core Grammar of the verb points to the different types of construction in which a verbal predicate can appear. In the case of the predicate "bake", apart from participating in the For-Benefactive construction, it can also appear in the "Unexpressed Second Argument Construction (The kitchen smelled so wonderful while they were baking); the Instrument Subject Construction (This oven bakes wonderful bread); the Material Subject Construction (This flour bakes a delicious loaf) and the Benefactive Object Construction (She baked them a cake)" (Fumero-Pérez & Díaz-Galán, 2017: 39).

+ORGANISM_00, that is defined in the FunGramKB ontology as "an animal, plant, human or any other living thing" (Fumero-Pérez & Díaz-Galán, 2017: 40).

Table 3.2: Feature-based constructional rule for the predicate "bake" in the For-Benefactive construction (Fumero-Pérez & Díaz-Galán, 2017: 40)

	Constructional rule for the For-Benefactive construction
GRAMMAR BUILDING	L1-constr [Type FBEN] CLS [Aktionsart CACC] Variables [Type x] [Role agent] [Type y] [Type w] [Role beneficiary] [Phrase PP] [Syntax argument adjunct] [Prep for] [Prefer +ORGANISM_00] [COREL scheme *(e1:<EVENT>) (f1:w) Beneficiary)]

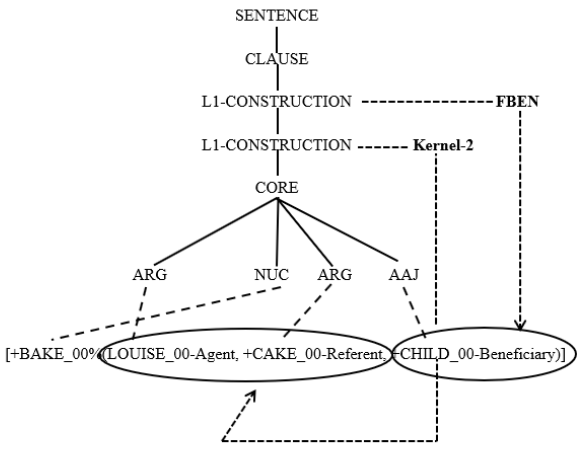
Finally, Table 3.3 presents the RRG lexical representation of the logical structure of the predicate "bake" in the example "Louise baked a cake for the kids". As can be seen, the lexical representation accounts for the causative nature of the event that is introduced by the FBEN construction (Fumero-Pérez & Díaz-Galán, 2017: 40):

Table 3.3: Lexical representation for the predicate "bake" in the For-Benefactive construction (Fumero-Pérez & Díaz-Galán, 2017: 40)

GRAMMAR BUILDING	**Lexical representation:** RRG logical structure [[**do'**(Louise, Ø)] CAUSE [BECOME **bake'**(cake)]] PURP [BECOME **have'**(the kids, cake)]

All the information generated in the GDE is now used to yield the parsed tree for our example in the following phase of the parsing process (Table 4):

Table 4: Syntactic parsing phase of the example "Louise baked a cake for the child" (Fumero-Pérez & Díaz-Galán, 2017: 41)

SYNTACTIC PARSING	**Parsed tree: ARTEMIS syntax-semantic linkage** 

In the final phases of the parsing process (Table 5), the CLS is first extracted, and that same information is then presented in a purely semantic conceptual representation using the formal FunGramKB representation language COREL. The operators used in the CLS show that the illocutionary force of our example is "declarative", the tense is "past", the type of level-1 construction is the "FBEN" (which is a "kernel-2" construction that corresponds to monotransitive verbal

predicates) and the *Aktionsart* is causative accomplishment "CACC". At the end of the representation, the FunGramKB ontological concepts to which the different lexical items are linked are given:[5]

Table 5: The "CLS extraction" and "COREL scheme presentations" phases for the example "Louise baked a cake for the child" (Fumero-Pérez & Díaz-Galán, 2017: 41)

CLS Extraction	$<^{DECL}_{IF}<^{PAST}_{TENSE}<^{FBEN}_{CONSTR\_L1}<^{CONSTR\_}_{KER2\_L1}<^{CACC}_{AKT}$<[+BAKE_00(%LOUISE_00-Agent,+CAKE_00-Referent,+CHILD_00-Beneficiary)] >>>>>>[5]
COREL scheme representation	+(e1: +BAKE_00 (x1: %LOUISE_00)$_{THEME}$ (x2: +CAKE_00) $_{REFERENT}$ (f1: (e2: +DO_00 (x1)$_{AGENT}$ (e1)$_{REFERENT}$ (f2: +CHILD_00)Beneficiary)) Purpose)

In the implementation process of ARTEMIS, the RRG descriptive apparatus has had to undergo two important adaptations. On the one hand, in RRG, the representation of grammatical categories like aspect, negation and directionals (at the level of the nucleus), directionals, event quantification, modality and negation (at the level of the core) and status, tense evidentials and illocutionary force (at the level of the clause) is given in the Operator Projection that is represented in a distinct projection from the one representing predicates and their arguments (i.e., the Constituent Projection) (Van Valin, 2005: 12).[6] This Operator Projection has been substituted in ARTEMIS by Attribute-Value Matrixes (AVMs) (Cortés-Rodríguez & Mairal-Usón, 2016: 95) that capture the corresponding values for grammatical categories (see Table 6 for two examples), and that adhere to the principle of linearity of processing established by the computational parsing application "so that a tag or label can be assigned to each of the constituents in the sentence, and the machine can analyse them in a strict sequential order" (Martín-Díaz & González-Orta, 2020: 10). Additionally, the Operator Projection has also been replaced by unification mechanisms (Boas & Sag, 2012; Sag et al., 2003) where "morphosyntactic parsing is carried out jointly by a set of production rules and a number of feature unification operations intended to satisfy the structural and semantic constraints encoded in the AVMs" (Cortés-Rodríguez & Mairal-Usón, 2016: 96).

5 The symbol '%' that can be found in the conceptual logical structures is used to introduce an ontological concept for specific entities in the conceptual model (Onomasticon) of FunGramKB (see Figure 1).

6 The Constituent Projection in RRG defines three syntactic units in the structure of the clause: the nucleus (which includes a verbal, an adjectival or a nominal predicate), the core (which contains the nucleus and its arguments), and the periphery (which includes constituents that are not predicate arguments).

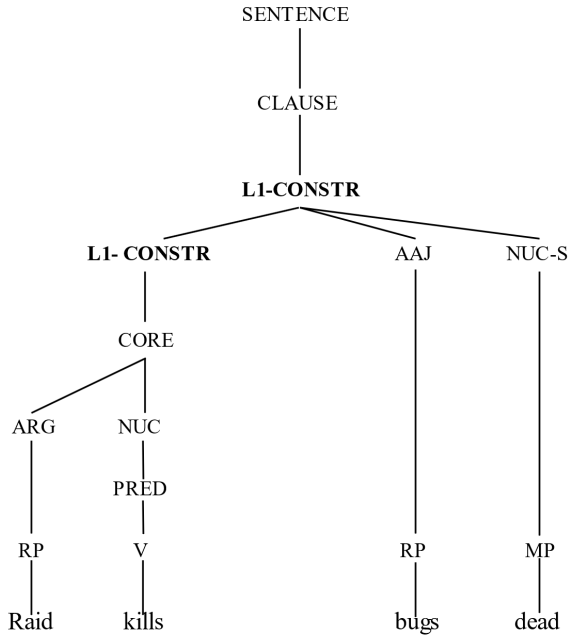
Table 6: Examples of the AVMs for the categories AUX and VERB (Cortés-Rodríguez & Mairal-Usón, 2016: 106-107)[7]

AVM for the category "auxiliary verb"	**AVM for the category "verb"**
<Category Type="AUX"> <Attribute ID="Aspect" /> <Attribute ID="Illoc" /> <Attribute ID="Num" /> <Attribute ID="Per" /> <Attribute ID="Tense"/> </Category>	<Category Type="VERB"> <Attribute ID="Aspect" /> <Attribute ID="Concept" /> <Attribute ID="Illoc" /> <Attribute ID="Num" /> <Attribute ID="Per" /> <Attribute ID="Recip" /> <Attribute ID="Reflex" /> <Attribute ID="Template" /> <Attribute ID="Tense" </Category>

On the other hand, and as put forward by Periñán-Pascual and Arcas-Túnez (2014), the RRG layered structure of the clause has been modified with the insertion of an intermediate constructional node, the level-1 construction node (L1-CONSTR), between the CORE and the CLAUSE nodes (see Figure 2 and the refined tree presented in Table 4), thus reflecting the constructional orientation of the LCM and their four-layered architecture of constructions, all of which are stored in the grammatical module (Grammaticon) of FunGramKB (see Figure 1).

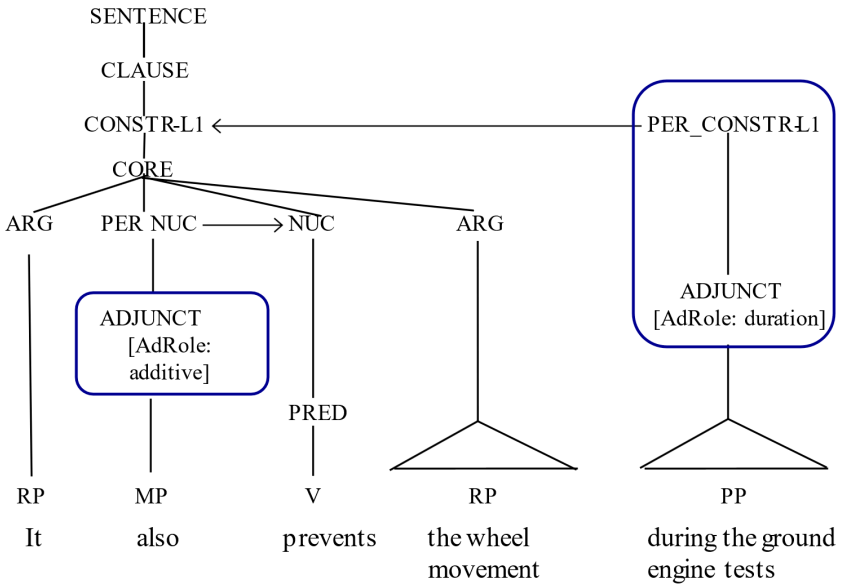
7 AVMs are encoded in XML format, similar to that of other platforms for the analysis of human language data, as is NLTK (Natural Language Toolkit; Bird et al., 2009).

Figure 2: The enhanced LSC of an English resultative level-1 construction



Adverbials in RRG are regarded as peripheral components that modify the three different levels of the Constituent Projection of the clause: in the nuclear periphery, focusing (*only, also, just*), degree (*completely, slightly*) and frequency adverbs (*daily, normally, never*) are located; in the core periphery, the following adverbial modifiers can be found: contingency (*although, in order to*), process (*because, apart from, loudly*), pace (*quickly*), space (*downstairs, above, from*) and temporal modifiers (*during, since, before*); finally, in the clause periphery we can place illocutionary (*briefly, frankly*), evidential (*apparently, presumably*) and epistemic adjuncts (*basically, certainly*) (see Rodríguez-Juárez & Cortés-Rodríguez (forthcoming) for a revised typology of adjuncts). As already mentioned, the addition of an intermediate constructional node (*L1-CONSTR*) in FunGramKB reflects the constructional orientation of the LCM, and since core adverbials share the same semantic typology and positional preferences as the adverbials found in level-1 argumental constructions, core adverbials will be reanalysed as peripheral constituents modifying the L1-CONSTR and, as a result, will be also referred to as CONSTR-L1 adjuncts or L1 adjuncts. Figure 3 shows an example of adverbials modifying the nuclear and L1-CONSTR nodes:

Figure 3: Tree of the enhanced LSC with the insertion of the CONSTR-L1 node of an Airbus sentence



After having summed up the relevant aspects of the grammatical models on which FunGramKB and ARTEMIS are grounded, and briefly outlined the architecture of the knowledge base and the parser, we will now move on to present the AVMs and the parsing rules that have had to be designed to respond to the conditions imposed by the ASD-STE100 controlled natural language that is used in the Airbus corpus.

3. Description of the corpus and methodology

The Airbus corpus[8] is made up of a collection of raw texts that belong to the domain of aircraft maintenance and the subdomain of aeronautical English–aircraft maintenance and that deal with instructions, descriptions and warning notices. It is a closed, synchronic and untagged corpus made up of 2,486 files / 6.697.387 bytes (xml format) that contains 687,345 word tokens (Felices-Lago & Alameda-Hernández, 2017: 109). The corpus has been written, as mentioned in the introduction of this paper, following the lexical and syntactic constraints established by the ASD-STE100 (2017) controlled language. Adverbs in ASD-STE100 are considered as a part of speech together with other seven parts of speech (verb, noun, pronoun, article, adjective, preposition and conjunction) and are briefly defined as "a word that modifies a verb, an adjective,

8 The Airbus corpus is at the disposal of our research group courtesy of Airbus in Seville.

or another adverb. It answers the questions, 'how?, 'where?', 'when', 'how often?' and 'how much?'" (ASD-STE100, 2021, Part 2 Dictionary: 2-0-4). Some orientation is also given as to how to form adverbs from adjectives: "Frequently (but not always), you can make an adverb from an adjective when you attach '-ly' ending to it. The comparative and superlative forms of adverbs are also made with 'more' and 'most'. Thus, they are not given in the dictionary" (ASD-STE100, 2021: 2-0-6). In fact, the words "more" and "most" are listed in the dictionary as independent approved words that can be used to form the comparative and superlative forms of adverbs. Other regular and irregular comparative and superlative adverbs formed by adding "-er" or "-est" are not approved in the ASD-STE100 dictionary. For example, the comparative adverb "later" is listed in the dictionary, but in lower case letters, which indicates that it is not an approved word that should be substituted by the adverb "subsequently", as in "Make sure that the tool is *subsequently* available (and not "available *later*") for the installation procedure" (ASD-STE100, 2021, Part 2 Dictionary: 2-1-L2). ASD-STE100 (2021) includes 182 approved examples of adverbs, prepositions and conjunctions (Appendix 1) that can be codified as adjuncts realised by adverb phrases ("The MLG door uplock closes *mechanically* and opens *hydraulically*"), prepositional phrases ("Make sure that there are no objects or persons *below the aircraft*") and adverbial clauses ("Put the sensor (6) in the correct position *while you keep the spacer (10) and sensor (2) in position*"), respectively.

Prior to the design of the rules for the internal realization of adjuncts and for their placement in the layered structure of the clause (syntactic rules), as well as to the creation of the catalogue of Attribute-Value Matrixes, a preliminary study was carried out devoted to the study of the semantic and syntactic analysis of a representative sample of adjuncts in the Airbus corpus (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming). To be precise, a sample of 99 adjuncts (67 examples of adverbs, 24 prepositions that could be used to introduce prepositional phrases, and 8 conjunctions introducing adverbial clauses) were analysed in 7,603 example sentences according to the position they occupied in the sentence and to their semantic input.[9] Of these 99 adjuncts (see Appendix 2), 84.84% were examples of approved words included in the Dictionary of ASD-AST100, and only 15.15% were examples of adverbs that are not approved in this STE (already, completely, at the same time, even, except, far, in case of, in general, individually, inside, never, normally, partially, remotely, sometimes), which shows that, although the corpus mostly adheres to the rules of the specification document ASD-STE100 (2017), some deviations can be attested.

9 For a detailed account of the methodology followed in the compilation of the corpus of Airbus adjuncts, see Rodríguez-Juárez & Cortés-Rodríguez (forthcoming).

The design of the parsing rules will reflect the positional and peripheral preferences of the two adjunct types (nuclear adjuncts (NUC) and core/L1-CONSTR adjuncts (CORE/L1)) that were registered in the Airbus corpus and that are presented in Table 7. The possible adjunct positions within the sentence are distributed along four main positions: the extra-clausal positions (i.e., the left-detached position (LDP) and the right-detached position (RDP,) as in "*In this condition* (LDP), the EBCU is in standby mode, *until the crew member applies pressure to one of the brake pedals* (RDP)) and the initial ("*Always* use gloves for protection") , medial ("It *also* adjusts the hydraulic flow in the return operation") and final ("Lower the trailing arm (25) positions *carefully*") positions.

Table 7: Positions and peripheral preferences of adverbials in the Airbus corpus (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming)

PERIPHERAL PREFERENCES	**POSITIONS**				
	LDP	**INITIAL**	**MEDIAL**	**FINAL**	**RDP**
+High **(100%-51%)**				CORE/L1 (58.12%)	
± Middle **(50%-20%)**	NUC (24.59%)		NUC (48.23%) CORE/L1 (24.79%)	NUC (25.53%)	
−Low **(19%-0%)**	CORE/L1 (15.26%)	NUC (0.95%) CORE/L1 (0.90%)			NUC (0.71%) CORE/L1 (0.93%)

Thus, the previous research into the semantics and syntax of adverbials in the Airbus corpus served as a basis for the design of the rules for the internal realization of adjuncts in the Airbus corpus and for their placement in the layered structure of the clause (syntactic rules) as well as for the elaboration of the catalogue of Attribute-Value Matrixes that are needed in the implementation of this CNL in the parser ARTEMIS. The integration of these production rules in the GDE of ARTEMIS will ultimately be helpful for an effective and more detailed automatic parsing of sentences in which

optional peripheral constituents appear, thus increasing the explanatory potential of this prototype.

4. Our contribution to the implementation of Airbus adjuncts in ARTEMIS

One of the main aims of this study is to contribute to the development of the NLP prototype ARTEMIS by providing, on the one hand, the catalogue of Attribute-Value Matrixes (AVMs) for peripheral constituents (Section 4.1), and, on the other hand, the syntactic rules for both the internal realization of the different peripheries (Section 4.2) and for the localization of the peripheries at the different layers of the constituent structure that are necessary for the parsing of the syntactic and semantic representation of Airbus adjuncts (Section 4.3). It is important to bear in mind that, differently from other syntactic parsers, the rules that we have designed are based on a solid linguistic model and are aimed at retrieving not only the syntactic but also the semantic structure of a given fragment of language.

4.1. Catalogue of AVMs

As already mentioned in Section 2, for a satisfactory and effective formalization of adverbials in ARTEMIS, a methodological adaptation of the RRG descriptive apparatus had to be embraced that consisted in the replacement of the RRG Operator Projection by AVMs, which are feature-bearing structures that encode the selectional and semantic information of the different types of grammatical constituents in the format of attributes and values. The AVMs for the category Adjunct and the sentence nodes for the nuclear peripheries (PER_NUC) and the core/level-1peripheries (PER_CORE/L1) have been refined with the addition of the attribute "weight" in the case of nuclear and core peripheries:

ADJUNCT

<Category Type="ADJUNCT">
  <Attribute ID="AdjunctRole" />
  <Attribute ID="concept" />
  </Category>

PER_NUC

<Category Type="PER_NUC">
  <Attribute ID="AdjunctRole" />
  <Attribute ID="concept" />
  <Attribute ID="weight" />
  </Category>

PER_CORE/L1

<Category Type="PER_L1">

<Attribute ID="AdjunctRole" />

<Attribute ID="concept" />

<Attribute ID="weight" />

</Category>

A closer look at the attributes included in these AVMs shows that the values under the attribute "Adjunct Role" display all the possible semantic types that have been distinguished for the three types of adverbials in the different layers of the LSC (see Rodríguez-Juárez & Cortés-Rodríguez, forthcoming), as can be seen below, where the type of periphery has been facilitated in brackets ((N) for nuclear, (C) for core and (CL) for clausal peripheries):

<Attribute ID="**AdjunctRole**" obl="+" num="s">

<Value>**Additive**</Value> (N)

<Value>**Beneficiary**</Value> (C)

<Value>**Company**</Value> (C)

<Value>**Concession**</Value> (C)

<Value>**Conditional**</Value> (C)

<Value>**Definite frequency**</Value> (N)

<Value>**Degree/Amplifiers**</Value> (N)

<Value>**Degree/Diminisher**</Value> (N)

<Value>**Direction**</Value> (C)

<Value>**Distance**</Value> (C)

<Value>**Duration**</Value> (C)

<Value>**Epistemic**</Value> (CL)

<Value>**Evidential**</Value> (CL)

<Value>**Exception**</Value> (C)

<Value>**Illocutionary**</Value> (CL)

<Value>**Indefinite freq.**</Value> (N)

<Value>**Instrument**</Value> (C)

<Value>**Limiter**</Value> (N)

<Value>**Location/Position**</Value> (C)

<Value>**Manner**</Value> (C)

<Value>**Means**</Value> (C)

<Value>**Pace**</Value> (C)

<Value>**Path**</Value> (C)

<Value>**Purpose**</Value> (C)

<Value>**Reason**</Value> (C)

<Value>**Result**</Value> (C)

<Value>**Source**</Value> (C)

<Value>**Span**</Value> (C)

<Value>**Time position**</Value> (C)

</Attribute>

The value for the attribute "concept" in the AVMs indicates that the parser has to address the FunGramKB core ontology to find and retrieve the concept associated with the lexical entry of the head word functioning as adjunct; this is encoded as an instruction by means of the string [FIND: core > concept > concept | CHECK: %\w*]. The attribute "concept" is then expressed as follows:

<Attribute ID="**Concept**" obl="*" num="s"

<Value>[FIND: core > concept > concept | CHECK: %\w*]</Value>

</Attribute>

Finally, we have added an attribute, "weight", with 6 values, that is going to determine the likelihood or probability of occurrence of an adverbial in a particular position of the LSC. In the AVM, the weight values are open, but they are codified at the very beginning of the syntactic parsing where each position is given a specific value based on the analysis of the Airbus positional preferences and frequencies that we have established for the different peripheral constituents (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming). The weight that is attached to each position is dependent on the frequency of occurrence of adjuncts in that particular position. Thus, the higher the frequency of occurrence, the less marked the assignment of an adjunct to that

position is, and, as a result, the lower the weight that is assigned to this position in the ARTEMIS parsing rules:

<Attribute ID="**Weight**" obl="*" num="1">
<Value>0</Value>
<Value>1</Value>
<Value>2</Value>
<Value>3</Value>
<Value>4</Value>
<Value>5</Value>
<Value>6</Value>
</Attribute>

We reproduce here (Table 8) the weight-based priority that has been assigned to each position and for each type of adjuncts as a result of the different rates of frequency that have been registered in the previous study of the Airbus corpus (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming):

Table 8: Assignment of weight to the different positions per adjunct type based on frequency rates in the Airbus corpus (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming)

POSITION	Core/L1 adjuncts	WEIGHT	NUC adjuncts	WEIGHT
RDP	0.93%	6	0.71%	6
FIN	58.12%	3	25.53%	5
MED	24.79%	5	48.23%	4
INI	0.90%	6	0.95%	6
LDP	15.26%	6	24.59%	5

WEIGHTS	FREQUENCIES
1	83.34% - 100%
2	66.68% - 83.33%
3	50.02% - 66.67%
4	33.36% - 50.01%
5	16.70% - 33.35%
6	0.00% - 16.69%

The assignment of a weight-based approach to positional preferences is substantiated by psycholinguistic evidence supporting the use of frequency factors by the human sentence parser in order to iron out local ambiguity (cf. Pickering & van Gompel, 2006). Additionally, Periñán-Pascual and Arcas-Túnez (2014: 186) suggest that "it would be more effective to apply the 'weight-based priority' from the beginning of the syntactic parsing with the purpose of minimizing global syntactic ambiguity". The inclusion of the weight attribute for adverbials guarantees that it will be activated at the earlier stages in parsing operations.

4.2. Parsing rules for the internal realization of airbus peripheries in ARTEMIS

The rules for the internal realization of the different peripheries incorporate and distribute the values assigned to the different attributes of the AVMs (Table 9). In the case of adjuncts, the syntactic rule includes the attributes "concept" and "adjunct role" with all the possible semantic types. The rule also codifies its different realizations as modifier phrases (MP) (like *also*), prepositional phrases (PP) (like *behind the handle*), referential phrases (RP) (like *today*) and clauses (CL) (like *until the hinge pin is installed*):

Table 9: Rule for the internal realization of adjunct peripheries in ARTEMIS

ADJUNCT[**adjunctrole**=additive | amplifier | beneficiary | company | concession | condition al | definitefrequency | diminisher | direction | distance | duration | exception | indefinitefreque ncy | instrument | limiter | location | manner | means | pace | path | purpose | reason | result | sou rce | span | timeposition, **concept**=?]
→ **MP** || **PP**[prep=?p, concept=?] || **RP**[cnt=?, concept=?, def=?, dei=?, num=?, per=?, quant=?] || **CL** [akt=?, concept=?, emph= ?, illoc= ?, sta=?, t=?, tpl=?]

The rule for the internal codification of nuclear peripheries shows a restriction in the array of possible semantic types admitted in this periphery and incorporates the attribute "weight" with all the six possible values (Table 10). The rule also indicates that nuclear peripheries are realised by adjuncts:

Table 10: Rule for the internal realization of nuclear peripheries in ARTEMIS

PER_NUC[**adjunctrole**=additive	degreeamplifier	definitefrequency	degreediminisher	indefinitefrequency	limiter, **concept**=?, **weight**=1	2	3	4	5	6]
→ **ADJUNCT** [adjunctrole=additive	definitefrequency	degree	indefinitefrequency	limiter, concept=?, weight=1	2	3	4	5	6]

The rule for core/L1 peripheries also includes a restricted and distinct list of semantic types that can occur at this level and shows that this periphery may be realised by one, two or even by the concatenation of three adjuncts (Table 11). The rule also specifies the attribute "concept" and the potential "weight" that can be assigned to each position:

Table 11: Rule for the internal realization of core/L1 (concatenated) peripheries in ARTEMIS

1. ONE ADJUNCT

PER_L1[**adjunctrole**=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, **concept**= ?, **weight**=1 | 2 | 3 | 4 | 5 | 6] →

ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6] | |

2. CONCATENATION OF TWO ADJUNCTS

PER_L1[**adjunctrole**=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, **concept**= ?, **weight**=1 | 2 | 3 | 4 | 5 | 6] →
ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6]
ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6]

3. CONCATENATION OF THREE ADJUNCTS

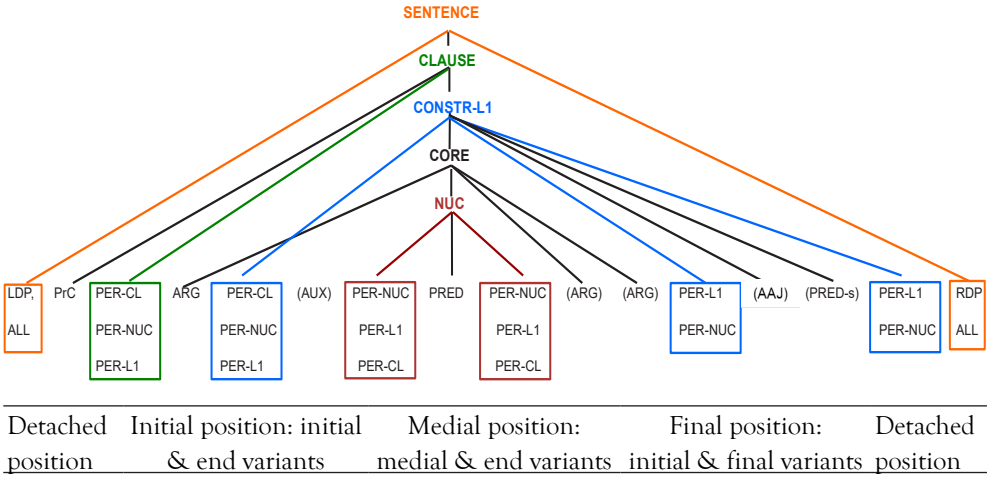
PER_L1[**adjunctrole**=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, **concept**= ?, **weight**=1 | 2 | 3 | 4 | 5 | 6] →

ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6]
ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6]
ADJUNCT[adjunctrole=beneficiary | company | concession | conditonal | direction | distance | duration | exception | instrument | location | manner | means | pace | path | purpose | reason | result | source | span | timeposition, concept: ?, weight=1 | 2 | 3 | 4 | 5 | 6]

4.3. Parsing rules for the localization of Airbus peripheries in the layers of the LSC

This section deals with the process followed in the elaboration of the syntactic rules for the localization of the peripheries in the different layers of the constituent structure of the clause that will be stored in the GDE of the parser ARTEMIS. The peripheries are represented as daughter constituents that belong to the central layers in the LSC. In this way, the computational requirements for the parsing of these constituents in the syntactic rules are met since these rules can only capture linear ordering and immediate dominance (Cortés-Rodríguez & Rodríguez-Juárez, 2019: 74). The scheme provided in Figure 4 for standard English shows that the detached extra-clausal positions (LDP/RDP) are attached to the Sentence, the initial positions are assigned to the Clause and also to the L1-Constructional layer, the medial positions are assigned to the Nucleus, and the final positions are assigned to the L1-Constructional layer. These dominance relations have to be reflected in the parsing rules:

Figure 4: Peripheries and positional preferences in the abstract LSC for standard English (with dominance relations) (Cortés-Rodríguez & Rodríguez-Juárez, 2019: 75)



Before presenting the rules for the localization of adjuncts in the different layers of the LSC, it should be reminded that in the controlled natural language of the Airbus corpus, there is only one instance of a clausal adjunct (*possibly*) in the sentence "The aircraft is possibly not parallel to the ground", as opposed to standard English. This is again one difference in comparison with standard English where clausal subjective adjuncts like *frankly*, *presumably* or *cleverly* are quite frequent. Since ASD-STE100 is a

CNL that aims at providing information as objectively as possible, these types of clausal subjective adjuncts do not occur. Thus, our rules will only display the localization of core/L1 and nuclear peripheral constituents.

At the layer of the nucleus, the parsing rule for Airbus nuclear adjuncts is formulated as follows:

NUC → (PER_L1) $_{\text{Weight:5}}$ (PER_NUC) $_{\text{Weight:4}}$ **PRED** (PER_NUC) $_{\text{Weight:4}}$ (PER_L1) $_{\text{Weight:5}}$

The parentheses indicate optionality of constituents since otherwise the number of possibilities would multiply their linearization possibilities. This rule shows that in the Airbus corpus at the layer of the NUC we can locate nuclear and core/L1 adjuncts in medial positions (as shown in Figure 4), each with a particular weight. As mentioned above, in the AVMs, the weight values are left open, but in the parsing rules each position is given a specific weight. So, in the case of L1 adjuncts in medial position, the weight assigned is 5, which shows that this occurrence is less frequent (24.79%) than the one registered for nuclear adjuncts with a weight of 4 (48.23%). Examples extracted from the Airbus corpus of the possible combinations are given below:

(PER_NUC) PRED:

7. Each system *continuously* monitors the other system.

(PER_L1) PRED:

8. This test sequence *automatically* does a check of the actuators.

PRED (PER_L1):

9. Connect *mechanically* the brake unit to the main wheel.

PRED (PER_NUC):

10. Let the landing gear extend *fully*.

PRED (PER_NUC) (PER_L1):

11. Their routing is *also near to the structure of the leg assembly* ...

At the layer of the core/L1 periphery, the adjuncts can be located both in initial and in final positions, with different arrangements as can be seen in the rule:

PER_L1/CORE → (ARG) (PER_NUC) $_{\text{Weight:6}}$ (PER_L1) $_{\text{Weight:6}}$ (AUX) **PRED** (ARG) (ARG) (PER_L1) $_{\text{Weight:3}}$ (PER_NUC) $_{\text{Weight:5}}$ (AAJ) (PRED-s) (PER_L1) $_{\text{Weight:3}}$ (PER_NUC) $_{\text{Weight:5}}$

Below we show examples of a nuclear periphery in initial position after the argument and before the auxiliary (example 12), of a nuclear periphery in final position (example 13) and of a concatenation of a nuclear and a core/L1 adjunct in final position after the argument (example 14):

(ARG) (PER_NUC) (AUX) PRED:

12. The hitch pin *only* can be installed if the kneeling lock valve is correctly closed.

PRED (ARG) (PER_NUC):

13. The Normal Braking System operates the brake units *usually*.

PRED (ARG) (PER_NUC) (PER_L1):

14. Do not push the pin (11) *completely through the lug*.

At the layer of the clause, the nuclear and core/L1 peripheries have to be located in initial positions:

CL → (PER-NUC) $_{\text{Weight:6}}$ (PER_L1) $_{\text{Weight:6}}$ **CONSTR-L1**

We have registered examples of nuclear and core/L1 adjuncts in this position but, as can be seen from the weight assigned to them, they are not found very often (weight 6). No examples of a concatenation of two distinct peripheries before the L1-construction have been registered:

(PER_NUC) CONSTR-L1:

15. *Always* the top bleed valve (8) of the related brake unit (1) must be used ...

(PER_L1) CONSTR-L1:

16. *When the aircraft is on-ground* the MLG shock absorber is compressed**.**

At the level of the sentence, examples of all the possible arrangements have been registered in the Airbus corpus:

SENTENCE → (LDP) CLAUSE (RDP)

In the left and right detached positions, which are pragmatically motivated positions that are usually separated by commas or pauses from the rest of the clause, we have found examples of both nuclear and L1 adjuncts:

LDP → (PER_L1) $_{\text{Weight:6}}$ (PER_NUC) $_{\text{Weight:5}}$

(PER_L1):

17. *In normal mode*, the system sends all the fault messages to ...

(PER_NUC):

18. *In general,* these materials are flammable, poisonous and ...

RDP → (PER_L1) $_{\text{Weight:6}}$ (PER_NUC) $_{\text{Weight:6}}$

(PER_L1):

19. The ADCN interchanges data of the Normal Braking System (*through the BACS software).*

(PER_NUC):

20. This type of equipment (...) can cause damage to equipment, *specially to: electrical equipment ...*

We have also registered examples of both LDP and RDP together in the same sentence:

(LDP) CLAUSE (RDP):

21. *In this condition,* the EBCU is in standby mode, *until the crew member applies pressure to one of the brake pedal assemblies.*

Despite the introduction of weighted options of realization of peripheral types in a given position, syntactic ambiguity will still persist, which will be translated in more than one final parse tree for a given adverbial. However, each tree will be provided with a different weighted value corresponding to its probability to become the winning option. In those cases of cooccurrence of several adverbials in a same sentence, the sum of the weighted values will determine which is the most probable overall syntactic analysis. This option enables the parser to minimise global syntactic ambiguity by solving all cases of local ambiguity in the structure.

5. Conclusions

Controlled natural languages present differences with respect to natural languages in the sense that they try to be simpler by avoiding ambiguity and implicit meaning and, as a result, aim at the comprehensibility and readability of technical texts. These simplified languages make them suitable candidates to be used in the first phases of the computational implementation of natural language processing. In the present study, we have chosen the ASD-STE100 CNL used in the Airbus corpus and have focused on the behaviour of Airbus adjuncts to test the validity of the parser ARTEMIS. Thus, our main aim to enrich the research conducted so far into the development of the grammatical module (GDE) of ARTEMIS through the design of the production rules for the different constituents of the clause has been fulfilled with our proposal of the set of AVMs and parsing rules for ASD-STE100 adverbial units, which have had to be adapted to respond to the reality imposed by the Airbus controlled natural language. In order to account for the large flexibility and variety of adjuncts in terms of the place they can occupy in the sentence and their semantic variability, we have incorporated a weight factor that will facilitate the automatic processing and computational parsing of Airbus adjuncts, thus achieving our secondary aim.

With this research we have contributed to completing the studies that have already been carried out on the development of the GDE in ARTEMIS and on the design of the AVMs and the syntactic rules for ASD-STE100 phrasal and clausal constituents, in which no analysis of adjuncts had been offered so far. Although our findings are limited in the sense that they are restricted to the adverbs of the ASD-STE100 CNL, we believe that they can serve as a stepping-stone towards the analysis of these and other constituents in other English-based CNLs that are relevant to the field of computational linguistics. On the other hand, English-based CNLs are in fact subsets of their "base language" (English) (Khun, 2014: 123), and, consequently, the work done so far can also be valid to do further research into the eventual generation of the logical conceptual structures of natural language expressions when we move from controlled natural languages into non-controlled natural ones. In this process, the scope of research should be widened by the analysis of data relative to the position and behaviour of clausal subjective adverbials (e.g. *frankly, possibly*, etc.) in standard English, as they are not used in texts produced in ASD-STE100.

6. References

Aerospace and Defence Industries Association of Europe (2017). *ASD-STE100 Simplified Technical English. Specification ASD-STE-100. International specification for the preparation of technical documentation in a controlled language*. Issue 7, January 2017. Brussels.

Aerospace and Defence Industries Association of Europe (2021). *ASD-STE100 Simplified Technical English. Specification ADE-STE-100. International specification for the preparation of technical documentation in a controlled language*. Issue 8, April 30, 2021. Brussels. https://asd-ste100.org/

Bird, S., Klein, E & Loper, E. (2009). *Natural Language Processing with Python*. Sebastopol: O'Reilly Media Inc.

Boas, H.C., & Sag, I. (eds.) (2012). *Sign-Based Construction Grammar*. Stanford: CSLI.

Clark, P., Harrison, P., Murray, W.R. & Thompson, J. (2010). Naturalness vs. predictability: a key debate in controlled languages. *Proceedings 2009 Workshop on Controlled Natural Languages (CNL'09)*.

Cortés-Rodríguez, F. J. & Mairal-Usón, R. (2016). Building an RRG computational grammar. *Onomázein*, *34*, 86-117.

Cortés-Rodríguez, F.J. & Rodríguez-Juárez, C. (2018). Parsing phrasal constituents in ASD-STE100 with ARTEMIS. *Voprosy Kognitivnoy Lingvistiki*, *3*, 97-109.

Cortés Rodríguez, F.J. & Rodríguez-Juárez, C. (2019). The syntactic parsing of ASD-STE100 adverbials in ARTEMIS. *Revista de Lingüística y Lenguas Aplicadas*, *14*, 59-79.

Díaz-Galán, A. (2018). Deep parsing for the aviation industry: Adjusting ARTEMIS for parsing simple clauses in ASD-STE100. *Voprosy Kognitivnoy Lingvistiki*, *3*, 83-96.

Díaz-Galán, A. & Fumero-Pérez, M.C. (2016). Developing parsing rules within ARTEMIS: The case of *do* auxiliary insertion. In C. Periñán-Pascual & E.M. Mestre-Mestre (eds.), *Understanding Meaning and Knowledge Representation: From Theoretical and Cognitive Linguistics to Natural Language Processing* (pp. 283-302). Cambridge: Cambridge Scholars Publishing.

Díaz-Galán, A. & Fumero-Pérez, M.C. (2020). An account of constructions in ASD-STE100: Formalizing non-propositional meaning in aviation instructional texts. *Revista Electrónica de Lingüística Aplicada, 19*, 24-41.

Felices-Lago, A. & Alameda-Hernández, A. (2017). The process of building the upper-level hierarchy for the aircraft structure ontology to be integrated in FunGramKB. *Revista de Lenguas para Fines Específicos, 23*(2), 86-110.

Fuchs, N.E., Kaljurand, K. & Kuhn, T. (2008). Attemp to Controlled English for knowledge representation. *Reasoning Web, LNCS*, 5224, 104-124.

Fumero-Pérez, M.C. & Díaz-Galán, A. (2017). The Interaction of parsing rules and argument- predicate constructions: Implications for the structure of the Grammaticon in FunGramKB. *Revista de Lingüística y Lenguas Aplicadas*, *12*, 33-44.

Fumero-Pérez, M.C. & Díaz-Galán, A. (2019). Designing the lexical rules for the parsing of ASD-STE100 function words in ARTEMIS from a Role and Reference Grammar perspective. *Journal of English Studies*, *17*, 149-174.

González-Orta, M. & Martín-Díaz, A. (2022). ARTEMIS: Parsing non-peripheral complex sentences in ASD-STE100. *Onomázein*, *56*, 80-99.

Kailuweit, R., Künkel, L. & Staudinger, E. (eds) (2018). *Applying and Expanding Role and Reference Grammar.* Freiburg: Freiburg University Library [NIHIN Studies].

Kuhn, T. (2014). A survey and classification of controlled natural languages. *Computational Linguistics*, *40*(1), 121-170.

Luzondo-Oyón, A. & Ruiz de Mendoza Ibáñez, F.J. (2015). Argument structure constructions in a natural language processing environment. *Language Sciences*, *48*, 70-89.

Mairal-Usón, R. & Ruiz de Mendoza Ibáñez, F.J. (2008). New challenges for lexical representation within the lexical-constructional model (LCM). *Revista Canaria de Estudios Ingleses*, *57*, 137-155.

Martín-Díaz, M.A. (2017). An account of English *yes/no* interrogative sentences within ARTEMIS. *Revista de Lenguas para Fines Específicos*, *23*(2), 41-62.

Martín-Díaz, M.A. (2018). Parsing *wh*-interrogative sentences within ARTEMIS. *Revista Electrónica de Lingüística Aplicada*, *17*(1), 78-98.

Martín-Díaz, M.A. (2019). Parsing adverbial complex sentences in ASD-STE100 within ARTEMIS. *Complutense Journal of English Studies*, *27*, 159-181.

Martín-Díaz, M.A. & González-Orta, M. (2020). Clausal arguments and peripheries in ASD-STE100: The parsing of subordination in ARTEMIS. *Revista Electrónica de Lingüística Aplicada*, *19*(1), 1-23.

Periñán-Pascual, C. (2012). En defensa del procesamiento del lenguaje natural fundamentado en la lingüística teórica. *Onomázein, 26*, 13-48.

Periñán- Pascual, C. (2013a). Towards a model of constructional meaning for natural language understanding. In B. Nolan & E. Diedrichsen (eds.), *Linking Constructions into Functional Linguistics: The Role of Constructions in Grammar [Studies in Language Companion Series 145]* (pp. 205-230). Amsterdam: John Benjamins.

Periñán-Pascual, C. (2013b). A knowledge-engineering approach to the cognitive categorization of lexical meaning. *Vigo International Journal of Applied Linguistics, 10*, 85-104.

Periñán-Pascual, C. & Arcas-Túnez, F. (2014). The implementation of the FunGramKB CLS Constructor. In B. Nolan & C. Periñán-Pascual (eds.), *Language*

Processing and Grammars: The Role of Functionally Oriented Computational Models (pp. 165-196). Amsterdam: John Benjamins,

Pickering, M.J. & van Gompel, R.P.G. (2006). Syntactic parsing. In M.J. Traxler & M.A. Gernsbacher (eds.), *Handbook of Psycholinguistics* (pp. 455-503). Amsterdam: Elsevier.

Rodríguez-Juárez, C.& Cortés-Rodríguez, F.J. (forthcoming, 2022). A Role and Reference Grammar account of adjuncts in the Airbus Corpus: A quantitative-based study. *Atlantis*, *44* (2).

Ruiz de Mendoza Ibáñez, F.J. (2013). Meaning construction, meaning interpretation and formal expression in the Lexical Constructional Model. In B. Nolan & E. Diedrichsen (eds.), *Linking constructions into functional linguistics: The role of constructions in grammar* (pp. 231-70). Amsterdam: John Benjamins.

Ruiz de Mendoza Ibáñez, F.J. & Mairal-Usón, R. (2008). Levels of description and constraining of factors in meaning construction: an introduction to the Lexical Constructional Model. *Folia lingüística: Acta Societatis Linguisticae Europaeae*, *42*(2), 355-400.

Ruiz de Mendoza Ibáñez, F.J. & Galera Masegosa, A. (2014). *Cognitive Modeling. A Linguistic Perspective*. Amsterdam: John Benjamins.

Sag, I.A., Wasow, T. & Bender, E.M. (2003). *Syntactic Theory: Formal Introduction*. Stanford: CSLI Publications.

Schwitter, R. (2010). Controlled natural languages for knowledge representation. *Colin 2010: Poster Volume*, 1113-1121.

Van Valin, R.D. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.

Van Valin, R.D. (ed.) (2008). *Investigations of the Syntax-Semantics-Pragmatics Interface*. Amsterdam: John Benjamins.

Van Valin, R.D. & LaPolla, R.J. (1997). *Syntax. Structure, Meaning and Function*. Cambridge: Cambridge University Press.

White, C. & Schwitter, R. (2009). An update on PENG Light. *Proceedings of ALTA*, 80-88.

APPENDICES

Appendix 1

List of approved adverbs, prepositions, and conjunctions in ASD-STE100 (2017):

ADVERBS

1. accidentally
2. accurately
3. across
4. aft
5. again
6. almost
7. also
8. always
9. apart
10. approximately
11. automatically
12. axially
13. back
14. brightly
15. carefully
16. chemically
17. clearly
18. clockwise
19. constantly
20. continuously
21. correctly
22. counterclockwise
23. diagonally
24. differently
25. digitally
26. dimly
27. directly
28. down
29. downstream
30. easily
31. electrically
32. electromagnetically
33. electronically
34. equally
35. externally
36. first
37. forward
38. freely
39. frequently
40. fully
41. gradually
42. here
43. horizontally
44. how
45. hydraulically
46. immediately
47. in
48. inboard
49. incorrectly
50. independently
51. initially
52. internally
53. intermittently
54. irregularly
55. in one (TN = technical noun) ... and then the other
56. in progress
57. last
58. laterally
59. lightly
60. linearly
61. locally
62. longitudinally
63. loosely
64. magnetically
65. manually
66. mechanically
67. moderately
68. momentarily
69. more
70. much
71. no
72. not
73. off
74. on
75. only
76. out
77. outboard

78. outdoors
79. overboard
80. permanently
81. pneumatically
82. possibly
83. quickly
84. radially
85. randomly
86. rearward
87. regularly
88. safely
89. same
90. satisfactorily
91. slowly
92. smoothly
93. specially
94. structurally
95. subsequently
96. suddenly
97. sufficiently
98. symmetrically
99. temporarily
100. then
101. there
102. thus
103. through
104. tightly
105. together
106. too
107. unsatisfactorily
108. unusually
109. up
110. upstream
111. vertically
112. very
113. visually
114. yes

PREPOSITIONS

1. abaft
2. about
3. above
4. across
5. adjacent to
6. after
7. aft of
8. against
9. along
10. around
11. as
12. at
13. away from
14. because of
15. before
16. behind
17. below
18. between
19. by
20. down
21. downstream of
22. during
23. for
24. forward to
25. from
26. in
27. inboard of
28. in front of
29. into
30. minus
31. near
32. of
33. off
34. on
35. onto
36. opposite
37. outboard of
38. out of
39. over
40. plus
41. through
42. thru
43. to
44. until
45. up
46. upstream of
47. with
48. without

CONJUNCTIONS

1. after
2. although
3. and
4. as
5. as ... as
6. because
7. before
8. but
9. if
10. or
11. since
12. than
13. that
14. until
15. unless
16. when
17. where
18. while

APPENDIX 2

List of word types (99) that can be realised as adverbial expressions (codified as adverbs, prepositional phrases or clauses); raw representative sample (7,603); final sample (5,180) after not valid instances were eliminated (manual filtering process) (Rodríguez-Juárez & Cortés-Rodríguez, forthcoming).

AIRBUS CORPUS		REPRES. SAMPLE	FINAL SAMPLE
WORD TYPES	WORD TOKENS	WORD TOKENS	WORD TOKENS
99	**67,556**	**7,603**	**5,180**
above	17	16	6
accurately	5	5	5
across	3	3	3
after	141	104	82
again	92	75	75
against	16	15	13
along	25	24	21
already	2	2	1
also	172	120	120
although	1	1	1
always	45	40	40
away	96	77	11
apart (from)	1	1	1
approximately	130	98	2
around	39	36	11
as	105	83	61

AIRBUS CORPUS		REPRES. SAMPLE	FINAL SAMPLE
WORD TYPES	WORD TOKENS	WORD TOKENS	WORD TOKENS
99	67,556	7,603	5,180
at the same time	134	100	100
automatically	45	40	40
because	36	33	33
before	830	270	258
behind	3	3	1
below	144	106	46
between	420	205	92
by	97	78	19
carefully	70	60	60
clearly	42	38	38
clockwise	25	24	23
completely	2	2	2
continuously	6	6	6
correctly	483	219	219
counterclockwise	20	19	19
directly	152	110	110
down	520	226	12
downstream	10	10	4
during	614	242	242
easily	16	15	15
electrically	24	23	22
equally	1	1	1
even	3	3	1
except	5	5	5
externally	1	1	1
forward	268	160	1
freely	6	6	6
far	8	8	2
for	5908	375	375

AIRBUS CORPUS		REPRES. SAMPLE	FINAL SAMPLE
WORD TYPES	WORD TOKENS	WORD TOKENS	WORD TOKENS
99	**67,556**	**7,603**	**5,180**
from	3438	358	90
fully	197	132	37
hydraulically	27	25	20
if	6945	378	335
immediately	102	81	75
in	9880	384	384
in case	2	2	2
in general	83	69	83
incorrectly	3	3	3
independently	8	8	7
individually	2	2	2
initially	1	1	1
inside	26	24	6
internally	2	2	2
lightly	5	5	1
linearly	1	1	1
manually	30	28	28
mechanically	30	28	28
momentarily	4	4	4
much	87	71	71
near	700	255	246
never	1	1	1
normally	1	1	1
off	621	243	29
on	15780	390	156
only	230	146	7
onto	3	3	3
out	76	64	37
over	18	17	3

AIRBUS CORPUS		REPRES. SAMPLE	FINAL SAMPLE
WORD TYPES	WORD TOKENS	WORD TOKENS	WORD TOKENS
99	67,556	7,603	5,180
partially	7	7	7
permanently	4	4	4
possibly	1	1	1
rearward	2	2	1
remotely	1	1	1
safely	5	5	5
since	1	1	1
slowly	64	55	55
smoothly	2	2	2
some times	5	5	5
specially	9	9	1
structurally	3	3	2
suddenly	9	9	9
sufficiently	2	2	2
temporarily	8	8	8
then	619	243	2
through	485	219	207
x times	34	31	26
to	13929	389	389
until	135	101	101
usually	5	5	5
when	1849	329	299
while	48	43	43
with	1231	302	131
without	12	12	4

APPENDIX 3

List of abbreviations:

AAJ	Argument Adjunct
ACA	Active Accomplishment
ACE	Attempto Controlled English
AdRole	Adjunct Role
ARG	Argument
ASD-STE100	AeroSpace and Defence-Simplified Technical English Specification
AUX	Auxiliary
AVM	Attribute-Value Matrix
CACC	Causative Accomplishment
CL	Clause
CLN	Controlled Natural Language
CLS	Conceptual Logical Structure
CORE/L1	Core/Level-1 CONSTRUCTION Adjuncts
COREL	Conceptual Representation Language
CPL	Computer Processable Language
FBEN	For-Benefactive construction
FIN	Final position
GDE	Grammar Development Environment
Iloc	Illocutionary Force
INI	Initial position
L1	Level 1
L1-CONSTR	Level-1 Construction
LCM	Lexical Constructional Model
LDP	Left-Detached Position
LSC	Layered Structure of the Clause
MED	Medial position
MP	Modifier Phrase

NLP	Natural Language Processing
NUC	Nucleus / Nuclear Adjunct
NUC-S	Secondary Nucleus
Num	Number
obl	Obligatory
PENG	Processable English
Per	Person
PER-CL	Clausal Peripheries
PER_L1/ CORE	Level-1/Core Peripheries
PER_NUC	Nuclear Peripheries
PP	Prepositional Phrase
PRED	Predicate
RDP	Right-Detached Position
Recip	Recipient
Reflex	Reflexive
RP	Referential Phrase
RRG	Role and reference Grammar
STE	Simplified Technical English
V	Verb

How complex is professional academic writing? A corpus-based analysis of research articles in 'hard' and 'soft' disciplines

Javier Pérez-Guerra[1]
Department of English
Universidade de Vigo
jperez@uvigo.es

Elizaveta A. Smirnova
Foreign Languages Department
HSE University / Universidade de Vigo
easmirnova@hse.ru

Abstract

This study focuses on the analysis of linguistic complexity in professional academic writing in light of the empirical evidence provided by a 1,597,000-word corpus of 'hard' (life and physical sciences) and 'soft' (arts and social) scientific research articles published in leading peer-review journals. Specifically, this investigation aims both to describe the complexity features of texts written by professional authors and to test the hypothesis that linguistic complexity varies across disciplines. Since previous studies have revealed that automatic complexity indices do not sufficiently succeed in providing a comprehensive description of complexity of texts, in this paper complexity has been measured in two ways: quantitatively through the indexes provided by Lu's (2010) L2 Syntactic Complexity Analyser, and through the more qualitative analysis of a selection of metrics associated with clausal and phrasal complexity in seminal studies. The data show, first, that syntactic complexity indices (basically, strategies of coordination and subordination) are statistically relevant to the characterisation of specifically the soft-science disciplines; second, that there is a continuum across subdisciplines within the broad distinction of soft versus hard genres; and, third, that the soft genre demonstrates a more stable productivity of clausal-complexity strategies, while phrasal-complexity features are more pervasive in the hard-science subcorpus.

Keywords: academic writing, complexity, corpus, hard sciences, soft sciences.

1 I am grateful to the Spanish State Research Agency and the European Regional Development Fund (MCIN/AEI/10.13039/501100011033 – PID2020-117541GB-I00), and Xunta de Galicia (grant no. ED431C 2021/52) for generous financial support.

Resumen

Este estudio se centra en el análisis de la complejidad lingüística en la escritura académica profesional a la luz de las pruebas empíricas aportadas por un corpus de 1.597.000 palabras de artículos de investigación científica '*hard*' (ciencias físicas y de la vida) y '*soft*' (humanidades y ciencias sociales) publicados en las principales revistas sometidas a revisión por pares. En concreto, esta investigación pretende describir las características de la complejidad de los textos escritos por autores/as profesionales y poner a prueba la hipótesis de que la complejidad lingüística varía según las disciplinas. Dado que estudios anteriores han revelado que los índices automáticos de complejidad no consiguen proporcionar una descripción exhaustiva de la complejidad de los textos, en este trabajo la complejidad se mide de dos maneras: cuantitativamente a través de los índices proporcionados por el L2 Syntactic Complexity Analyser de Lu (2010), así como a través de un análisis más cualitativo de una selección de métricas asociadas en estudios seminales a la complejidad clausal y sintagmática. Los datos muestran, en primer lugar, que los índices de complejidad sintáctica (básicamente, la coordinación y la subordinación) son estadísticamente relevantes para la caracterización de las disciplinas específicas de las ciencias '*soft*'. En segundo lugar, se demuestra que existe un continuo entre las subdisciplinas dentro de los géneros '*soft/hard*'. En tercer lugar, este estudio concluye que el discurso académico '*soft*' demuestra una productividad más estable de las estrategias de complejidad clausal, mientras que los rasgos de complejidad sintagmática son más dominantes en el subcorpus de las ciencias '*hard*'.

Palabras clave: escritura académica, complejidad, corpus, ciencias exactas, ciencias sociales.

1. Introduction

Different aspects of linguistic complexity have been explored from the perspective of corpus linguistics. As regards the complexity of writing, the literature has focused on the degree of complexity of L2 writing with respect to L1 writing (among others, Hinkel, 2003; Ai & Lu, 2013; Lambert & Nakamura, 2019), on correlations between text complexity, language proficiency and task types (Crossley & McNamara, 2012; Biber et al., 2016; Casal & Lee, 2019), or on the development of complexity in writing after intensive instruction (Crossley & McNamara, 2014; Mazgutova & Kormos, 2015). Despite the potential pedagogical implications of investigating the realisation of complexity features in expert writing, complexity in professional academic writing has been comparatively understudied to date. Research articles can be taken as a benchmark for optimal academic writing and can provide learners with "a rich and authentic introduction to the complexities and nuances of the genre" (Kelly-Laubscher

et al., 2017: 3). As claimed by Swales (1990: 177), research-oriented writing constitutes a core genre in academic discourse, comprising various subgenres such as dissertations, monographs and presentations. According to Casal et al. (2021: 2), the investigation of the degree of syntactic complexity in research articles can provide "important insights into the role that syntactically complex structures play in disciplinary RA [Research Article] writing practices". This is because it evinces the variation demonstrated by complex structures as demanded by their specific functional goals. It is from this angle that the study of complexity traits in research articles published in specialised peer-reviewed journals may help learners master the conventions of the genre.

This study undertakes a quantitative analysis of features associated with linguistic complexity, conducted on a 1,597,000-word corpus of research articles in four ('soft') arts and social sciences (business studies, linguistics, history, and political science), and four ('hard') life and physical sciences (mathematics, engineering, chemistry, and physics), published in leading journals. The goal is twofold: first, to describe the complexity features of research articles written by professional authors and, second, to test the hypothesis that linguistic complexity varies across disciplines. Specifically, two research questions (RQ) are addressed here:

- RQ1. Which indices of linguistic complexity can serve as proxies for the characterisation of 'hard' and 'soft' sciences?

- RQ2. Do 'hard' and 'soft' scientific writings differ as regards the realisation of linguistic complexity features? Are these differences observable across specific disciplines?

This paper is organised as follows. Section 2 outlines the previous works on complexity in academic discourse and disciplinary variation. Section 3 describes the data and method of analysis. Section 4 presents the analysis and results of the study, which are summarised and discussed in Section 5.

2. Literature review

Linguistic complexity is a diverse notion whose scientific analysis has been carried out by exploring a number of lexical, structural and syntactic features (for example, Bulté & Housen, 2012) traditionally associated with production difficulty, proficiency and/or sophistication (Ortega, 2003: 492). This section summarises previous studies on linguistic complexity in academic writing specifically (Section 2.1) and informs of computational tools providing indices of complexity relying on taxonomies of linguistic features (Section 2.2).

2.1. Linguistic complexity in academic writing

Two dimensions can be identified in the study of linguistic complexity: phrasal and clausal complexity. Phrasal complexity is realised, for example, by the "dense use of embedded phrases functioning mostly as modifiers of a head noun", while clausal complexity is evinced by, for instance, the "dense use of dependent clauses functioning as clause constituents (complement clauses) or clausal modifiers (adverbial clauses)" (Biber & Gray, 2016: 141). In their seminal paper, Biber et al. (2011) measured complexity features in two corpora: a written corpus of research articles and a spoken corpus of face-to-face conversations. They concluded, on the one hand, that most of the measures evincing clausal subordination were more common in conversation than in academic writing. This conclusion is also in line with the findings presented in Biber et al. (1999). On the other hand, they demonstrated that academic writing featured more complex (specifically, noun) phrases. Therefore, it is suggested that grammatical complexity should be associated not with an extensive use of dependent clauses, typical of conversation, but with "linguistic units with phrases embedded in phrases" (Biber et al., 2021).

Gray (2015) explored linguistic complexity by paying attention to disciplinary variation. Her study focused on phrasal and clausal complexity in research articles in six disciplines, grouped into 'hard' sciences (physics and biology), social sciences (applied linguistics and political science) and humanities (history and philosophy). Gray concluded that clausal complexity is more prominent in the humanities and less salient in the hard sciences. This finding is in line with Staples et al.'s conclusion (2016), who studied (T-unit/clause-based) clausal and phrasal complexity in university students' writing in a number of disciplines. These authors found that "writers in arts and humanities disciplines and, to a lesser extent, the social sciences use more clausal features than writers in the life and physical sciences" (2016: 31). They also showed that, as university student writing develops, phrasal complexity increases and clausal complexity decreases. Gardner et al. (2019) clustered a number of linguistic complexity features in university student writing across four dimensions, and each dimension or cluster was explored according to the discipline, genre and level. Nesi & Gardner (2019) concluded that clausal complexity is more prominent in the so-called 'soft' disciplines and in the more conversational genres. In detail, in their study, complexity was shown to be achieved through the use of epistemic adverbials and stance nouns complemented by *that*-clauses in the texts in the soft sciences and by stance verbs in the conversational texts, such as narrative recounts. The authors contend that the findings might have important implications for the way in which academic writing is taught at universities.

Linguistic complexity in native and non-native academic writing has been the focus of a number of studies that justify this case study. To give a few examples, Gray

(2013) studied complexity variation in three types of research articles (theoretical, quantitative and qualitative) by investigating linguistic features related to elaboration/involvement *versus* informational density, one of the dimensions suggested by Biber already in his (1988) seminal study, and showed that the scope of complexity is broader than mere discipline variation since it also correlates with differences in the purpose and the type of evidence used in a particular study. Wu et al. (2020) explored syntactic complexity in research papers authored by ELF (English as a Lingua Franca) writers as opposed to those written by native speakers. By adopting Lu's (2010) indices, Wu et al. showed that ELF authors tend to use longer sentences, more coordinate phrases and complex nominals, and rely on nominal phrases to a greater extent than native writers. In a similar vein, Ruan (2018) explored phrasal complexity in journal abstracts by native English and non-native Chinese writers through the use of elaborated noun phrases (e.g. noun phrases with at least one pre-modifying element or a post-modifying prepositional phrase), and found that the non-native writers used more complex and elaborated noun phrases, particularly, employing significantly more noun premodifiers and multiple noun sequences, whereas the native writers opted for a more frequent use of the post-modifying *of* phrases. In their recent study, Yin et al. (2021) compared syntactic complexity indices provided by Lu's L2 Syntactic Complexity Analyser in emerging international research articles authored by L2 novice writers and those corresponding to publications by expert researchers. Their analysis revealed significant differences between the two types of texts; for instance, it was determined that novice writers use fewer verb phrases per T-unit and fewer instances of subordination, which might be explained by L1 transfer and by a lack of expertise in the use of such structures by the emerging writers.

2.2. Tools for measuring linguistic complexity

Over recent decades, several web-based tools have been developed for the automated analysis of the degree of complexity evinced by texts. To give a few examples, SyB http://sifnos.sfs.uni-tuebingen.de/SyB-0.1/#analyzer, developed at the university of Tübingen on the basis of the Common Text Analysis Platform (Chen & Meurers, 2016), provides 13 complexity indices of lexical, syntactic and discoursal phenomena. Coh-Metrix (http://tool.cohmetrix.com) detects basic cohesion, lexical, syntactic and semantic-discursive features, along with other metrics reporting textual lexical diversity and readability (approximately 200 metrics overall), which makes this tool especially useful for the study of text cohesion (as in, for example, Graesser et al., 2004; McNamara et al., 2010). The Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) (Kyle, 2016) measures a number of syntactic sophistication and complexity indices in learner writing. The indices include features already provided by Lu's (2010) Syntactic Complexity Analyser, discussed below, as well

as indices related to fine-grained clausal complexity, fine grained phrasal complexity and syntactic sophistication, employing the Stanford Neural Network Dependency Parser (version 3.5.1; Chen & Manning, 2014) and a Python XML parser that counts the relevant structures.

The tool used in this study is the L2 Syntactic Complexity Analyser (L2SCA henceforth; http://aihaiyang.com/software), developed by Lu (2010). It provides the frequencies and the ratios of complexity indices in Wolfe-Quintero et al. (1998) and Ortega (2003). L2SCA employs the Stanford parser (Klein & Manning, 2003) and is able to identify sentences, clauses, T-units[2], phrases, etc., using Tregex (Levy & Andrew, 2006), a utility that matches patterns in trees (https://nlp.stanford.edu/software/tregex.shtml). Lu tested the measures generated by the analyser in his (2017) study on syntactic complexity in L2 writing and found that a large number of the L2SCA metrics were "predictive of holistic measures of writing quality" (2017: 505). The indices of phrasal and clausal complexity that correlated with writing scores to a greater extent were: mean length of sentence, mean length of T-unit, mean length of clause, ratio of dependent clauses per clause and ratio of complex nominals per clause.

The measures provided by automated complexity analysers have been called into question in the literature. For example, Lambert & Kormos (2014: 2) contend that different types of subordinate constructions (e.g. adverbial clauses, complement

2 T-units, very similar to AS-units, are defined as the "shortest grammatically terminable units into which a connected discourse can be segmented without leaving any residue" (Hunt, 1964: 34), more specifically as "one main clause plus any subordinate clause or nonclausal structure that is attached to or embedded in it" (Hunt, 1970: 184; similarly in 1965: 36). 'Clause' is defined as "a structure with a subject and a finite verb (...), and includes independent clauses, adjective clauses, adverbial clauses, and nominal clauses" in Lu (2010: 481). To give an example, (i) illustrates a single T-unit (the whole utterance) consisting of one main clause (*I don't know ... to see something else*), which contains one embedded dependent clause (*why I was expecting...*), and the latter, in turn, includes the embedded dependent clause *to see something else*.

(i) I don't know [[why I was expecting [to see something else]].

The concept of c-unit, close to that of T-unit, also includes non-clausal structures which have communicative value, like *Coffee, please*. Lintunen & Mäkilä (2014: 394) develop their concept of U-unit, which is valid for the segmentation of spoken productions. In their words, a U-unit is "one independent clause or several coordinated independent clauses, with all dependent clauses or fragmental structures attached to it, separated from the surrounding speech by a pause of 1.5 seconds or more, or, especially in occurrences of coordination, a clear change in intonation and a pause of 0.5 seconds or more (depending on the average length of boundary pauses in the sample), containing one semantic unity" (2014: 385). Their analysis of the indexes of written and spoken L2 productions leads to the conclusion that "the choice of segmentation unit strongly affected the results, and that spoken language complexity may not be as different from written language complexity as it had been claimed in several earlier studies".

clauses controlled by verbs, complement clauses controlled by nouns) "emerge at different points in the developmental process [so,] the use of measures that conceptualise subordination as a unitary process can mask, rather than illuminate, developmental variation during task performance". They also point out that estimates of subordination can be inaccurate since, for example, clauses introduced by disjunctive markers such as *I see* or *I think* are parsed as subordinate by the analysers, whereas they do not necessarily illustrate syntactic subordination. In this vein, Wijers (2018) explored subordinate clauses in the writing of Dutch-speaking learners of Swedish as a foreign language and of native speakers, and concluded that the subordination ratios (the number of subordinate clauses divided by the total number of clauses, and the number of subordinate clauses divided by the number of T-units) are not efficient predictors of syntactic complexity since the differences in such ratios in learner and native writings were not significant, whilst the texts in which these subordination strategies were employed were clearly dissimilar. Similarly, Kyle & Crossley's (2018) in their study of TOEFL exams come to the conclusion that although the measures of phrasal complexity are better predictors of writing quality than clausal indexes, more fine-grained taxonomies of phrasal and clausal complexity per dependency/subordination type contribute to the explanatory power of the statistical model.

Some studies suggest that mode and task rank higher than proficiency as predictors of complexity. For instance, Biber et al.'s (2016: 662) study on complexity in spoken and written TOEFL exams demonstrated that "task-type differences on standardised language exams – associated with both speech versus writing and with different communicative purposes – are systematically associated with linguistic differences, especially with the use of grammatical complexity features". This finding is in line with previous research by the same team. In particular, Biber et al. (2011: 29; similarly, Biber & Gray, 2011) claimed that complexity cannot be seen as "a single unified construct", therefore, it is not reasonable to believe that any single measure will be able to "adequately represent this construct". They found that T-unit- and subordination-based (i.e. clausal) measures are not typical of academic writing but of conversational discourse, whereas nominal/prepositional (i.e. phrasal) measures are good indicators of academic writing.

As regards the influence of mode on the preference for the complexification strategies computed by the analysers, proficiency has been claimed not to run necessarily parallel to complexity, at least in the way the latter is measured by the analysers. In an influential study, Crossley et al. (2014) carried out a multifactorial analysis of L1 academic writing (argumentative essays), with such variables as score/grade, topic, writer's geographical area, timing and (handwritten or typed) production, as well as with the indexes provided by Coh-Metrix. It was found that the variable reflecting

the grade assigned to the essays explained only 5 percent of the total variance. Other studies give support to a certain parallelism between phrasal and clausal complexity. To give an example, with the objective of testing the measures provided by his analyser, Lu (2017: 505) operationalised syntactic complexity in L2 writing and concluded that "[a] large subset of the measures incorporated in L2SCA has been found to be predictive of holistic measures of writing quality as well. (...) [T]he following measures significantly correlated with writing scores: mean length of sentence, mean length of T-unit, mean length of clause, dependent clauses per clause, and complex nominals per clause", that is, indexes of phrasal and clausal complexity. By contrast, Lambert & Nakamura's (2019) study compared productions by L2 English produced by Japanese learners and native speakers through variables such as the proportion of simple, compound (coordination) and complex (nominal, adverbial and relative subordination) utterances, and phrasal-complexity measures like the ratio of words and modifiers per noun phrase. They prove the negative relationship between phrasal and clausal syntax: "[a]s the (...) measures of clausal complexity increase, the (...) measures of phrasal complexity tended to decrease and vice versa" (2019: 10), which is in keeping with the claims of Biber and colleagues. However, and this actually contradicts Biber's previous findings, they also observe a "negative relationship between proficiency and all [the] measures of phrasal syntax (words, modifier types, modifier tokens, and subordinated nouns per NP)": as proficiency increased, NPs became syntactically simpler, at least in the intermediate level of students' proficiency.

Substantial differences among disciplines and sub-registers or sub-genres have been reported in the literature as regards the use of complexity measures. Thus, Gardner et al. (2019: 670) analyse L1 university assignments and find that "the writing situation – disciplinary group [Arts and Humanities, Life Sciences, Physical Sciences and Social Sciences], genre family, discipline, and level of study – is key to interpreting each dimension" resulting from the application of the multidimensional analysis of 39 lexico-grammatical features associated with different aspects of linguistic complexity – see also Staples et al. (2016) in this respect.

Finally, in a recent paper on specifically developmental stages in L1 and L2 writing, Biber et al. (2020) argue that the omnibus complexity measures employed by automated complexity analysers fail to provide a comprehensive description of complexity, which must be compensated for by also describing the "multiple structural types, syntactic functions, and systematic patterns of variation across spoken and written registers" (2020: 13). In this vein, even though the current study does not investigate developmental stages associated with native academic writing, the analysis of the automated complexity indices will be enhanced by a more qualitative treatment of selected features evincing clausal and phrasal complexity.

3. Data and methodology

The analysis of linguistic complexity in professional academic writing has been conducted on a corpus of research papers published in peer-reviewed journals. Four disciplines fall into the category of so-called 'hard' science, and four into that of 'soft' science, thus representing a broad cross-section of academic discourse. The labels 'hard' and 'soft', commonly attributed to Storer (1967), are used to compare scientific fields on the basis of perceived methodological rigour, exactitude and objectivity. In a nutshell, the applied, empirical, experimental and natural disciplines (e.g. astronomy, biology, mathematics, physics) are considered 'hard', whereas the social sciences (e.g. history, linguistics, literature, sociology, political science) are categorised as 'soft'. Even though as stated by some researchers, the hard/soft division does not always adequately reflect the existing variation in the structure of knowledge in different disciplines (see, for example, Becher & Trowler, 2001; Nesi, 2002), this division can serve as a useful shorthand when attempting to describe the diversity of academic discourse (Dang, 2018). In this study, the hard-science subcorpus comprises articles in chemistry, physics, mathematics and engineering, and the soft-science subcorpus consists of texts in business studies, history, linguistics and political-science research articles. All the articles were published in leading academic journals, indexed in Scopus Quartile 1, in 2016-2020. This makes our corpus more up-to-date and homogeneous than, for instance, the Corpus of Academic Journal Articles (CAJA; Kosem, 2010), in which the category of journal articles also includes reports, reviews and progress reports published between 1993 and 2008. In our corpus we have aimed at warranting balance in terms of the number of tokens within each discipline. The texts were formatted for further textual analysis; for example, tables, formulas, graphs, charts, metadata and reference lists were removed from the documents. The size and details of the corpus are given in Table 1.

Table 1: Corpus

Discipline	No. texts	Word totals	*Journals*
HARD SCIENCES			
Chemistry	34	197,806	*Cell Chemical Biology (CCB)* *Chem* *Chemical Science (CS)* *Trends in Analytical Chemistry (TrAC)*
Physics	44	200,206	*Physics Letters B (PL)* *Reviews in Physics (RP)* *European Physical Journal C (EPJ)* *Nuclear Physics B (NPh)*
Mathematics	28	199,380	*Compositio Mathematica (CM)* *The Journal of Differential Geometry (JDG)* *Acta Mathematica (ActaM)* *Applied Mathematics and Computation (AMC)*
Engineering	34	198,926	*Automatica (Auto)* *Materials Characterisation (MC)* *International Journal of Engineering Science (IJES)* *Engineering (Engin)*
Totals	140	796,318	
SOFT SCIENCES			
Business	20	197,956	*The Journal of Management (JM)* *The Journal of Management Studies (JMS)* *Academy of Management Annals (AMA)* *Journal of Business Research (JBR)*
Linguistics	22	200,997	*Applied Linguistics (AL)* *Lingua (Ling)* *Modern Language Journal (MLJ)* *Language in Society (LS)*
History	21	199,394	*Contemporary European History (CEH)* *The Journal of Modern History (JMH)* *Journal of Global History (JGH)* *History of the Family (HF)*
Political science	25	202,040	*Political Analysis (PA)* *World Politics (WP)* *American Journal of Political Science (AJPS)* *British Journal of Political Science (BJPS)*
Totals	88	800,387	

This study undertakes both the quantitative analysis of measures automatically generated by the complexity analyser and the qualitative scrutiny of a number of syntactic patterns associated with syntactic complexity. Firstly, to accomplish the quantitative analysis, the corpus texts were processed using Lu's L2SCA (see Section 2.2). This software was chosen because, as stated in Lu (2010), the analyser's precision/recall rates and F-score are high, with an accuracy of 0.83+. Also, the complexity indices in L2SCA were identified specifically for the analysis of academic discourse, which makes them optimal for the purposes of this research. L2SCA provided the 14 indices given in Table 2 along with their descriptions, as in Lu (2011: 43). Such indices were categorised into:

i. metrics of structural complexity: indices reporting the length of units (sentences, T-units, clauses), measured by counting the number of words

ii. metrics of syntactic complexity: indices reflecting syntactic depth and dependency, that is, those based on coordination and subordination ratios as well as on clausal/T-unit embedding within other superordinate units

iii. metrics of categorial complexity: indices expressing the pervasiveness of nominal and verbal categories in the text, which have been identified as key measures of genre complexity in the literature (see Section 2.2).

Table 2: L2SCA syntactic complexity indices

Structural complexity		MLS	mean length of sentence (no. of words)
		MLT	mean length of T-unit (no. of words)
		MLC	mean length of clause (no. of words)
Syntactic complexity	Coordination	CPC	coordinate-phrase/clause ratio
		CPT	coordinate-phrase/T-unit ratio
	Subordination	CS	clause/sentence ratio
		CT	clause/T-unit
		TS	T-unit/sentence ratio
		DCC	dependent-clause/clause ratio
		DCT	dependent-clause/T-unit ratio
		CTT	complex-T-unit/T-unit ratio
Categorial complexity	Predicates	VPT	verb-phrase/T-unit ratio
	Nominals	CNT	complex-nominal/T-unit ratio
		CNC	complex-nominal/clause ratio

The indices and the disciplines have been modelled statistically (see Section 4 for the detailed description of the analyses) by means of a number of multivariable methodologies (regression, Random Forests and clusters) in an attempt to both weigh the contribution of the individual complexity indices to the hard/soft distinction, and to determine similarities/differences among the hard and the soft disciplines as far as linguistic complexity is concerned.

Secondly, the corpus was tagged with TagAnt (Anthony, 2015), which employs the TreeTagger tagset (available at https://www.sketchengine.eu/english-treetagger-pipeline-2/) and processed with AntConc (Anthony, 2014), in order to carry out the qualitative analysis of the clausal and the phrasal complexity features in Table 3, based on the taxonomy in Staples et al. (2016).

Table 3: Clausal/phrasal complexity indices

Feature	Example
Clausal-complexity features	
Finite adverbial clauses of purpose introduced by the conjunctions *in order that*, *so that*	Moreover, p can be chosen *so that the next property is satisfied* (CM-2016-3).
Finite adverbial clauses of condition introduced by the conjunctions *if*, *unless*, *in the event that*, *provided that*	*Unless action was taken,* it might grow into a serious danger in a very short time (JMH-2016-2).
Finite adverbial clauses of concession introduced by the conjunctions *although*, *even though*, *despite the fact that*	*Although the above papers did not discuss switching speed,* typical integrated thermo-optical and electro-optical switching can reach GHz rates (RP-2016-5).
Finite adverbial clauses of time introduced by the conjunctions *after*, *before*, *when*, *until*, *as soon as*, *as*	No action could reasonably be taken *before the Western powers had given the signal* (CEH -2016-5).
Finite adverbial clauses of place introduced by the conjunction *where*	Once transcribed, the data are encoded with instances *where gestures…enacted by the right, left, and both hands being marked up* (AL-2016-2).
Finite adverbial clauses of reason introduced by the conjunctions *because*, *since*, *as*	*Since we observed a time-dependent accumulation of very long chain fatty acids,* we also investigated the expression levels of genes that showed significant accumulations during early necroptosis (CCB-2017-1).

Finite adverbial clauses of result introduced by the conjunction *so that*	The learning conditions were counterbalanced, *so that each participant learned half of the critical items in the WW condition and half in the ME condition* (AL-2016-4).
Finite adverbial clauses of manner introduced by the conjunctions *as if*, *as though*, *as*	In Western Europe the political discussions on Eureka were mostly conducted *as if the Soviet Union and the Cold War divide did not exist* (CEH-2016-2).
Finite adverbial clauses of contrast introduced by the conjunctions *while*, *whereas*	*While these adverbs have a basic restrictive function* which can be accounted for at the RL (Section 3.2.1.), they will be shown to have a number of different functions at the IL (Ling-2016-2).
Wh-complement clauses	The Chiefs of Staff of fronts and armies and scouting units do not know *where captives came from*...(JMH-2016-4).
Verb + *that*-clauses	Lagrangian is considered in the Einstein and in the Jordan frame, and we *demonstrated that several cosmological scenarios can be realised* (PL-2017-3).
Phrasal complexity features	
Nouns	Finally, phonetic *cues* as contrastive *stress* have been pointed out as another *factor* in determining the *interpretation* of *pronouns* (Ling-2016-1).
Attributive adjectives	The *crucial* role of *topological* defects was observed in a *new* type of phase transition in *two-dimensional* systems (PL-2016-1).
Premodifying nouns	Due to *space* limitations, we report the analyses including the six dummies for *company* level, but not including the statistical details of these dummies (JM-2016-3).
Of-genitives	In the absence *of enforceable global governance regimes*, the social responsibilities *of corporations* take on a new explicit political dimension (JMS-2016-2).

The qualitative analysis of the data required extensive manual pruning and the careful interpretation of the examples, for example, in order to determine the semantic

type of the adverbial clauses – to illustrate this, the conjunction *as* can introduce adverbial clauses of time (1), of reason (2) and of manner (3):

1. Moreover, *as the concentration of FabX was increased*, a second product appeared (CCB-2016-1)

2. *As the ammosamides share the same core structural features of lymphostin (Figure 1)*, they provide a unique opportunity to explore pyrroloquinoline alkaloid biosynthesis in a distinct genomic context (CCB-2016-2)

3. it should be possible to generate glycOMVs displaying a wide array of biomedically relevant glycotopes found on the surfaces of bacteria and human cells, *as we demonstrated here with T antigen and PSA* (CCB-2016-3)

4. Analysis of complexity

This section deals with the analysis of the complexity indices of the hard- and soft-science papers automatically generated by L2SCA (Section 4.1), and with the variation in the realisation of the syntactic complexity features in the same corpus (Section 4.2).

4.1. Automated complexity metrics

The mean values for each of the indexes per category (hard/soft) and discipline are given in Table 4.

Table 4: L2SCA syntactic complexity indices in hard/soft sciences

Index	Hard sciences					Soft sciences					Mean
	chemistry	physics	mathematics	engineering	*mean*	business	linguistics	history	political-sc	*mean*	
MLS	30.61	25.71	24.48	26.10	*26.72*	29.17	30.99	51.52	32.35	*36.06*	*31.43*
MLT	28.70	24.36	23.32	24.57	*25.23*	27.68	28.26	46.08	29.14	*32.83*	*29.06*
MLC	18.68	15.14	13.83	16.28	*15.98*	16.13	15.07	23.38	15.02	*17.42*	*16.70*
CPC	0.52	0.30	0.23	0.44	*0.37*	0.63	0.43	0.44	0.32	*0.45*	*0.41*
CPT	0.80	0.48	0.38	0.65	*0.58*	1.08	0.79	0.86	0.62	*0.84*	*0.71*
CS	1.65	1.72	1.76	1.61	*1.68*	1.83	2.07	2.22	2.16	*2.07*	*1.88*
CT	1.55	1.62	1.68	1.51	*1.59*	1.73	1.88	1.96	1.94	*1.88*	*1.74*
TS	1.07	1.06	1.04	1.06	*1.06*	1.06	1.10	1.13	1.11	*1.10*	*1.08*
DCC	0.34	0.36	0.40	0.32	*0.35*	0.41	0.44	0.44	0.46	*0.44*	*0.40*
DCT	0.54	0.60	0.68	0.50	*0.58*	0.73	0.85	0.89	0.91	*0.85*	*0.71*
CTT	0.39	0.41	0.46	0.36	*0.40*	0.50	0.54	0.54	0.56	*0.54*	*0.47*
VPT	2.17	2.17	2.22	2.01	*2.14*	2.50	2.69	2.72	2.71	*2.66*	*2.40*
CNT	3.82	3.37	2.92	3.24	*3.34*	3.9	3.89	4.76	3.85	*4.10*	*3.72*
CNC	2.47	2.10	1.74	2.18	*2.12*	2.27	2.08	2.44	2.00	*2.20*	*2.16*

In an attempt to determine the relative weights of the complexity indices within a multivariate model, a binomial linear regression analysis was applied to the data, implemented via the functions 'glm' in the 'stats' package (R Core Team, 2022).[3] More specifically, the model was intended to identify the complexity indices that have a significant determining effect on the categorisation of research writings into the hard- or the soft-science categories (i.e. the dependent variable). When the 14 complexity indices identified by L2SCA (Table 2) entered the model, the collinearity among indices revealed by the function 'vif' (Variance Inflation Factor, 'car' package; Fox & Weisberg, 2019) was severe. This statistical function indicates the presence in the model of predictors or variables (i.e. indices) that are not significantly informative because of mutual convergence with the values provided by other factors. This weakness of the model was overcome by operationalising the reduction of the number of indices in the initial model with no statistically significant loss of explanatory power. Firstly, taking backward steps (step(emode_glm, direction="backward"; 'MASS' package) caused such a harmless reduction, which led to a reduced model with only the L2SCA indices MLT, MLC, VPT, DCC, CS, CT and CPT. Secondly, the regression model (glm(formula = hardsoft ~ mlt + mlc + c_s + vp_t + c_t + dc_c + cp_t, family = binomial) revealed that the complexity indices MLC and MLT (Mean length of T-unit) did not contribute significantly to the characterisation of hard/soft academic writings, as shown in (4):

(4) Summary of glm model (significance conventions: '***': 0.001, '**': 0.01, '*': 0.05, '': 0.1)

```
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.7559     4.1081  -1.158 0.246996
mlc          -0.3761     0.2432  -1.546 0.122000
mlt           0.2336     0.1432   1.632 0.102762
dc_c         15.2052     4.0120   3.790 0.000151 ***
c_t         -12.2593     3.0563  -4.011 6.04e-05 ***
vp_t          3.1525     0.7440   4.237 2.27e-05 ***
c_s           5.2790     0.9986   5.287 1.25e-07 ***
cp_t          3.1392     0.5108   6.145 7.98e-10 ***
```

3 In an attempt to check the effect associated with the random variability of the indices per individual text, a generalised linear mixed-effects model was implemented on top of the fixed model, where the variable 'file', encoding for each of the 234 file names in the corpus, served as the random variable. For that purpose, we used the 'glmer' function in R (glmer(hardsoft ~ (1|file) + dc_c + c_t + vp_t + c_s + cp_t, data=data, family=binomial, control= glmerControl(optimiser="bobyqa")); 'lme4' package, Bates et al., 2015). The difference between the mixed- and the fixed-effects models' AICs (395.02 and 393.29, respectively) did not prove to be statisticaly significant (p=0.6072), which also leads to the conclusion that the mixed-effects model does not add explanatory force to the generalised linear model with the 4 fixed factors (indices).

In fact, dropping MLC and MLT from the dataset did not trigger a substantial statistical loss of the model's explanatory power, the difference between the models' AICs with and without MLC+MLT not being significant. As a consequence, we have opted for the simplest model summarised in (5), with only the indices VPT (Verb phrases per T-unit), DCC (Dependent clause ratio), CS (clause/sentence ratio), CPT (Coordinate phrases per T-unit) and CT (clause/T-unit).

(5) Summary of definitive glm model (significance conventions: '***': 0.001, '*': 0.05)

```
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -11.0875     1.2430  -8.920  < 2e-16 ***
dc_c         13.9837     3.8285   3.652  0.00026 ***
c_t          -8.1453     1.8276  -4.457 8.32e-06 ***
vp_t          3.3050     0.7395   4.469 7.84e-06 ***
c_s           5.1488     0.9829   5.239 1.62e-07 ***
cp_t          3.1483     0.4855   6.484 8.91e-11 ***
```

The VIF results are now low, ranging from 1.15 to 8.53, which reflects a lack of severe collinearity in the definitive model. Also, both the C(oncordance) 0.895 and Nagelkerke R^2 0.586 discrimination indices provided by 'lrm' ('rms' package; Harrell, 2021) indicate that the model is very good at explaining the variation (C>0.9 reveals the model's outstanding fit and predictive power, and R^2>0.4 it plausibility) and, consequently, adequate for the research question.

Logistic regression was used to assess the significance of the contribution of the indices to the overall categorisation of texts into hard and soft sciences. Random Forests (function 'cforest', 'party' package; Hothorn et al., 2006), first applied to linguistic analysis by Tagliamonte & Baayen (2012), rank predictors according to their impact on the explanation of the variation. Figure 1 presents the Random Forests corresponding to the model's fixed predictors. The goodness of fit of the model, as indicated by a C-index of 0.9487 computed by the function 'somers2' ('Hmisc' package; Harrell et al., 2020), is excellent and, as expected, slightly better than that of the regression model (C=0.895).

Figure 1: Dot chart of conditional variable importance

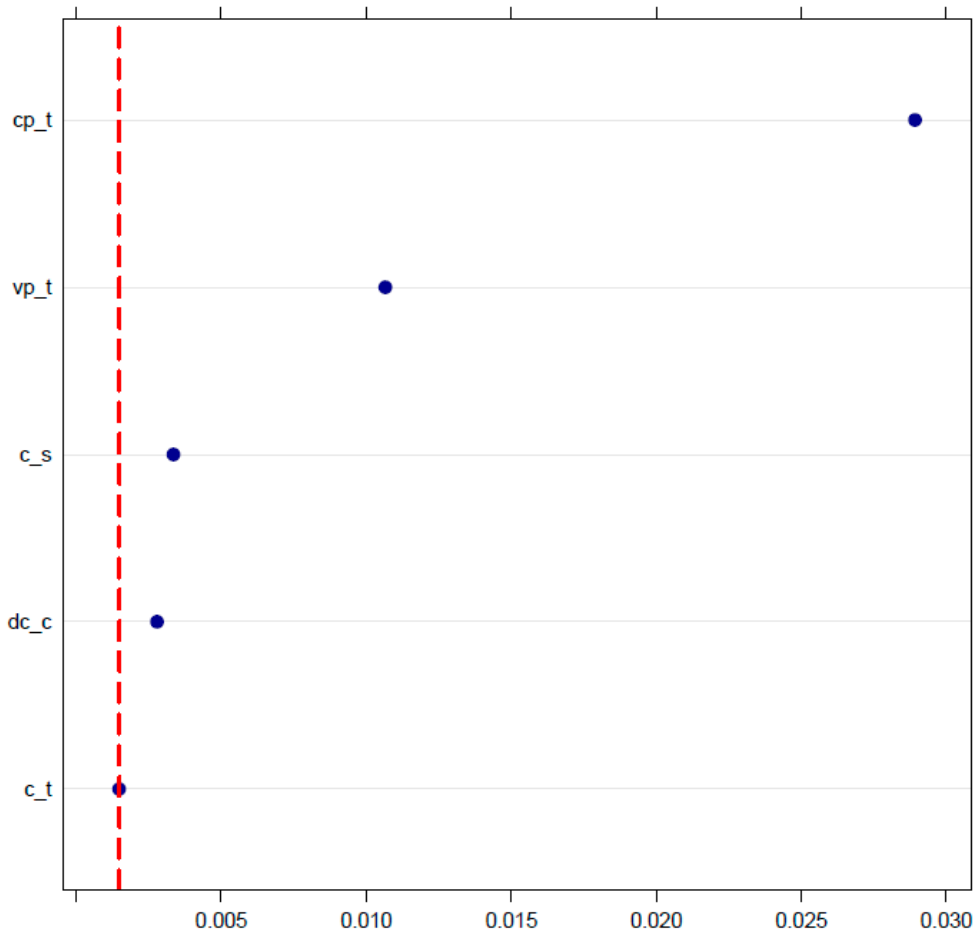
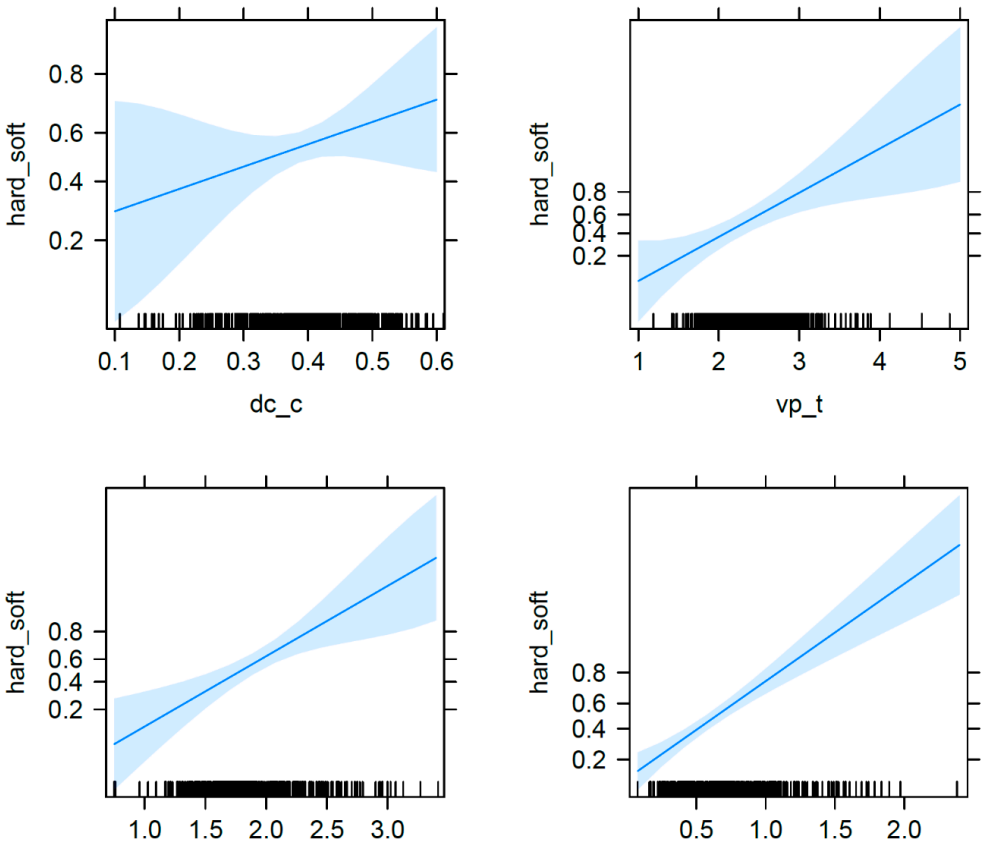


Figure 1 reflects, first, that the CT index exerts a very weak influence on the model, as indicated by its very low conditional importance value, and thus paves the way for the disposal of this variable. Second, Figure 1 evinces the significant impact of the indices CPT, VPT and DCC on the variation hard/soft science. The effects plots in Figure 2 provide a more detailed picture of the correlation between the significant indices and the categorisation of the research articles.

Figure 2: Effect plots



The interpretation of the findings revealed by the statistical analysis of the complexity indices per broad discipline, that is, hard and soft sciences, is as follows. Firstly, as a consequence of the overlap of units such as sentence, clause and T-unit in specifically formal academic writing, the regression analysis showed severe collinearity among the initial 14 complexity indices. The reduction of the indices led to a model with only 4 indices evincing different dimensions of linguistic complexity:

i. syntactic complexity mirrored by pervasive coordination, as reflected by the index CPT, which calculates the ratio of coordinated phrases per T-unit

ii. syntactic complexity determined by subordination within clausal units, as evinced by the index DCC, which expresses the number of subordinate dependent clauses in matrix clauses, and in sentences, which has been corroborated by the statistical significance of the index CS, a telling indicator of the ratio of clauses per sentence

iii. (iii) categorial complexity associated with the frequency of, specifically, verbal constituents in T-units, here captured by the index VPT.

Random Forests and the analysis of effects have demonstrated, on the one hand, that, out of the indices that proved to be very strong in the model, those measures evincing complexity triggered by coordination (CPT) and by the profusion of verbal categories (VPT), contribute to the variation of hard *versus* soft science to a greater extent than DCC and CS. On the other hand, the probability of higher values in the four complexity indices increases in academic writings categorised as soft science. In other words, greater ratios of coordination, subordination and the 'verby' status of texts can be taken as proxies for the categorisation of a research paper within the domain of social sciences and humanities.

These results give support to Biber et al. (2011: 29; similarly, Biber & Gray, 2011) when they claim that "complexity is not a single unified construct, and it is therefore not reasonable to suppose that any single measure will adequately represent this construct". However, some remarks are in order here as regards the interpretation of our findings in light of the conclusions drawn by Biber and colleagues. In their multidimensional analysis of academic writing *versus* other more informal genres, Biber et al. (1999, 2013) found that high(er) phrasal complexity and low(er) clausal complexity are characteristic features of academic English (as well as of newspaper and magazine writings). By contrast, the type of complexity evinced in personal, professional (even academic) spoken genres, as well as in popular written (novels, personal essays) discourse, is fundamentally clausal. Specifically, they contend that T-unit- and subordination-based (i.e. clausal) measures are not typical of academic writing but of conversational discourse, whereas nominal/prepositional (i.e. phrasal) measures are good indicators of academic writing. The statistical modelling of the complexity indices reported in this section has shown that subordination, coordination and the 'verby' status of sentences (or, better, T-units) are defining features of soft academic writing. As we see it, this conclusion does not invalidate a dominantly phrasal characterisation of academic writing when compared to more informal speech-based/related discourse, but gives support to the multifaceted nature of academic writing.

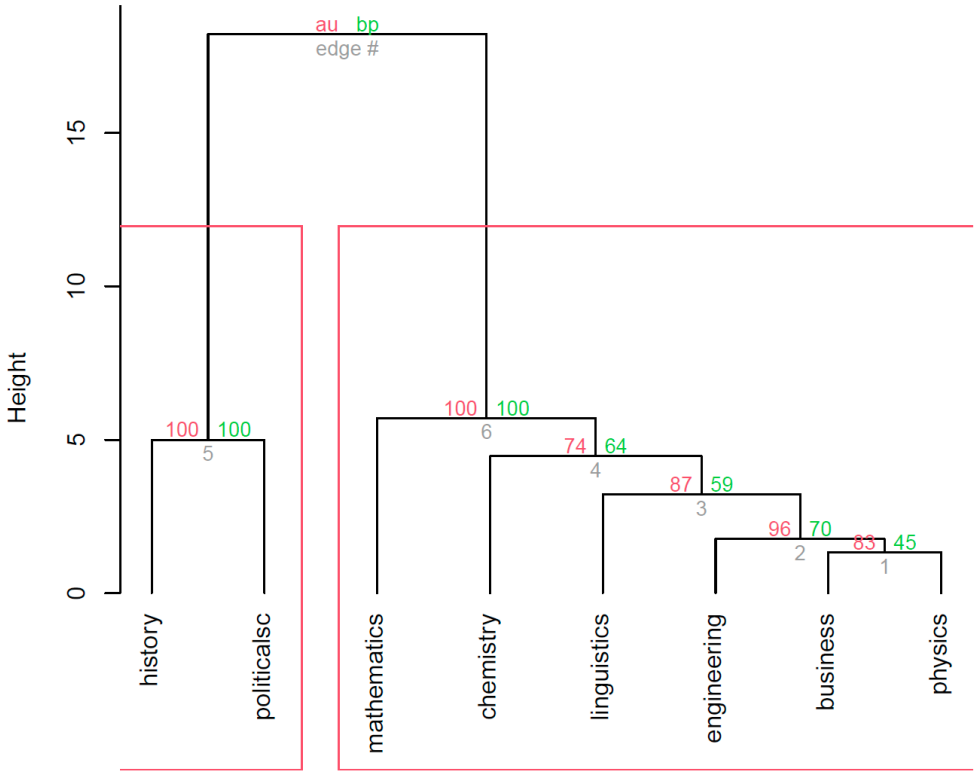
Our data confirm that, within the academic genre, the complexity strategies that serve to categorise soft- and hard-science articles are different, a continuum being observed between applied, empirical, experimental and natural disciplines, and social sciences and humanities as regards the productivity of, for example, coordination or subordination. This finding is in keeping with previous studies that highlighted substantial differences among academic disciplines, and sub-registers or sub-genres as regards the use of complexity measures. To give a few examples, Gardner et al. (2019: 670) analysed successful university student writing (about two thirds of which

was written by students who declared English as their first language) and found that "the writing situation – disciplinary group [arts and humanities, life sciences, physical sciences and social sciences], genre family, discipline, and level of study – is key to interpreting each dimension", resulting from the application of the multidimensional analysis of lexico-grammatical features associated with different aspects of linguistic complexity. Also, Hardy & Friginal (2016) confirmed that there is a continuum of academic paper types even within each of the dimensions recognised by Biber's multidimensional analysis. Our results also align with, for example, Nesi & Gardner's (2019), who, as already mentioned in Section 2.1, concluded that clausal complexity is more prominent in the so-called soft disciplines and in the more conversational genres[4].

A hierarchical agglomerative clustering algorithm was applied ('hclust' function, 'ward.D2' method, 'pvclust' package; Suzuki et al., 2019) to identify subgroups of disciplines based on the pervasiveness of all the complexity indices. The analysis used a Behavioural Profiles approach, in which the data are represented as vectors of proportions of each level of each variable ('bp' function, 'pvclust' package). With this technique, the numerical differences between vectors are operationalised as 'distances' ('dist' function, 'canberra' method, 'pvclust' package), which determine how the disciplines are grouped into clusters. The optimal number of clusters was calculated by means of the function 'silhouette' ('cluster' package; Maechler et al., 2019). The clusters are represented as tree leaves or branches of dendrograms (Levshina, 2015: 316), the most similar of which (i.e. those with the smallest 'distance') are merged together. Figure 3 displays a dendrogram of the clustering of the varieties.

4 The detected differences between subdisciplines are discussed in detail in Section 4.2.

Figure 3: Hierarchical cluster analysis of disciplines



The eight disciplines were grouped into the two statistically optimal clusters represented by the boxes in Figure 3. The stability of the clusters and their fit with the data was measured by the function 'pvclust' ('pvclust' package), which quantifies the uncertainty in the clusters by implementing multiscale bootstrap resampling to calculate the Approximately Unbiased (AU) *p*-values of each – the closer AU *p* is to 1, the greater the statistical significance of the cluster.

This technique identified two groups of disciplines: on the one hand, history and political science (both soft sciences), and all the remaining disciplines, on the other. To check the plausibility of such a grouping of disciplines, we carried out a qualitative analysis of textual samples randomly selected from the corpus of research papers. In detail, we extracted complete sentences, amounting to at least 150 words, from texts belonging to the disciplines of linguistics (in (6)) and business (in (7)), clustered together with the other hard-science texts, and of history (in (8)), a discipline included in the first (soft-science) cluster. The frequencies of finite verbal groups (in italics in

(6)-(8)), of finite subordinate clauses (starting with '[' in (6)-(8)) and of coordinating conjunctions (in boldface italics), have all been explored in the samples in an attempt to compute the most significant complexity indices of the hard *versus* soft distinction.

(6) [Because L2 proficiency *is* an important predictor of contextual vocabulary learning ***and*** only limited control over participants' proficiency *was* possible at recruitment, their L2 lexical proficiency *was* further estimated using the following published instruments: LexTALE *was* used as a measure of receptive vocabulary knowledge, and Laufer ***and*** Nation's (1999) vocabulary levels test of controlled-productive ability (PVLT) *was* used to measure their productive vocabulary knowledge (Table 1). PVLT *was* measured at the 2,000 ***and*** 5,000 word frequency levels, ***and*** the average score *was* used in the data analyses. Furthermore, [because larger working memory *tends* to positively correlate with word learning in L1 ***and*** L2, ***and*** [because both word-writing ***and*** meaning deliberation *may* consume the limited processing resources needed to create form-meaning associations (Barcroft 2006), participants' working memory *was* measured using an Operation Span (O-Span) task. Individual L2 vocabulary scores ***and*** working memory (O-Span) scores *were* included as covariates in the data analyses of the immediate ***and*** delayed tests. (AL2016-4)

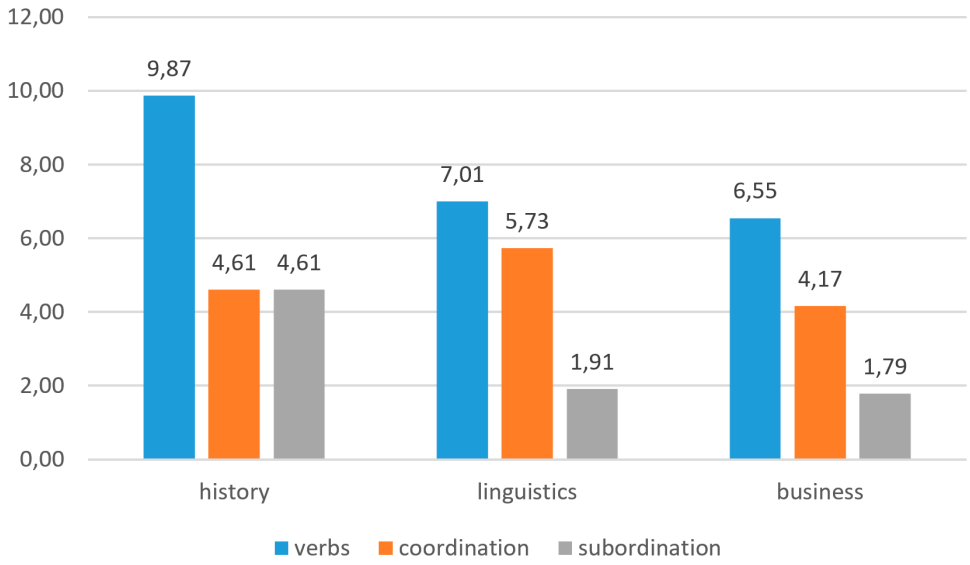
(7) In the previous section, we *have demonstrated* [that political CSR pursued by an organisation *is* related to individual-level CEO characteristics, i.e., their value orientation ***and*** subsequent behaviour. We *argued* [that CEOs with a social welfare orientation *are* likely to display an integrative responsible leadership style ***and*** *motivate* their organisations to engage in substantive political CSR (high-involvement MSI, second-order social innovation), [while CEOs with a strong sense of fiduciary duty *are* more likely to practice an instrumental responsible leadership style ***and*** *pursue* incremental political CSR (low-involvement MSI, first-order social innovation). We *will examine* in the following factors at the individual, organisational ***and*** societal level [that *moderate* the expected relationships between value orientations ***and*** leadership styles, emphasizing the relationship between social welfare orientation ***and*** an integrative responsible leadership style. Against the backdrop of political CSR we *discuss* exemplary, multilevel contingencies: at the individual level, the ability of CEOs to cope with complexity *is* a precondition to be able to respond to the complex institutional ***and*** relational environment of global business. (JMS2016-2)

(8) Apart from 'work on the party line' ***and*** similar tasks within the category of ideological work ***and*** propaganda, the committee *stressed* [that 'the entire work *is* focused on preparing ***and*** grooming the fighters for the struggle for national liberation'. A clandestine unit of 150 people *was created* ***and*** *organised* as a partisan military detachment, ready to join the partisans [when the moment *came*. It *was* clear [that these preparations *were* most relevant for young, able-bodied men ***and*** the few women

[who *could join* the partisan units in battle. [What *was* less clear *was* [what *was going to happen*, [when such opportunity *arose*, to the people unsuitable for partisan life: the sick, the elderly ***and*** children. The moment *came* soon enough. On 8 September 1943 the news of the armistice between Italy ***and*** the allies *reached* the camp. The exact chronology of events in the camp over the next few days *remains* murky. (CEH-2016-3)

The normalised frequencies of the three complexity strategies in the texts, displayed in Figure 4, corroborate the deviation of the linguistics and the business texts from the soft-science of history ones, the only exception to this clear-cut trend being the frequency of coordination in the linguistics extract, and, in consequence, the plausibility of the clusters sketched in Figure 3.

Figure 4: Complexity measure in subdiscipline samples



4.2. Syntactic complexity features

This subsection is devoted to a qualitative analysis of the frequencies of the features associated with clausal and phrasal complexity, as described in Section 3. Table 5 provides the raw and normalised frequencies (per 100,000 words) of the 16 clausal/phrasal complexity features as well as the measures of statistical significance (*p*-values resulting from chi-square test) of the variation hard- *versus* soft-sciences – statistical significance has been conventionalised as follows: '***' when $p \leq .001$ and '**' when $p \leq .01$.

Table 5: Clausal/phrasal complexity features in hard/soft sciences: raw, (normalised frequencies)

Feature	Hard sciences	Soft sciences	χ^2	*p*-value	
Finite adverbial clauses:					
Purpose	16 (2.01)	17 (2.13)	–	1	
Condition	1545 (194.10)	836 (104.50)	213.83	<0.00001	***
Concession	241 (30.28)	538 (67.25)	110.92	<0.00001	***
Time	733 (92.09)	1442 (180.25)	226.56	<0.00001	***
Place	1842 (231.41)	678 (84.75)	541.82	<0.00001	***
Reason	1195 (150.13)	1142 (142.75)	1.44	0.2309	
Result	176 (22.11)	46 (5.75)	75.61	<0.00001	***
Manner	113 (14.20)	198 (24.75)	22.26	<0.00001	***
Contrast	618 (77.64)	828 (103.50)	29.13	<0.00001	***
Adverbial clauses total	*6479 (813.94)*	*5725 (715.63)*	*50.00*	*<0.00001*	***
Wh-complement clauses	53 (6.66)	371 (46.38)	235.33	<0.00001	***
Verb+ *that*-clauses	3566 (447.99)	3964 (495.50)	18.87	0.000014	***
Nouns	224367 (28186.81)	219641 (27455.13)	60.43	<0.00001	***
Attributive adjectives	51917 (6522.24)	56727 (7090.88)	177.01	<0.00001	***
Premodifying nouns	40102 (5037.94)	30583 (3822.88)	1274.2	<0.00001	***
Of-genitives	6995 (878.77)	9219 (1152.38)	290.61	<0.00001	***

Most of the differences in the use of the complexity features in hard and in soft sciences are statistically significant. As for the clausal complexity features, overall, adverbial clauses were found to be more frequent in the corpus of the hard-science papers. Unlike in Staples et al. (2016), which examined finite adverbial clauses 'in bulk', Table 5 provides the frequencies of the different semantic types of adverbial clauses, which manifest significant differences across the two broad hard/soft categories. As regards the two features evincing complementation strategies, *wh*-clauses and *that*-clauses prevail in the soft research papers. Finally, the trends revealed by the data as far as phrasal complexity is concerned are, first, the preference for verbal, adjectival and prepositional phrases in the soft-science texts and, second, for nominal categories in the hard sciences.

Table 6 gives the distribution of the complexity features across individual disciplines, with raw and normalised frequencies per 100,000 words. To measure the dispersion of the complexity features in the corpus we used Juilland's D, which is considered to be the most reliable dispersion coefficient (Rayson, 2003: 94), ranging from 0 (a perfectly uneven distribution) to 1 (a perfectly even distribution). SD and Juilland's D were calculated for relevant frequency values.

Table 6: Clausal/phrasal complexity features across disciplines: raw, (normalised frequencies)

Feature	Hard sciences						Soft sciences					
	Chemistry	Engineering	Maths	Physics	SD	D-dispersion	Business	History	Linguistics	Political science	SD	D-dispersion
CLAUSAL COMPLEXITY												
Finite adverbial clauses:												
Purpose	2 (1.01)	4 (2.01)	3 (1.51)	7 (3.5)	0	0.59	6 (3.03)	6 (3.02)	4 (1.99)	1 (0.5)	0	0.59
Condition	64 (32.32)	300 (150.75)	954 (479.40)	227 (113.5)	0	0.44	143 (72.22)	176 (88.44)	175 (87.06)	342 (169.31)	0	0.60
Concession	109 (55.05)	65 (32.66)	25 (12.56)	42 (21)	0	0.59	209 (105.56)	104 (52.26)	96 (47.76)	129 (63.86)	0	0.66
Time	146 (73.74)	203 (102.01)	171 (85.93)	213 (106.5)	0	0.66	371 (187.37)	297 (149.25)	266 (132.34)	508 (251.49)	0	0.61
Place	120 (60.61)	384 (192.96)	694 (348.74)	654 (327)	0	0.63	117 (59.09)	130 (65.33)	189 (94.03)	242 (119.80)	0	0.63
Reason	176 (88.89)	198 (99.50)	565 (283.92)	256 (128)	0	0.37	271 (136.87)	183 (91.96)	334 (166.17)	354 (175.25)	0	0.63
Result	7 (3.54)	25 (12.56)	110 (55.28)	34 (17)	0	0.35	4 (2.02)	2 (1.01)	25 (12.44)	15 (7.43)	0	0.44
Manner	5 (2.53)	18 (9.05)	24 (12.06)	66 (33)	0	0.48	45 (22.73)	74 (37.19)	38 (18.91)	41 (20.3)	0	0.6
Contrast	159 (80.30)	165 (82.91)	64 (32.16)	230 (115)	0	0.55	187 (94.44)	161 (80.90)	232 (115.42)	248 (122.77)	0	0.65
Adverbial clauses total	*788 (397.98)*	*1362 (684.42)*	*2560 (1286.43)*	*1769 (884.5)*	*0*	*0.62*	*1353 (683.33)*	*1133 (569.35)*	*1359 (676.12)*	*1880 (930.69)*	*0*	*0.64*
Wh-complement clauses	14 (7.07)	10 (5.03)	13 (5.53)	16 (8)	0	0.64	185 (93.43)	28 (14.07)	78 (38.81)	80 (39.6)	0	0.58
Verb + *that*-clauses	628 (317.17)	811 (407.54)	1419 (713.07)	708 (354)	0	0.60	1054 (532.32)	848 (426.13)	952 (473.63)	1110 (549.5)	0	0.66
PHRASAL COMPLEXITY												
Nouns	58014 (29300.00)	57920 (29105.53)	51802 (26031.16)	56631 (28315.5)	0.12	0.67	58820 (29707.07)	49868 (25059.3)	52834 (26285.57)	58119 (28771.78)	0.12	0.66
Attributive adjectives	13902 (7021.21)	13189 (6627.64)	10579 (5316.08)	14247 (7123.5)	0.03	0.66	14696 (7422.22)	14512 (7.292.46)	12943 (6439.30)	14576 (7215.84)	0.03	0.66
Premodifying nouns	11023 (5567.18)	11460 (5758.79)	7448 (3742.71)	10171 (5085.5)	0.02	0.65	9375 (4734.85)	5646 (2837.19)	6902 (3433.83)	8660 (3936.36)	0.02	0.65
Of-genitives	1765 (894.41)	1725 (866.83)	1817 (913.07)	1688 (844)	0	0.67	2481 (1253.03)	2288 (1149.75)	2165 (1077.11)	2285 (1131.19)	0	0.67

Table 6 shows, on the one hand, that the distribution of phrasal complexity features is more even than that of clausal complexity features, which suggests that there is less disciplinary variation in their use. On the other hand, Table 6 reveals that the distribution of the clausal complexity features is more balanced in soft sciences than in the hard disciplines. The hard science that stands out in this respect is mathematics, with an extensive use of adverbial clauses of condition, place and reason. In the soft category, political-science texts noticeably demonstrate greater frequencies of adverbial clauses of time and reason.

Some final remarks are in order here concerning our fine-grained analysis of the automated indices per subdicipline and the, on occasions, intrinsic characteristics of academic writing. First, the frequencies for long sentences and for complex nominals, as identified by the analyser, were very salient in the chemistry texts, and this is partially due to the pervasiveness of outstandingly long names of chemical entities and processes in the field (see, in this respect, Dai et al., 2015), as illustrated in, respectively (9) and (10):

(9) In addition, *methyliminodiacetic acid (MIDA)-protected boronate esters* were well tolerated (Chem-2016-4)

(10) We have demonstrated a new pathway of *unsaturated fatty acid synthesis* that is catalyzed by an enzyme, FabX, that has dual dehydrogenase/isomerase activities (CCB-2016-1)

Also, the coordination indices turned out to be outstanding in chemistry, when compared with the other hard sciences. This could be explained by the large number of descriptions of chemical experiments in the textual data, which usually involve several steps and deal with several chemical entities, as in (11).

(11) FabX was monitored at 280 nm *and* eluted at 15.26 min. (...) The standards were vitamin B12 (1.35 kDa), myoglobin (horse, 17 kDa), ovalbumin (chicken, 44 kDa), g-globulin (bovine, 158 kDa), *and* thyroglobulin (bovine, 670 kDa) (CCB-2016-1)

Another deviant hard-science subdiscipline was mathematics, where the high frequency of condition (12), place (13) and reason (14) adverbial clauses correlates with the latter's extensive use in the comments for calculations and formulas.

(12) *If we apply the recurrences σ1σ3 ... σn−1 to an arbitrary ca0, a1, ... ,an*, we obtain a linear combination of lower coefficients (CM-2016-2)

(13) For example, let G = SL4(C), and let w = w0s3, *where w0 is the longest element in W = S4, the symmetric group with four letters* (CM-2016-1)

(14) *Since the off-diagonal factors of R are all even in x0, x2, ... , xn and P (x) is even*, the diagonal factors of R(cid:91)R(cid:92) must be even as well (CM-2016-2)

Mathematicians' writing was characterised by Davis & Hersh (1981: 36) as giving "an impression that, from the stated definitions, the desired results follow infallibly by a purely mechanical procedure". Such an impression can be supported by the use of adverbial clauses, whose rhetorical function in mathematics is to explain calculations and formulas. Interestingly, time and reason adverbial clauses abounded also in the political-science research papers, specifically in comments for mathematical and statistical models (see examples (15) and (16)).

(15) Conversely, multiple imputation will offer small gains in bias reduction *when variables of theoretical interest have a low proportion of missing values* (PA-2016-1).

(16) Such bias inducers may not be troublesome in practice, however, either *because they can be identified for exclusion, as is sometimes the case for post treatment variables*, or *because the bias they induce tends to be small* (PA-2016-3).

Physics was found to employ a large number of attributive adjectives which, as in chemistry, are often part of terms used in the discipline. To give a few examples, *interatomic* bonds, *local* minimum, *magnetic* field, *massive* gravity, *black* hole.

5. Summary and conclusions

This study has looked at the linguistic complexity of professional academic writing by analysing automatically generated complexity measures and frequencies of complexity features in a corpus of research papers in four 'hard' and four 'soft' sciences. As regards the first research question 'do hard and soft disciplines differ as regards the selection of linguistic complexity indices or metrics?', the automated analysis of the data and their statistical modelling have shown that soft sciences demonstrate more signs of syntactic complexity, particularly of subordination and coordination ratios, than the hard-science genre. Also, the data have revealed the (statistically significant) pervasiveness of verbs in the soft academic writings, as compared to the hard-science texts. In response to the second research question 'do hard and soft scientific writings differ as regards the productivity of linguistic complexity features?', it has been found that the clausal-complexity indices, particularly, the amount of adverbial clauses, are more revealing in the corpus of the hard-science papers, where they also demonstrate a greater degree of variation within the category. As for phrasal complexity, this has been shown in the preference for verbal, adjectival and prepositional phrases in the soft-science texts, and for nominal categories in the hard sciences, where the latter often instantiate genre-specific terminology.

As far as the limitations of the current study are concerned, first, although the significance of our results has been statistically verified at all times, the limited size of the corpus suggests that the empirical results must be taken with a pinch of salt. Second, as already claimed by Hyland (2004: 30) and also corroborated by this investigation, the classification of sciences into hard *versus* soft is not able to capture disciplinary variation to the fullest. Therefore, to provide a fuller picture of the realisation of complexity features in academic discourse, additional sub-disciplines and sub-genres are needed.

All in all, despite the recognised limitations and differences among subdisciplines, we contend that the teaching of EAP/ESP writing will greatly benefit from the scientific study of linguistic complexity in academic genres, and that the rigorous description of the core complexity strategies adopted in professional academic writing will guide the production of discipline-specific language-learning materials that will effectively address the needs of learners of different sciences.

6. References

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier & P. Thompson (eds), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). Amsterdam: John Benjamins.

Anthony, L. (2014). AntConc (Version 3.4.4) [Computer Software]. Tokyo: Waseda University.

Anthony, L. (2015). TagAnt (Version 1.2.0) [Computer Software]. Tokyo: Waseda University.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories* (2nd ed.). Philadelphia: Open University Press.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. London: Longman.

Biber, D. & Gray, B. (2011). Is conversation more grammatically complex than academic writing? In M. Konopka, J. Kubczak, Ch. Mair, F. Štícha & U.H. Waßner (eds), *Grammatik und Korpora 2009: Dritte Internationale Konferenz. Grammar & Corpora 2009: Third International Conference* (pp.7-61). Tübingen: Narr Verlag.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5-35.

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond T-units – A response to WeiWei Yang. *TESOL Quarterly*, *47*(1), 192-201.

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, *37*(5), 639-668.

Biber, D., Gray, B., Staples, S., & Egbert, J. (2020). Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes*, *46*, Article 100869.

Biber, D., Gray, B., Staples, S., & Egbert, J. (2021). *The register-functional approach to grammatical complexity: Theoretical foundation, descriptive research findings, application*. New York: Routledge.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (eds), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). Amsterdam: John Benjamins.

Casal, J.E., & Lee, J.J. (2019). Syntactic complexity and writing quality in assessed first year L2 writing. *Journal of Second Language Writing*, *44*, 51-62.

Casal, J.E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes*, *52*, Article 100996.

Chen, D., & Manning, C.D. (2014). A fast and accurate dependency parser using neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 740-750). Doha, Qatar.

Chen, X., & Meurers, D. (2016). CTAP: A web-based tool supporting automatic complexity analysis. In *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity at COLING, Osaka, Japan, 11th December.* (pp. 113-119). Osaka, Japan: The International Committee on Computational Linguistics.

Crossley, S.A., & McNamara, D.S. (2012). Predicting second language writing proficiency: The role of cohesion, readability, and lexical difficulty. *Journal of Research in Reading*, *35*(2), 115-135.

Crossley, S.A., & McNamara, D.S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, *26*, 66-79.

Crossley, S.A., Allen, L.K., & McNamara, D.S. (2014). A Multi-Dimensional Analysis of essay writing. What linguistic features tell us about situational parameters and the effects of language functions on judgements of quality. In T. Berber Sardinha & M. Veirano Pinto (eds), *Multi-Dimensional Analysis, 25 years on: A tribute to Douglas Biber* (pp.197-238). Amsterdam: John Benjamins.

Dai, H.J., Lai, P.T., Chang, Y.C., & Tsai, R. T.H. (2015). Enhancing of chemical compound and drug name recognition using representative tag scheme and fine-grained tokenisation. *Journal of Cheminformatics*, *7*, 1-14.

Dang, T.N.Y. (2018). The nature of vocabulary in academic speech of hard and soft-sciences. *English for Specific Purposes*, *51*, 69-83.

Davis, P.J., & Hersh, R. (1981). *The mathematical experience*. Boston: Birkhauser.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*. Thousand Oaks: Sage.

Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, *40*(4), 646-674.

Gray, B. (2013). More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles. *Corpora*, *8*(2), 153-181.

Gray, B. (2015). On the complexity of academic writing: Disciplinary variation and structural complexity. In V. Cortes & E. Csomay (eds), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber* (pp. 49-78). Amsterdam: John Benjamins.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193-202.

Hardy, J.A., & Friginal, E. (2016). Genre variation in student writing: A Multi-Dimensional Analysis. *Journal of English for Academic Purposes*, *22*, 119-131.

Harrell, F.E.Jr. (2021). Regression Modeling Strategies. https://github.com/harrelfe/rms

Harrell, F.E.Jr. with contributions from Charles, D. et al. (2020). Hmisc version 4.3-1. https://CRAN.R-project.org/package=Hmisc

Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly, 37*(2), 275-301.

Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006). Survival ensembles. *Biostatistics*, *7*(3), 355-373.

Hunt, K.W. (1964). *Differences in grammatical structures written at three grade levels: The structures to be analysed by transformational methods*. Report no. CRP-1998. Tallahasser: Florida State University.

Hunt, K.W. (1965). *Grammatical structures written at three grade levels*. NCTE Research Report No. 3. Champaign, IL: National Council of Teachers of English.

Hunt, K.W. (1970). Recent measures in syntactic development. In M. Lester (ed), *Readings in applied transformational grammar* (pp. 179-192). New York: Holt, Rinehert and Winston.

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.

Kelly-Laubscher, R. F., Muna, N., & van der Merwe, M. (2017). Using the research article as a model for teaching laboratory report writing provides opportunities for development of genre awareness and adoption of new literacy practices. *English for Specific Purposes*, *48*, 1-16.

Klein, D., & Manning, Ch.D. (2003). Fast exact inference with a factored model for Natural Language Parsing. In S. Becker, S. Thrun & K. Obermayer (eds), *Advances in neural information processing systems* (pp. 3-10). Cambridge, MA: MIT Press.

Kosem, I. (2010). *Designing a model for a corpus-driven dictionary of academic English*. PhD, Aston University.

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. PhD, Georgia State University, Atlanta, GA.

Kyle, K., & Crossley, S.A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal, 102*(2), 333–349.

Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics*, *35*(5), 607-614.

Lambert, C., & Nakamura, S. (2019). Proficiency-related variation in syntactic complexity: A study of English L1 and L2 oral descriptive discourse. *International Journal of Applied Linguistics*, *29*(2), 1-17.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Levy, R., & Andrew, G. (2006). Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*. (pp. 2231-2234). Genoa: ELRA.

Lintunen, P., & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language, 12*(4), 377-399.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, *45*(1), 36-62.

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, *34*(4), 493-511.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik Maechler, K. (2019). cluster: Cluster analysis basics and extensions. R package version 2.1.0.

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3-15.

McNamara, D.S., Louwerse, M.M., McCarthy, P.M., & Graesser, A.C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes*, *47*(4), 292-330.

Nesi, H. (2002). An English spoken academic wordlist. In A. Braasch & C. Povlsen (eds) *Proceedings of the Tenth EURALEX International Congress. Vol. 1*. (pp. 351-358). Copenhagen.

Nesi, H., & Gardner, S. (2019). Complex, but in what way? A step towards greater understanding of academic writing proficiency. In C. Danjo, I. Meddegama, D. O'Brien, J. Prudhoe, L. Walz & R. Wicaksono (eds.), *Online Proceedings of the 51st Annual Meeting of the British Association for Applied Linguistics: Taking Risks in Applied Linguistics, 6-8 September, 2018*. https://custom.cvent.com/01664CE00C344F7BA62E39C4CFE91FA8/files/0f77de05eb81461a8037170680562243.pdf (accessed on 22.06.2019)

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*, 492-518.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org

Rayson, P. (2003). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. PhD thesis, Lancaster University. https://eprints.lancs.ac.uk/id/eprint/12287/1/phd2003.pdf (accessed on 20.12.2020)

Ruan, Z. (2018). Structural compression in academic writing: An English-Chinese comparison study of complex noun phrases in research article abstracts. *Journal of English for Academic Purposes*, *36*, 37-47.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*(2), 149-183.

Storer, N.W. (1967). The hard sciences and the soft: Some sociological observations. *Bulletin of the Medical Library Association*, *55*(1), 75-84.

Suzuki, R., Terada, Y., & Shimodaira, H. (2019). pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R package version 2.2-0.

Swales, J.M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tagliamonte, S.A., & Baayen, R.H. (2012). Models, forests and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change*, *24*, 135-178.

Wijers, M. (2018). The role of variation in L2 syntactic complexity: A case study on subordinate clauses in Swedish as a foreign language. *Nordic Journal of Linguistics*, *41*(1), 75-116.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawaii Press.

Wu, X., Mauranen, A., & Lei, L. (2020). Syntactic complexity in English as a lingua franca academic writing. *Journal of English for Academic Purposes*, *43*, Article 100798.

Yin, S., Gao, Y., & Lu, X. (2021). Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. *System*, *97*, Article 102427.

Polarization in the Spanish press: A study of political sectarianism during the COVID-19 pandemic

Roberto A. Valdeón
South China Business College (SCBC), Guangzhou, China;
University of Oviedo, Spain; University of the Free State, South Africa
valdeon@uniovi.es

Abstract

Drawing on the concept of "political sectarianism" proposed by Finkel et al. and on Entman's classification of media biases, this paper aims to study polarization in three mainstream Spanish news websites, i.e. *El País*, *El Mundo* and *Abc*, during the COVID-19 pandemic through a cross-sectional analysis. For that purpose, a corpus of 900 articles was gathered during June 2020, following the end of the state of emergency and the strict lockdown enforced by the central government to fight the COVID-19 pandemic. The analysis relies on concepts used by communication and discourse analysts, i.e. framing and narratives. I have also drawn on the work of Bednarek on semantic choices. The study, which focuses on the headlines and the leads, shows that *El Mundo* and *Abc* were more critical of the central government's handling of the health crisis, while *El País* was critical of Madrid's conservative government. This points to the existence of distortion and content bias in the three newspapers. However, the most significant finding is the fact that *El País* is the only medium that clearly opted for a selection bias, omitting any news that could have been construed as critical of the central coalition government.

Keywords: Polarization, political sectarianism, media biases, *El Mundo, El País, Abc.*

Resumen

Partiendo del concepto de "sectarismo político" de Finkel et al, así como de la clasificación de los tipos de parcialidad en los medios de comunicación de Entman, este artículo lleva a cabo un estudio transversal de la polarización en tres medios españoles, i.e. *El País*, *El Mundo* and *Abc*, durante la pandemia de COVID-19. Para ello, se reunió

un corpus de 900 artículos publicados en el mes de junio de 2020, tras el fin del estadio de emergencia y el confinamiento estricto decretado por el gobierno central. El análisis se centra en conceptos utilizados en estudios de comunicación y análisis del discurso, tales como enmarcado y narrativa, así como el trabajo de Bednarek para el análisis de los elementos semánticos. El estudio, que se centra en los titulares y las entradas, revela que *El Mundo* y *Abc* fueron más críticos con el gobierno central mientras que *El País* fue especialmente crítico con el gobierno regional madrileño. Esto refleja la existencia de dos elementos de parcialidad: de distorsión y de contenido. Sin embargo, el resultado más importante es el hecho de que *El País* es el único que se caracteriza por parcialidad en la selección de la información al eliminar toda noticia que pudiera interpretarse como crítica con el gobierno central.

Palabras clave: Polarización, sectarismo político, parcialidad de los medios, *El Mundo, El País, Abc.*

1. Introduction

In countries such as the United States, the United Kingdom, Italy, Greece and Spain, politics has become a battle ground for antagonistic debates and confrontation. Spain's political system has been polarized since the economic recession of 2008 and the emergence of two competing forms of populism (Vampa, 2020): far-left *Podemos* (literally "We can") and, more recently, far-right Vox ("Voice" in Latin). *Podemos*, a party characterized by its many ideological contradictions (Caravantes, 2021), had stood out as the only far-left party to play a key role in Europe (Ramiro and Gómez, 2016), and was the only populist national party in Spain until the electoral success of far-right Vox in 2018 (Turnbull-Dugarte, 2019) and 2019 (Turnbull-Dugarte, Rama and Santana, 2020). The polarization resulting from (or reflected by) the appearance of these two parties in Spain differs from political polarization in the US, where the division between the Democrats and the Republicans remains the driving force of American politics. In contrast, Spanish politics has been radically transformed not only by incorporating extremist political parties in its parliament, but especially in its government. In December 2019, Spain's Socialist Party reached an agreement with *Podemos*, whose leaders have been linked to the authoritarian regimes of Latin America and Iran (Torre, 2017), to form the first coalition government since Spain's return to democracy.

Thus, events such as the polarisation of Spanish politics caused by the recession, the economic measures taken by the conservative government of the Popular Party after winning a landslide election following the failed economic policies of the Socialist Party (Martín and Urquizu-Sancho, 2012; Salmon, 2017), the emergence of the *indignados* movement (Perugorría and Tejerina, 2013) and the creation of the extreme-left party

Podemos first and far-right Vox later (Turnbull-Dugarte, 2019; Turnbull-Dugarte, Rama and Santana, 2020) as well as a number of other political problems such as the Catalan secessionist crisis have produced a media climate where news companies seem to be more interested in addressing their own partisan readership than in offering reliable information, which has led readerships to mistrust the information provided by news media, e. g. during the COVID-19 pandemic (Martín-Llaguno et al., 2022).

Ideological affinity between editors and political parties seems to be another crucial factor in the increased sectarianism of news media in Spain (for a discussion see Valdeón, 2022). Indeed, public institutions and private corporations have typically attempted to influence gatekeeping decisions (Champagne, 2005; Vos & Finneman, 2017) for a number of reasons, including the existence of ideological affinity between their leaders and the possibility to obtain financial gain and exert power. In Spain, for instance, the controversy over to the (in)efficiency of the measures taken by the central government and the allegations that Prime Minister Pedro Sánchez and deputy Prime Minister Pablo Iglesias took advantage of the health crisis to impose undemocratic measures was widely reported by news outlets (Martín-Llaguno et al., 2022), although the representation and interpretations of those measures have varied considerably.

To be sure, political polarization in Spain is reflected in both public and private media, including companies such as *Libertal Digital* and *Público*, which support *Vox* and *Podemos* respectively. More recently *Podemos* launched a party-controlled digital medium called *La Última Hora*, edited by controversial party member Dina Bousselham (see below). This article does not aim to study these media, but rather to analyze how political sectarianism has permeated through mainstream news outlets and how it is reflected in the on-line versions of three mainstream national newspapers: *Abc*, *El País* and *El Mundo*. They are the three major print newspapers whose internet circulation is also among the highest. *El Confidencial*, an online only medium, comes close to *Abc*, but I decided to select only those with both print and internet editions.

Founded in 1903, *Abc* has had a tumultuous history. It has traditionally supported the Spanish monarchy, it was banned during the Spanish republican period[1], and, after Spain's return to democracy, has been associated with the conservative Popular Party. In contrast, *El País*, founded in 1976, was the first left-leaning newspaper to appear after the death of dictator Francisco Franco. From its inception, the name of the newspaper was accompanied by the slogan "Diario independiente de la manaña" [Independent morning newspaper], recently replaced by "El periódico global" [The

1 See *Gaceta de Madrid* (22 October 1931) which published a law banning "la apología del régimen monárquico" [the defense of the monarchy] (p. 420).

global newspaper]. *El País* has been traditionally supportive of the Socialist Party. Finally, *El Mundo*, first published in October 1989, is a centre-right daily that has exposed a number of corruption scandals, both in the Popular and the Socialist Parties. *El Mundo* has been involved in a controversy over the allegations that the Basque terrorist organization ETA was linked to Madrid's 2004 Islamic terrorist attacks.

2. Theoretical framework

In connection with the tendency towards political polarization, Finkel et al. (2020) have proposed the term "political sectarianism" to describe the situation in the United States. Defined as "the tendency to adopt a moralized identification with one political group and against another" (2020: 533), Finkel and his colleagues identified three core ingredients of political sectarianism, namely othering, aversion and moralization. Othering refers to the "tendency to view opposing partisans as essentially different or alien to oneself", aversion is "the tendency to dislike and distrust opposing partisans" and moralization considers "opposing partisans as iniquitous" (2020: 533). Although Finkel et al. discuss these elements in relation to US party politics, they can also be related to news media as "people who are already sectarian selectively seek out congenial news, but consuming such content also amplifies their sectarianism" (2020: 534).

In fact, this article aims to analyze sectarianism in three mainstream news media in Spain. Since Finkel et al.'s proposal is at an initial stage, I would like to suggest that political sectarianism can be studied in connection with the notion of "bias", which Entman (2007) posited as being crucial in media production. Entman mentioned three different types of bias, namely distortion bias (referring to a process that distorts or falsifies reality), content bias (which favors one side rather than providing balanced treatment to two or more sides of a political conflict) and decision-making bias (which is related to the reasons why journalists select specific material). Entman used this concept to integrate three other notions widely used in communication studies: agenda-setting, framing and priming. In this article, I will draw on the tripartite concept of bias to explore political sectarianism in three mainstream Spanish media in relation to framing, which will be studied from the perspective of narrative theory.

The focus was on the headlines and the leads. Following Molek-Kozakowska (2013: 185), my assumption was that lexical choices and syntactic patterns might be crucial to present the topics, since headlines do not only include names and concepts with high 'news value' (Dor, 2003: 712-713) and 'connect' the story to previously known facts and events (2003: 714), they also *frame* the story in an appropriate fashion (2003: 715). In connection with this, Dor has noted that headlines normally try to answer:

What kind of story is this? Is this, for example, a politics-oriented story, a human interest story, an entertainment story? As everybody who has ever worked with journalistic materials knows, the answer to these questions does not lie in the objective world, but *in the construction of the story by its writer and editor* (my italics) (2003: 715).

This means that headline writers are responsible for the framing mechanisms, following the editorial line of the newspaper. In fact, and given the features of today's information society, readers very often have their first and only impressions from headlines (Jovanović, 2018: 5).

For this purpose, the concept of framing was used to complement those of polarization and bias. This concept has been widely applied in journalism and in other disciplines in the social sciences and the humanities. Framing is related to gatekeeping processes as journalists select some aspects of a perceived reality to make them more salient (Entman, 1993: 52). These processes influence the way in which audiences interpret events and, consequently, contribute to shaping their opinions (De Vreese, 2005). Entman (2007: 164) has refined the concept as "the process of culling a few elements of perceive reality and assembling a narrative that highlights connections among them to promote a particular interpretation".

As gatekeeping involves selecting news, on a macrolevel I will also draw on narrative theory. Based on Somers & Gibson's sociological approach (1994), discourse scholars (e.g. Baker, 2006) have studied the construction of public narratives in a variety of texts, including the news. Narratives are built by selecting a small number of events from the vast number available to writers or, as Baker puts it, "to elaborate a coherent narrative, it is inevitable that some elements of experience are excluded and others privileged" (2006: 71). Thus, media (and other institutions and individuals) resort to selective appropriation of topics to disseminate a narrative by repeatedly exposing audiences to them. The elaboration of these narratives results from or is related to framing mechanisms. In this paper, I will consider the selection and deselection of news events as well as the textual (linguistic and visual) mechanisms used to promote specific frames. This will serve to understand the features and functions of frames as the "organizing principles that are socially shared and persistent over time, that work symbolically to meaningfully structure the social world" (Reese, 2001: 11). Accordingly, framing is related to and can contribute to analyzing the tripartite notion of bias mentioned above.

On a microlevel, framing can be realized by various mechanisms, linguistic (e.g. headlines, leads) and non-linguistic (e.g. charts, photographs). The use of specific terms to refer to news actors or events are examples of labelling (Baker, 2006: 122) or framing. Reese mentions the use of terms such as the "war on terror" rather than the

"war on terrorism" and "prolife" rather than "prochoice" as examples of competing frames (Reese, 2007: 152). Also of relevance is the work by Bednarek and Caple (2019: 60-61), who have identified negativity and evaluative language as features capable of constructing negative representation of facts. Bednarek and Caple argue that negative vocabulary (such as confusion, damage, deaths) does not "automatically and expressly tell us that the writer disapproves of the reported events" (2019: 60) since "negativity" is one of core values of news production. However, they also point out that negative items can be used to express evaluation of various types (2019: 232), including likes, dislikes, preferences and anxieties (Bednarek, 2015: 47). Visual elements are not the focus of this article, but references to them will draw on the work by Kress and van Leeuwen (2006). Both Bednarek and Caple (2019) and Kress and van Leeuwen (2006) draw on Halliday's functional linguistics, which will also be used here.

3. Research questions and methodology

In line with the above and drawing on Finkel et al.'s notion of sectarianism (2020) and Entman's concept of bias, this paper aims to analyze the extent to which political polarization is reflected in *Abc*, *El País* and *El Mundo* in the wake of the lifting of the state of emergency declared by the Spanish central government to fight the spread of the COVID-19 pandemic. For that purpose, I will examine the news items published by the three newspapers during June 2020. The declaration of the state of emergency on March 14th was a highly controversial decision that followed the government's initial refusal to ban mass public gatherings despite repeated warnings and recommendations by international organizations (Martín-Llaguno et al., 2022). In May 2020, after the state of emergency was lifted, the country entered the period referred to as "nueva normalidad" [new normal], in which the strict lockdown was gradually eased. However, for the study it was decided that May was too close to the state of emergence while by June the polarization resulting from "denying and minimizing COVID-19" first (Martín-Llaguno et al., 2022: 412) and declaring the state of emergency later might have subsided (for a discussion of the polarization during the month of March see the valuable research by Martín-Llaguno et al., 2022).

The initial assumption was that *El País* would be critical of Madrid's regional conservative government and support the left/far left-leaning positions of the central government. Recent studies have in fact underscored the ideological affinity between the current leaders of the Socialist-led coalition government and the editorship of *El País*. In contrast, *Abc* and *El Mundo* were more likely to be more critical of the central government than of Madrid's conservative regional cabinet. To put these initial assumptions to the test, three research questions aimed to consider the topics that the newspapers focused on and the ways in which these media framed the health,

economic, social and political crises resulting from the pandemic. Thus, research questions focused on topics and events rather than on journalistic text types. More precisely the questions were phrased as follows:

R1. Did the three newspapers report on similar issues, even if from a different perspective, or did they focus on different news events or topics?

R2. How did the newspapers frame the news events or topics?

R3. What type of biases defined the three newspapers?

In order to answer these questions, a corpus of 900 articles[2] comprising the top ten articles published every day for the whole of June (300 per newspaper) was collected. The "top ten articles" label refers to those that occupy the most prominent positions in the webpage of the newspapers, that is, they were the first articles that their respective readerships would see before scrolling down their web pages.

After gathering the corpus, the main topic of every article was used for a preliminary topic-based classification. Although initially the classification comprised fifteen distinct topics and an additional miscellaneous section under the label "others", the number of articles for seven of those categories was negligeable, so they were finally moved to the "others" section (see table 1). Subsequently, a qualitative analysis of the texts was conducted drawing on the theoretical framework mentioned above. The analysis was divided into macro- and micro-levels.

4. Results and analysis

4.1 Macrolevel: selection and deselection of news events

As mentioned, and based on the main topic of every article, the final classification comprises eight categories, namely COVID-19, central government, education, economy, Madrid's government, 8M-related news texts, scandals in *Podemos*, nepotism/corruption allegations against the central government, as well as the miscellaneous section called "others". The following table summarizes the numbers and percentages of articles devoted to each of these topics by the three newspapers.

2 As the corpus contains copyrighted material, it cannot be published as supplementary material. However, it is available to anyone interested in validating the study and can be sent privately.

Table 1: Topic-based classification of articles

	COVID	Government	Education	Economy	Madrid	8M-related	Podemos-related scandals	Nepotism/ Corruption	Other topics
El País	83 (27,6%)	32 (10.6%)	12 (4%)	30 (10%)	17 (5.6%)	4 (1.3%)	1 (0.3%)	0 (0%)	121 (40.3%)
El Mundo	85 (28.3%)	21 (7%)	12 (4%)	44 (14.6%)	8 (2.6%)	15 (5%)	27 (9%)	5 (1.6%)	83 (27.6%)
Abc	109 (36.3%)	24 (8%)	4 (1.3%)	36 (12%)	10 (3.3%)	7 (2.3%)	17 (5.7%)	11 (3.7%)	82 (27.3%)

It should also be noted that there is bound to be thematic overlapping among the articles included in these categories, notably among those in the first six. All six are somehow related to the COVID-19 pandemic. However, to differentiate the articles thematically the most salient topic was selected to decide the specific category. For example, the article entitled "El forense del 'caso 8-M' señala a miembros del Gobierno pese a no estar investigados" [The forensic expert of the '8-M case' points to cabinet members even though they are not being investigated] reports on the investigation into those responsible for the spread of the virus during the so-called feminist demonstrations of March 8, 2020, but the central topic is the role that these gatherings may have played in the pandemic.

As regards the main findings, the table shows that the main topic in the three newspapers was the COVID-19 crisis. The information in these articles covers a variety of issues such as the number of people infected and deaths, articles about symptoms and recent research developments, the evolution of the virus outside Spain as well as in specific Spanish regions and so on. The second category broadly refers to the actions taken by the central government during the crisis. On the whole, the first four categories account for 52.2% of the articles in *El País*, 53.9% in *El Mundo* and 56.6% in *Abc*, which means that all three newspapers devoted over 50% of their texts to these issues.

This is interesting when compared with the number of articles under the miscellaneous category "Others", which is almost 13 points higher in the case of *El País* than *El Mundo* and *Abc*. Also of interest is the fact that, while some of the "other" articles in *El País* dealt with serious news events (such as the Black Lives Matter demonstrations in the US), a sizeable part provided information and/or comments about human-interest stories or celebrity news, such as the financial difficulties of Montreal-based Cirque du Soleil, the death of the Spanish novelist Carlos Ruiz Zafón, information about the problems in the Annapurna, soap operas, and news pertaining to actor Antonio Banderas and New York-based movie director Woody Allen as

well as one of *El País*'s pet topics, that is, the former dictator Francisco Franco. In contrast, the number of human-interest stories of this type in the other newspapers is considerably smaller. For example, *El Mundo* reported on the latest news about the disappearance of three-year old Madelaine McCann in Portugal and two other similar events, while *Abc* provided its readership with information about the death of a Spanish singer and a Spanish actress, about the way in which the pandemic affected the lives of prostitutes, and the difficulties Chinese citizens in Spain faced in the early months of the pandemic. The tendency to focus on human-interest stories has been referred to as a tabloidization process of the press (Winston, 2002) and can be related to a deselection process of COVID-19 related information in favor of other issues (Baker, 2006: 122; Entman, 2007) that, in the case of *El País*, may have served to divert the attention from the controversies affecting the central government to be discussed below. The selection and deselection of material, or selective appropriation in Baker's terms (2006), is the principal mechanism used by the three newspapers to provide their readerships with interpretive frameworks. This exemplifies at least two of Entman's biases (2007), namely content and decision-making.

Consequently, the overall impression is that the information about the pandemic in its many facets appears more diluted in *El País* than in the two newspapers. In fact, on specific days articles focusing on other topics outnumbered those about the pandemic, e.g. on June 30 three texts dealt with the political situation in Venezuela and three more were devoted Hong Kong. Conversely, on that same day COVID-19 related texts were reduced to three. On June 11th, 14th, 18th and 19th, *El País* also published a larger number of articles about other topics rather than about the pandemic. In contrast, both *El Mundo* and *Abc* focused on COVID-19 related topics during the whole month.

In addition, *El País* on the one hand and *El Mundo* and *Abc* on the other differed in the treatment of the topics in four other categories. In March 2020, as the pandemic was spreading across the European continent, calls were made to ban mass rallies and meetings in line with the World Health Organization's recommendations[3]. However, Spain's government allowed political, cultural and sports gatherings until March 8, the date when feminist demonstrations had been called across the country. These events, particularly Madrid's rally, were later blamed for the contagion in the capital by opposition parties, who claimed that the coalition government was more interested in showcasing ideology than in public health (Martín-Llaguno et al., 2022). The controversy also affected Madrid's conservative regional government, which in

3 On February 25, WHO reminded governments that the strict measures taken in China were being effective to stop the spread of the virus. On March 3, it warned of a likely shortage of medical supplies putting health workers in danger if effective measures were not taken. On March 5, it encouraged governments to take aggressive measures against the spread of the virus because "this is not a drill".

turn became the target of the central government's criticisms. This controversy was differently reflected in the three newspapers selected for this study: *El País* omitted all references to it, while *El Mundo* and *Abc* devoted front pages to this issue, particularly in the first week of June, when the Ministry for Equality acknowledged in an off-the-record comment that the relative failure of Madrid's feminist rally may have been due to the fear of becoming infected by the coronavirus. *El País*'s deselection of any references to the controversy demonstrates a decision-making bias that does not stem solely from individual journalists as partisan actors (Patterson and Donsbagh, 1996), but rather from the combination of the institutional and media forces at play (Entman, 2007: 167).

In fact, between June 1st and 9th *El Mundo* published eleven texts reporting or commenting on the allegations that the feminist demonstrations may have been the cause of the fast spread of the virus in the capital. On June 3, the main headline read: "8-M: Cónclave en la Fiscalía General al más alto nivel para fijar la posición sobre la manisfestación" [8-M: High-level meeting at the Attorney General's office to agree on an official position about the demonstration] followed by an op-ed column entitled "La bola de nieve del 8-M: antes matar a Montesquieu que rectificar" [The 8-M snowball: rather kill Montesquieu than rectify]. The texts were accompanied by a photograph of the demonstration where the Prime Minister's wife, occupying the central position, is surrounded by Pedro Sánchez's female ministers. The photograph did not only serve to illustrate the news event, in fact it dominated the website to present the essence of the news text, thus offering an interpretive framework (Kress and van Leeuwen, 2006: 30) which becomes the evidence of the newspaper's narrative (Bednarek & Caple, 2019: 118-119): government members are responsible for the spread of the pandemic as they allowed the demonstration to take place. As regards the linguistic components, the first article reported on a meeting at the Attorney General's Office in which the main news actors met to agree on an official position concerning the demonstration: here the word "conclave" [conclave] points to an act of faith rather than an act of justice. The accompanying opinion column by Juanma Lamet largely contributed to underscore *El Mundo*'s position: the government preferred to kill Montesquieu rather than to admit to their mistake. The metaphorical reference to the separation of the legislative, the executive and the judiciary served the newspaper to frame the government negatively: metaphors such as "conclave" and "kill Montesquieu" are part of what has been called the "frame package" (van Gorp, 2007: 64) and are also examples of what Baker described as labelling (2006: 122) as a procedure contributing to framing news events for the readers. Labelling also allows writers to present a specific position within the interpretative framework used by the writer and the newspapers, exemplifying the structures of anticipation discussed by Baker (2006: 106) as crucial in framing events.

As regards *Abc*, seven articles were devoted to the controversy surrounding the Minister of Equality's off-the-record admission that the feminist demonstration had been a failure, and to report about the Civil Guard's investigation into allegations that the government had been aware of the dangers of allowing mass public gatherings of any type. For example, on June 5th *Abc* published a report questioning the decisions taken by the government in light of the information available in the month prior to the demonstrations. The article was complemented with a video of those gatherings. Thus, the framing process encompassed linguistic and visual components, the value of the video resulting from the editing process, which depicted both political rallies and sports meetings, providing a clear frame of interpretation for the intended audience (Kress and van Leeuwen, 2006: 203). This also related to the notion of framing space that, drawing on Goffman, Baker discusses (2006: 109). In Baker's view the different interactions of the participants, including the news actors but also the journalists, can be presented in ways that will serve to allocate the frame space allocated to the writer and what is acceptable in each medium. Official information and visual elements allow in the case on *Abc* and *El Mundo* to provide information within their own space frame and interpretative framework.

On the other hand, *El País* did not publish any texts about this topic. This is a clear example of deselection of information, pointing to the fact that the editorial line of the newspaper might have sided with the government's position: *El País* prioritized the interest of "particular holders of power" (Entman, 2007: 166) over the right to information (Martín-Llaguno et al., 2022). By omitting information *El País* elaborated on patterns of omission and addition that are meant to suppress or accentuate certain narratives (Baker, 2006: 114). Paradoxically, while the allegations against the government's decisions and the ensuing investigation were conspicuously absent from *El País*, the newspaper published two reports with these headlines on June 9th: "La Abogacía acusa a la juez del 'caso 8-M' de lanzarse a 'una búsqueda voraz de indicios' contra el Gobierno" [The state lawyer accuses the judge in charge of the '8-M' case to embark on a rapacious search for evidence] and "El forense del 'caso 8-M' señala a miembros del Gobierno pese a no estar investigados" [The forensic expert in the '8-M case' points his finger at cabinet members even though they are not being investigated]. In these articles, the newspaper pointed its finger at the judge and the forensic team, presenting the information as a vendetta against the government. In other words, the newspaper deselected information about the investigation into the case, but selected information involving the critique of the investigation. Thus, a selection bias is combined with distortion biases in order to produce specific views of the news event (Entman, 2007: 163).

In contrast, on June 9th *El Mundo* and *Abc* opted for providing their readerships with the highlights of a 2,000-page report produced by the investigators, totally absent

from *El País*. On the 9th *Abc*'s main article was entitled "La fiscalía quiere cerrar el 8-M mientras el forense ve una 'hecatombe' predecible" [The General Attorney's office wants to close the 8-M investigation while the forensic expert believes the 'disaster' was predictable], which reported on the state lawyers's attempts to undermine the judge's mandate to carry out the investigation. *Abc* complemented the text with an editorial column denouncing the government's efforts to control the Prosecutor's Office following the appointment of an outspoken and controversial former Socialist Minister of Justice as the new General Prosecutor.

During the same period, *El País*'s gave salience to Madrid's conservative government's handling of the pandemic. For instance, on June 6th two articles reported on allegations that the regional government prioritized younger patients over the elderly. The newspaper claimed that Madrid's regional government might have instructed doctors to ignore older patients and those living in nursing homes. Surprisingly, *El País* reported that the far-right party *Vox* demanded an inquest into the matter. Otherwise, *Vox* does not normally receive much coverage in the newspaper. A third text with little information was a rhetorical exercise over the future of Madrid's premier Isabel Díaz Ayuso. On June 18th, *El País* published an interview with the central government's representative in Madrid's region. The headline included a direct quotation that read "José Manuel Franco: 'La causa del 8-M iba dirigida a erosionar al gobierno'" [José Manuel Franco: 'The 8-M case aimed to undermine the position of the government]. The question-answer format allowed the interviewee to criticize his political opponents unabatedly and fiercely, and to defend all of the central government's decisions. Direct quotations, and interviews tend to be more emotional than reported speech (Bednarek and Caple, 2019: 58), but also to favour personal views to the detriment of information. More importantly, the interview served to give voice to the central government.

In addition to these issues, one of the most popular topics during the first month of the so-called de-escalation process was of political nature, i.e. a series of allegations of corruption against *Podemos*. These allegations tainted the Deputy Prime Minister, who was accused of tampering with a SIM card belonging to a female member of the party, which allegedly contained confidential and damaging information. Party members dismissed the allegations as unethical attempts by conservative institutions such as the judiciary and mainstream media, which *Podemos* consistently accused of being run by the extreme right, to undermine his position within the government. In fact, *Podemos* has used several metaphors such as "la casta" [the caste] (Briziarelli, 2016; Sanders et al., 2017: 552) to stigmatize traditional politicians, and "las cloacas" [the cesspit of the State] to denigrate some institutions. Repeated exposure (Baker, 2006: 72-73) to these words has familiarized readers with their intended meaning.

The approach to these allegations were very differently dealt with by the three newspapers. The most significant finding was the fact that *El País* did not to publish any articles on this topic. Conversely, *Abc* and *El Mundo* provided their readers with comprehensive information about the scandal as well as several op-ed columns. This clearly points to decision-making biases whose aim was to prevent exposure to this news event in the former and to provide repeated exposure in the case of the latter (Baker, 2006: 72-73), thus framing *Podemos* and, consequently the government, very differently. *Abc* and *El Mundo* published a total of 17 and 26 articles about the topic respectively. In other words, during the period under examination 5.7% and 9% of the top ten articles in these newspapers were devoted to allegations of corruption in *Podemos*. These articles provided information about the evolution of the investigation and gave prominence to the contradictions within the party's publicized ethical stance. Both newspapers underscored the fact that *Podemos* had gone from being the whip of the Socialist and the Popular parties to unashamedly share in the privileges of the system. Interestingly, *Abc* also included articles about corruption in conservative parties (e.g. on the 20th it posted an article about a financial scandal in Foro, a spin-off of the conservative Popular Party in the northern region of Asturias).

Particularly noteworthy is the three-day period June 26-28. On the 26th *El Mundo* reported that it was in 2016 when *Podemos* first became aware of the damage that the missing SIM card could inflict on the party. A second article unveiled another political scandal: *Podemos*'s lawyer has been in contact with the State Prosecutor, who might have disclosed confidential information to benefit the party. On the same day, *Abc* reported on the judge's attempt to replace the prosecutor in charge of the case because of the latter's close connection with *Podemos*. This was complemented with an op-ed in which journalist Carlos Herrera appropriated *Podemos*'s own terminology in a text titled "La cloaca eres tú" [The cesspit is you], where Herrera described Pablo Iglesias as a corrupt leader by placing "cesspit" as the Theme of the headline and the informal "tú" to refer to Iglesias as the Rheme. This selection of elements in a typical Theme-Rheme structure has been described as crucial in framing news events and actors (Fowler, 1991), even though the Theme here is known rather than new information (Halliday, 2004: 93). However, placing "cesspit" presupposes the readers' familiarity with a word that Podemos has repeatedly used.

In stark contrast, during the same period, *El País* focused on the COVID-19 pandemic in the US, on the writer Carlos García Gual, and on whether the movie *Cinema Paradiso* had become a modern classic. The only article about corruption in *Podemos* during this period was posted on June 27. The text, entitled "La sombra de Villarejo persigue a Iglesias" [Villarejo's shadow threatens Iglesias], presented Pablo Iglesias as a victim, or the Goal (Halliday, 2004: 184). The article served to claim that

the opposition parties were taking advantage of the situation to attack the Deputy Prime Minister. Conversely, in *El Mundo* and *Abc*, Iglesias appears as the Actor (Halliday, 2004: 180) rather as the passive recipient of Villarejo's actions: "Iglesias carga contra el juez..." [Iglesias criticises the judge] "Iglesias burla al juez..." [Iglesias makes fun of the judge] respectively. Interestingly, the following day, *El País* deselected the topic, while *El Mundo* provided detailed information about how the whole news story had unraveled in the previous months, and *Abc* offered its readership new information about the latest developments in the case.

In addition to the articles about the SIM card and other scandals concerning *Podemos*, *El Mundo* and *Abc* also published various texts on cases of corruption and/or nepotism in the Socialist Party: seven and eleven texts respectively. Most of these texts reported on the appointments of friends and relatives of members of the central government to highly paid public jobs with no merits other than being related to the Prime Minister or his ministers. In total, corruption scandals in *Podemos* and reports of nepotism featured in 10.6% of *El Mundo*'s articles and 9.4% of *Abc*'s. *El País* only devoted one article to these issues, a negligeable 0.3% of the total. The selection and deselection of information in the three newspapers is closely related to the requirements that Harcup and O'Neill (2017: 1471) have identified for news stories. While *Abc* and *El Mundo* selected topics that concerned the power elite and were relevant for the audience, *El País* chose to avoid them. These decision biases reflect the agendas of the newspapers (Harcup and O'Neill, 2001: 278-279).

To sum up, on a macrolevel we have seen that biases affect all three newspapers, with a decision-making bias exclusive to *El País*'s decision not to report about the scandals of corruption and nepotism tainting the coalition government. All three exhibit content and distortion biases by positioning themselves in favor or against the central government and Madrid's regional government. The decision-making bias in the case of *El País* also contributes to diluting the magnitude of the pandemic as human-interest stories regarding writers, actors and so on are given prominence.

Let us now consider the articles at a microlevel, focusing on the lexical choices and the strategies to present the news in the three newspapers.

4.2 Microlevel: lexical and grammatical choices

As exemplified in the previous section, on a microlevel, framing mechanisms in *Abc* and *El Mundo*, on the one hand, and in *El País*, on the other, differ considerable. In this section, I will illustrate in more detail the centrality of lexical and grammatical choices as framing devices capable of accentuating the narratives of the three newspapers.

As discussed in the previous section, the three media are characterized by decision-making biases, albeit in different ways. Decision biases are complemented with framing mechanisms that provide larger interpretive frameworks. For example, *Abc*'s decision to publish an article on the creation of a government agency for Ignacio Carnicero, one of Sánchez´s closest friends, was presented with a headline that stressed political nepotism by means of factual vocabulary (June 4): "Sánchez crea un nueva dirección general en Transportes para colocar a su mejor amigo" [Sánchez creates a new national transport agency for his best friend], where Sánchez is the Actor and Carnicero the Recipient (Halliday, 2004: 190-191). In addition, the lead provided the readers with background information about the appointment: Carnicero and Sánchez played in the same basketball team in their youth, and Carnicero appeared in some of Sánchez's electoral videos. The articles were accompanied by an editorial entitled "Un escudo social para el amigo de Sánchez" and an opinion column titled "¡Gracias, Pedrín!" [Thanks, Pete!]. The former used the metaphoric phrase "escudo social" [social shield], widely used during the pandemic by the coalition government to defend the laws passed aimed at providing financial aid to the unemployed. The writer subverted it in a radical manner (Baker, 2006: 94) in order to criticize the creation of a highly paid position for the Prime Minister's friend at a time when thousands of people were losing their jobs. This shows the importance of the use of stock phrases and keywords in the creation and consolidation of clusters of facts and judgements (Entman, 1993: 55) and exemplifies the importance of headlines in the construction of interpretive frameworks for new stories (Dor, 2003: 715). In addition, the choice of nouns contributes to the evaluation of the news story. While it has been pointed out that evaluation is typically provided by means of adjectives, adverbs and modal verbs (Bednarek and Caple, 2019: 190-191), these headlines illustrate the use of nouns ("amigo") and noun phrases ("escudo social") for the same purpose. The use of the diminutive ("Pedrín") adds a sarcastic tone to the negative evaluation present in the noun phrases. On the whole, these choices relate to what Bednarek and Caple call the parameter of emotivity in news writing (2019: 178-179), which in turn is associated with the ideological stance of the news organization.

Similarly, the visual elements are of paramount importance, especially after the shift from the text-dominated news to the consolidation of the image as central in news production. In line with this, it has been pointed out that images function as symbolic representations of the news content and the way in which a news company might want to present it to its target audience (Kress and van Leeuwen, 2006: 105-108, Bednarek and Caple, 2019: 138-139). For instance, the article entitled "Sánchez crea un nueva dirección general en Transportes para colocar a su mejor amigo" was accompanied by a powerful visual element: a combined image of two photographs in which the main actors, Prime Minister Sánchez and his friend Carnicero, who

look remarkably similar, wearing informal clothes and impeccably trimmed hair. In line with Kress and van Leeuwen's grammar of visual content, this photograph is characterized by the carriers (that is, the two politicians), and the possessive attributes (that is, their physical attributes, which serve to project the critical interpretation) (Kress and van Leeuwen, 2019: 107). This image is meant to impact readers: as has been argued, the visual element in a news text is not merely a given, but rather it has immense political significance related to questions of power (Frosh, 2011: 95) and, in this case, the newspaper's denunciation of the abuse of power.

In addition, *Abc* used two other mechanisms to present the government in a negative way, namely opinion columns and the inclusion of direct quotations in the headlines. Let us consider a few examples:

González: "A veces el Gobierno se parece mucho al camarote de los hermanos Marx"

[González: Sometimes the government looks like the Marx Brothers's cabin] (11 June)

Ayuso: "Iglesias nunca llamó para ayudar, siempre que pudo ha echado gasolina"

[Ayuso: Iglesias never called to help, he has added fuel to the fire] (14 June)

"El gobierno nunca ha tenido un plan contra el Covid"

[The government has never had an anti-COVID plan] (21 June)

These headlines correspond to hard news in which a number of politicians provided critical views of the central government, including former Socialist leaders. The use of quotes in the headlines has several functions. Direct quotations may indicate that the journalist has not manipulated the source texts providing readers with the 'voice of experience'. However, quotes also allow news workers to include indirect evaluation by selecting the words of the news actors (Bednarek & Caple, 2019: 112) that would better contribute to the framing process. Bednarek and Caple point out that the way in which the quotes are integrated is also crucial: in these examples, quotes are given prominence as they are used as *the* headlines, rather than as part of the headlines. Thus, the text is constructed around what politicians say (Bell, 1991: 153). In this case, *Abc* quoted the conservative leaders of Madrid and Andalucía to point out that the central government lacked a plan against the pandemic, and that the Deputy Prime Minister never provided any assistance during the process. More importantly, *Abc* quoted former Socialist Prime Minister Felipe González, who is credited with the modernization of Spain during the 1980s and much of the 1990s, in order to undermine the measures taken by the coalition government. In this quotation, González compares the coalition government with the famous scene of *A Night at the Opera*, a 1935 movie by Sam Wood with the Marx brothers, to highlight that the cabinet ministers kept contracting each other.

As regards *El País*, during the period analyzed, its headlines showed a much greater preference for evaluative lexis and syntax than *Abc* and *El Mundo*. Let us consider a few examples:

> España inicia una nueva era: reduce sus tropas en Irak y las retira de Afganistán (1 June)
>
> [Spain begins a new era: it reduces the number of troops in Iraq and withdraws its troops from Afghanistan]
>
> Una reconocida abogada fichada por La Moncloa dirigirá la CNMV (1 June)
>
> [A renowned lawyer appointed by Moncloa will head the CNMV]
>
> Sánchez: "El enemigo es el virus y la política debe combatirlo unida" (3 June)
>
> [Sánchez: "The enemy is the virus and politicians should unite to fight it]
>
> Salvador Illa: "Visto lo visto todos llegamos tarde a esto" (7 June)
>
> [Salvador Illa: "Given what we now know, everyone was late for this]
>
> Sánchez a Casado: "¿Va a perseverar en la bronca o va a tomar el camino de la unidad?" (10 June)
>
> [Sánchez a Casado: "Will you continue to use confrontation or will you take the road to unity]

These examples point in two directions. First, as was the case of *Abc* and *El Mundo, El País* on quotations by senior officials to produce news texts. However, *El País* depended almost exclusively on government sources. This has been explained through the notion of bureaucratic affinity (Fishman, 1988) between governments and the media, and has also been linked to the need to provide a feeling of authenticity to the news story (Bednarek & Caple, 2019: 111). It also contributes to the legitimization of specific narratives by selecting words that are given credence as a result of "someone said it" (Halliday, 2004: 446), and which reflect the ideological affinity between a specific news company and a government. Bednarek and Caple (2019: 265) have shown that quoting premiers such as Barack Obama serves to construe the newsworthiness of a news event. But in the case of *El País*, the sources quoted, i.e. the Spanish Prime Minister and Health Minister Salvador Illa, were used to support the government's positions and to undermine the leader of the opposition.

As regards the other two headlines, they were used to legitimize the government by means of positive evaluative adjectives, highlighting the actions of the government in a positive manner: the government's decision to withdraw Spanish troops from Afghanistan point to a new era ("nueva era") and the new government appointee is described as a "renowned lawyer", reminiscent of the feelgood stories used by the popular

press (Conboy, 2002: 174). In other words, these attitudinal epithets (Halliday, 2004: 319) are used as discursive practices that allow writers to construe or accentuate specific narratives by framing participants in a negative or positive way (Baker, 2006: 122).

Other notable examples of the use of quotations to provide readers with specific narratives of the events are "Las instrucciones en un hospital en Madrid: 'Vamos a denegar la cama a los pacientes con más riesgo de morir'" [Instructions in a hospital in Madrid: 'We will not give a bed to patients with the highest risk of dying'] (June 18) and "El asesor sanitario de Ayuso: 'Los ancianos quedaron abandonados a su suerte'" [Ayuso's health advisor: 'The elderly were left to their own devices'] (June 23). These headlines, which introduced the main stories of the day, focused on quotations that apportion the blame to Madrid's government by having recourse to officials accepting the blame or giving controversial instructions. Named sources tend to be preferred in these contexts, as in the second case, but unnamed sources may be used when named ones are not available or to avoid accusations of defamation (Bednarek & Caple, 2019: 111). However, these headlines are characterized by forceful statements that show agency on the part of the authorities ("Vamos a denegar") and passivity on the patients ("ancianos quedaron abandonados") (van Leeuwen, 2008: 23), where agency is expressed via a causative (Halliday, 2004: 513).

5. Concluding discussion

In this article, I set out to study whether and how polarization was reflected in three major Spanish newspapers after the Spanish central government lifted the state of emergency enforced in the first months of the COVID-19 pandemic. I aimed to explore whether these media exemplified "the tendency to adopt a moralized identification with one political group and against another" (Finkel et al., 2020: 533) by means of three questions. As Finkel et al.'s discussion of political sectarianism is at an early stage, I have drawn on Entman's notion of bias and narrative theory with regards to framing the actors and actions of news events as well as on the Bednarek and Caple's framework for the analysis of news discourse (2019) and Kress and van Leeuwen's grammar of visual content (2006).

The study has clearly shown that the three media opted to report on specific topics (Entman, 1993: 52) for the benefit of their intended audiences: they reflected clear decision-making biases (Entman, 2007). Thus, specific events were prioritized to the detriment of others in order to present the political situation in a way that served to "dislike and distrust opposing partisans" (Finkel et al., 2020: 533). The study has shown that *El País* supported the Socialist-led coalition government, whereas *Abc* and *El Mundo* were critical of the political, economic and health measures put forward by the central government.

As regards the framing of the events and news actors, two aspects are of particular interest. First, while *Abc* opted for macrolevel mechanisms that favored the selection of specific news events such as the negative effect of the measures taken by the government, it retained features of objectivity on a microlevel as the lexical choices included fewer evaluative words in the headlines than in the case of *El País*. In other words, the interpretive frames were shaped by the reiteration of topics that highlighted the incompetence of the government rather than on lexical choices. *El País*, on the other hand, opted for the deselection of news topics that could impact the central government (e.g. corruption and nepotism scandals), and the headlines of its articles included a high number of negative verbs, nouns and adjectives that framed the opposition negatively, while those applied to the government were positive. In this sense, as mentioned, the results fall in line with the ideological position of the three media: *El Mundo* and *Abc* were critical of the central government, *El País* was more complacent with its policies.

Although the influence of biases on news production is not new (Entman, 2007: 165), the most conspicuous finding regarding the third question is the fact that *El País* displays a greater number of bias markers. On a macrolevel, this newspaper opted for the suppression of news events that might have undermined the central government's credibility. This means that information about the allegations of political corruption in *Podemos* were absent in the newspaper's daily most prominent newsfeed. Criticism of the government's handling of the health and economic crises was also minimal, while negative reporting of Madrid's conservative government occupied a central position. Thus, *El País* favored one political side over the other (Finkel et al., 2020) illustrating both content and decision-making biases (Entman, 2007), while, on a microlevel, its news writers selected negative words (notably adjectives, but also verbs and nouns) that reinforced the negative presentation of Madrid's government and the positive representation of the central government. This exemplifies the concept of "othering" introduced by Finkel et al (2020: 533).

These selection/deselection processes are particularly noteworthy, as they are not merely examples of the *news slant* that characterizes news reports and editorials and which favor "one side over the other in a current or potential dispute" (Entman, 2007: 165). In fact, news slant can be observed in all three newspapers. It is observable in *Abc*'s and *El Mundo*'s reporting of the allegations of corruption against the central government and in their presentation of the measures taken by the central government and Madrid's regional government. However, it is *El País* that better illustrates decision-making biases as this newspaper completely deselected any news related to allegations of corruption and nepotism, and focused on a critique of Madrid's conservative government, whose decisions were in permanent conflict with those of the central government. This takes us to Finkel et al.'s notion of sectarianism as a sign not only of

the political situation in countries such as the US and Spain, but also of the media's alarming loss of objectivity at a time when the governments of a number of European national governments are attempting to control (or have effectively controlled news media).

Starting from Finkel et al's preliminary proposal of the concept of "political sectarianism", this article has applied it to news articles in three major Spanish newspapers. It has combined concepts used in communication studies, such as bias and framing, with the analysis of language and multimodal elements from a discourse studies perspective. As Finkel et al's proposal requires elaboration and was primarily applied to American political, I have suggested a possible expansion that may, in the future, be refined or redesigned in order to study polarization in other media, in different languages and to carry out diachronic studies.

6. References

Baker, M. (2006). *Translation and conflict. A narrative account*. London: Routledge.

Bednarek, M. (2015). *Emotion talk across corpora*. Basingtoke: Palgrave MacMillan.

Bednarek, M. & Caple, H. (2019[2012]). *News discourse*. London: Bloomsbury.

Bell, A. (1991). *The language of news media*. Oxford: Blackwell.

Briziarelli, M. (2016). To 'feel' and to 'understand' political struggle: The national-popular rhetoric of Podemos. *Journal of Communication Inquiry*, *40*(3), 287-304.

Caravantes, P. (2021). Tensions between populist and feminist politics: The case of the Spanish left populist party Podemos. *International Political Science Review*, *42*(5), 596-612.

Champagne P. (2005). The 'double dependency': The journalistic field between politics and markets. In R. Benson & E. Neveu (Eds.), *Bourdieu and the journalistic field* (pp. 48-63). Malden, MA: Polity.

Conboy, M. (2002). *The new and popular press*. London: Sage.

De Vreese, C.H. (2005). News framing: Theory and typology. *Information Design Journal + Document Design*, *13*(1), 51–62.

Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of Pragmatics*, *35*(5), 695-721.

Entman, R.M. (1993). Framing toward clarification of a fractured paradigm. *Journal of Communication*, *10*(2), 155–173.

Entman, R.M. (2007). Framing bias: Media in the distribution of power. *Journal of Communication*, *57*, 163-173.

Finkel, E.J., Bail, C.E., Cikara, M. et al. (2020). Political sectarianism in America. *Science*, *370*, 533-536.

Fishman, M. (1988). *Manufacturing the news*. Austin: University of Texas Press.

Fowler, R. (1991). *Language in the news: Discourse and ideology in the press*. London: Routledge.

Frosh, P. (2011). Framing pictures, picturing frames: Visual metaphors in political communications research. *Journal of Communication Inquiry*, *35*(2), 91-114.

Halliday, M.A.K. (2004). *An introduction to functional grammar*. London: Hodder Arnold.

Harcup, T. & O'Neill, D. (2001). What is news? Galtung and Ruge revisited. *Journalism Studies*, *2*(2), 261–280.

Harcup, T. & O'Neill, D. (2017). What is news? News values revisited (again). *Journalism Studies*, *18*(2), 1470-1488.

Jovanović, S.M. (2018). Headlines as fake news: Discursive deception in Serbia's *Daily Informer* (2012–2018). *Central and Eastern European Review*, *12*, 1–22.

Kress, G. & van Leeuwen, T. (2006). *Reading images. The grammar of visual design*. London: Routledge.

Martín, I. & Urquizu-Sancho, I. (2012). The 2011 general election in Spain: The collapse of the Socialist Party. *South European Society & Politics*, *17*(2), 347-363.

Martín-Llaguno, M., Ballestar, M.T., Sainz, J. & Cuerdo-Mir, M. (2022). From ignorance to distrust: The public "discovery" of COVID-19 around International Women's Day in Spain. *International Journal of Communication, 16*, 409-436.

Molek-Kozakowska, K. (2013). Towards a pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication*, *7*(2), 173–197.

Patterson, T. E. & Donsbagh, W. (1996). News decisions: Journalists as partisan actors. *Political Communication*, *13*(4), 455-468.

Perugorría, I. & Tejerina, B. (2013). Politics of the encounter: Cognition, emotions, and networks in the Spanish 15M. *Current Sociology*, *61*(4), 424-442.

Ramiro, L. & Gómez, R. (2016). Radical-left populism during the great recession: Podemos and its competition with the established radical left. *Political Studies*, *65*(1), 108-126.

Reese, S.S. (2001). Framing public life: A bridging model for media research. In S.S. Reese, O. Gandy & A. Grant (Eds.), *Framing Public Life* (pp. 7-31). Mahwah, NJ: Erlbaum.

Reese, S.S. (2007). The Framing project: a bridging model for media research revisited. *Journal of Communication*, *57*(1), 148-154.

Salmon, K. (2017). A decade of lost growth: Economic policy in Spain through the great recession. *South European Society and Politics*, *22*(2), 239-260.

Sanders, K., Molina Hurtado, M.J., & Zoragastua, J. (2017). Populism and exclusionary narratives: The 'Other' on Podemos's 2014 European Union election campaign. *European Journal of Communication*, *32*(6), 552-557.

Somers, M. R., & Gibson, G. D. (1994). Reclaiming the epistemological "Other": Narrative and the social constitution of identity. In C. Calhoun (Ed.), *Social theory and the politics of identity* (pp. 37-99). Oxford: Blackwell.

Torre, C. (2017). Hugo Chávez and the diffusion of Bolivarianism. *Democratization*, *24*(7), 1271-1288.

Turnbull-Dugarte, S.J. (2019). Explaining the end of Spanish exceptionalism and electoral support for Vox. *Research & Politics*, *6*(2), 1-8.

Turnbull-Dugarte, S.J., Rama, J. & Santana, A. (2020). The Baskerville's Dog suddenly started barking: Voting for Vox in the 2019 Spanish general elections. *Political Research Exchange*, *2*(1), 1-21.

Valdeón, R.A. (2022). Gatekeeping, ideological affinity and journalistic translation. *Journalism*, *23*(1), 117-133.

Vampa, D. (2020). Competing forms of populism and territorial politics: The cases of Vox and Podemos in Spain. *Journal of Contemporary European Studies*, *20*(3), 304-321.

van Leeuwen, T. (2008). *Discourse and Practice*. Oxford: Oxford University Press.

van Gorp, B. (2007). The constructionist approach to framing. *Journal of Communication*, *57*(1), 60-78.

Vos, T. P. & Finneman, T. (2017). The early historical construction of journalism's gatekeeping role. *Journalism*, *18*(3), 265-280.

Winston, B. (2002). Towards tabloidization? Glasgow revisited, 1975-2001. *Journalism Studies*, *3*(1), 5-20.

Contents:

An analysis of language anxiety in English and Basque-Medium Instruction: A study with primary school students
Irene Alonso Arruabarrena and David Lasagabaster Herrarte 9

Collaborative writing and patterns of interaction in young learners: the interplay between pair dynamics and pairing method in LRE production
María Basterrechea and Francisco Gallardo-del-Puerto 49

Metaphors set in motion in the context of L2 academic spoken discourse
Irene Castellano-Risco, Marta Martín-Gilete, Alberto Hijazo-Gascón and Iraide Ibarretxe-Antuñano 77

Computational analysis of adjuncts in ASD-STE100 for the NLP parser ARTEMIS
Francisco José Cortés-Rodríguez and Carolina Rodríguez-Juárez 107

How complex is professional academic writing?
A corpus-based analysis of research articles in 'hard' and 'soft' disciplines
Javier Pérez-Guerra and Elizaveta A. Smirnova 149

Polarization in the Spanish press: A study of political sectarianism during the COVID-19 pandemic
Roberto A. Valdeón 185

Universidade Vigo