ojs.uv.es/index.php/qfilologia/index

Rebut: 21.07.2022. **Acceptat**: 12.10.2022

Per a citar aquest article: Bellés-Calvera, Lucía & Caro Quintana, Rocío. 2022. "Is academic discourse accurate when supported by machine translation?". *Quaderns de Filologia: Estudis Lingüístics* XXVII: 171-201.

doi: 10.7203/QF.27.24671

Is academic discourse accurate when supported by machine translation?

¿Es el discurso académico preciso cuando se apoya en la traducción automática?

Lucía Bellés-Calvera
Universitat Jaume I
lucia.belles@uji.es
ORCID: https://orcid.org/0000-0002-1329-6395

Rocío Caro Quintana
University of Wolverhamtpon
R.Caro@wlv.ac.uk
ORCID: https://orcid.org/0000-0003-2275-2679

Resumen: El discurso académico ha despertado interés entre investigadores y profesores (Deroey, 2015; Mauranen, 2012; Hyland, 2010), en particular el uso de marcadores metadiscursivos. Sin embargo, se ha prestado poca atención a estas características apoyadas por la traducción automática (TA) en los contextos de AICLE. El objetivo del presente artículo es describir el uso y la frecuencia de los enfatizadores y atenuadores empleados en los ámbitos de la historia y la psicología y analizar la precisión de los equivalentes obtenidos en dos plataformas de TA, en concreto, DeepL y Google Translate. Para ello, se ha elaborado un pequeño corpus de dos seminarios y se han aplicado métodos cualitativos y cuantitativos para determinar la frecuencia y la precisión de los recursos lingüísticos bajo estudio. Los resultados han revelado que, si bien los elementos interaccionales proporcionados por la TA son precisos, pueden producirse omisiones y errores de traducción. Estas conclusiones pueden ser relevantes para los profesores de AICLE interesados en el discurso académico, así como para los investigadores de traducción que trabajan con corpus bilingües y multilingües y evalúan la exactitud de las herramientas de traducción.
Palabras clave: Aprendizaje integrado de contenido y lengua (AICLE); metadiscurso; enfatizadores; atenuadores; traducción automática neuronal.

Abstract: Classroom discourse has aroused interest among scholars and educators (Deroey, 2015; Mauranen, 2012; Hyland, 2010), particularly the use of metadiscoursal markers. However, little attention has been paid to these features when they are supported by machine translation (MT) engines in content and language integrated learning (CLIL) contexts. The aim of this paper is to describe the use and frequency of hedges and boosters employed in the fields of

History and Heritage and Psychology and analyse the accuracy of the equivalents obtained from two MT engines, namely DeepL and Google Translate. To this end, a small corpus consisting of two seminars was compiled and qualitative and quantitative methods were implemented to determine the frequency and the accuracy of the linguistic structures under study. The results revealed that even though the interactional devices provided by MT engines are highly accurate, some omissions and mistranslations may occur. These findings may be valuable for CLIL lecturers interested in classroom discourse, as well as for translation researchers working with bilingual and multilingual corpora who seek to assess the accuracy of translation tools.
Keywords: Content and language integrated learning (CLIL); metadiscourse; boosters; hedges; neural machine translation (NMT).

1. Introduction

In recent years, the promotion of multilingual education has become a priority in response to global challenges. Due to the value of multilingualism as a sustainable resource in the economic and social dimensions (Stavans & Hoffman, 2015; Cenoz, 2013; Edwards, 2004), international institutions such as the European Union have devised language learning policies with the aim of providing future generations with better career prospects (European Commission, 2020). In following these guidelines, European countries have strengthened linguistic diversity through the 2+1 target, that is, the teaching and learning of two foreign languages at various academic levels (Jover, Fleta & González, 2016; European Commission, 2015; Lasagabaster, 2012). For instance, the creation of the European Education Area and internationalization strategies developed by higher education institutions have supported the modernization of language teaching and learning and thus the development of multilingual classrooms (European Commission, 2015).

The role of English as a lingua franca in the academic world cannot be denied. The incorporation of this target language within educational programmes worldwide has been regarded as Englishization (Engwall, 2016; Hultgren, 2014). Not only has this phenomenon had an impact on classroom discourse, but also on international journals, conferences, and other academic events (Curry & Lillis, 2018; Hamel, 2013) where English works as a vehicular language. Major trends in this field have triggered the implementation of pedagogical approaches in the classroom setting (Graddol, 1999), such as content and language integrated learning (CLIL), content-based instruction (CBI) and English as a medium of instruction (EMI) (Merino & Lasagabaster, 2018).

The study presented here explores the use of metadiscourse devices in undergraduate programmes following a CLIL approach. These pedagogical initiatives seek to enhance the acquisition of disciplinary content and communicative skills through an additional language (Coyle, Hood & Marsh, 2010). On that basis, CLIL instruction can be conducted in any foreign, second or minority language, but the use of English as the vehicular language has become commonplace (Lanvers & Hultgren, 2018; Lasagabaster, Doiz & Sierra, 2014). Past research on CLIL has largely focused on the development of productive and receptive skills in primary and secondary education (Llinares, Morton & Whittaker, 2012; Ruiz de Zarobe, Sierra & Gallardo, 2011). Other issues such as teacher training courses, language learning attitudes and perceptions as well as the integration of open educational resources (OERs) have also become commonplace (Sylvén, 2013; Garone, Van de Craen & Struyven, 2020).

More recently, the focus has been on the use of metadiscoursal devices, either in spoken or written academic texts (Bellés-Fortuño, 2018; Molino, 2018; Ädel, 2018; Broggini & Murphy, 2017). Charles (2013), for example, noted that the analysis of academic discourse has delved into written productions, research articles, textbooks, and conference presentations. As regards spoken genres, it is lecture discourse that has received most attention since it is the most common teaching practice (Flowerdew, 1994; Swales, 1990). Studies on classroom discourse have contributed to identifying the types and frequency of interpersonal metadiscourse in educational settings (Bogdanović & Mirović, 2018; Zare & Tavakoli, 2016; Deroey & Taverniers, 2012; Morell, 2004), calling attention to teachers' speech while lecturing (Hyland, 2010; Bellés-Fortuño & Querol Julián, 2010). At this point, the use of digital resources should be highlighted. Tools such as slideshows, online videos and other resources support foreign language instruction through a set of verbal and non-verbal strategies which are decisive in building shared knowledge and getting students involved in the teaching-learning process (D'Angelo, 2018; Breuer & Archer, 2016).

However, research into the discourse of academic seminars in higher education settings is scarce. In fact, little is known about the role of machine translation engines in multilingual learning environments. This is the reason why the study presented here explores CLIL discourse in higher education, specifically in the domains of History and Heritage and Psychology. Drawing on Hyland's taxonomy of metadiscourse (2005), this paper examines the accuracy of two machine translation engines to provide translations of the

interactional markers used in CLIL seminars which are characterized by the use of multilingual patterns in Spanish, Catalan and English. Two CLIL seminars delivered as part of undergraduate programmes were audio-recorded and transcribed for the purpose of this study. In so doing, a small corpus containing boosters and hedges in the three languages was compiled to allow the identification and subsequent analysis of metadiscoursal devices when aided by machine translation systems in order to raise lecturers' and students' awareness towards these elements and their use in the classroom.

The analysis of instructional genres other than lectures could be of great value for scholars, policymakers, educators, and translators. The findings of this study could raise awareness of strategies that help lecturers accommodate their discourse for CLIL learners. It may also shed light on the role of machine translation engines as pedagogical and research tools in the field of corpus linguistics, especially when working with bilingual and multilingual corpora.

2. Literature review

The study of communication in the content classroom has enhanced the implementation of effective instructional models (Llinares, Morton & Whittaker, 2012). The structure and comprehension of academic texts is determined by the metadiscoursal devices educators and learners employ in their oral and written performances (Hyland, 2017). In other words, it is through the appropriate use of metadiscoursal markers that content assimilation and exposure to comprehensible input take place (Kuteeva & Mauranen, 2018). Classroom discussions and meaning-making processes are common practices in educational settings, but it is in seminar sessions where teacher-student interaction is most common (Soter, Wilkinson, Murphy, Rudge, Reninger & Edwards, 2008).

Seminars have been regarded as instructional practices that allow to discuss disciplinary topics in the classroom setting. Thus, rich opinions that call for reflection and feedback provision are shared (Harry, Gordon & Schmandt, 2012; Curzon, 2003). According to Barefoot and Fidler (1992) there are different types of seminar sessions, each of which have their own specific internal structure, namely extended-orientation seminars, academic seminars with generally uniform academic content across sections, academic seminars on various topics, pre-professional or discipline-linked seminars, basic study

skills seminars and seminars that are hybrid. Extended-orientation seminars are intended to introduce new students into academic life so that they get to know the campus and learn time management and planning strategies. Secondly, academic seminars with generally uniform academic content across sections are based on interdisciplinary or theme-oriented topics that are required within curricula not only to develop disciplinary knowledge, but also to improve students' academic skills. Thirdly, academic seminars on various topics are similar to the second category, but sections may differ in terms of content and topic. Fourthly, pre-professional or discipline-linked seminars are regarded as academic events devoted to preparing future professionals. Finally, basic study skills seminars are addressed to students who need more academic training, whereas hybrid seminars contain features of at least two of the previously mentioned categories.

The study of the devices used in these academic texts is known as metadiscourse analysis (Hyland, 2005). This subfield of discourse analysis, referred to as metatalk or metatext, serves as the basis to consider written and spoken contributions in terms of organization, evaluation, and interaction (Hyland, 2005). At the tertiary education level, scholars have mostly focused on corpora of written academic texts (e. g. university lecture transcripts, research articles), analysing the three metafunctions of language, which are ideational, interpersonal, and textual metadiscourse (Morell, 2020; Molino, 2017; Lee & Subtirelu, 2015; Farrokhi & Ashrafi, 2009). Nonetheless, further research is still needed that examines teaching genres that are meant to be more practical and interactive.

Despite the importance of metadiscourse in academic events, little is known about the accuracy of translated corpora when supported by machine translation tools. Focusing on this field, machine translation (MT) has been placed within the sub-field of computational linguistics, the practice of studying the use of computers when translating written and spoken texts from one language into another without human intervention (Hutchins, 1995). In MT the whole translation process is automatically generated by the computer, although a human translator is responsible for the post-editing process as the translation output is still not at the human level (Läubli et al., 2018). This concept implies that intervention by a post-editor who revises the texts provided by the MT engine is needed (Allen, 2003; Hutchins, 1995; O'Brien, 2022).

Throughout the years, several MT theories and approaches have emerged, including rule-based machine translation (RBMT), or statistical machine translation (SMT) and neural machine translation (NMT). Today the latter

is considered the state-of-the-art (Popovic & Castilho, 2019; Chu & Wang, 2020). NMT "considers translation as a task involving operations on numbers performed by mathematical systems called *artificial neural networks*" (Pérez-Ortiz, Forcada & Sánchez-Martínez, 2022: 141). In the last few years, some scholars have claimed that NMT is able to perform at the level of human translation in very particular domains and languages, focusing, for example, on news translated from Chinese or Czech into English (Hassan et al., 2018; Bojar et al., 2018; Popel, 2018). These findings, however, have been questioned in Läubli et al. (2020) since some aspects of the evaluation employed could be improved, particularly the use of non-expert annotators. Additional studies have also revealed that most annotators show a preference for human translation (Läubli et al., 2020). As regards other specialized areas, there is a lack of research on the implementation of NMT and metadiscoursal devices in the fields of History and Heritage and Psychology, which are the core of the current study.

The above notwithstanding, the use of MT engines has become commonplace within the translation industry, studies having demonstrated NMT to be a commercial solution when dealing with large-scale translation production (Shterionov et al., 2017). MT has also gained interest in the academic field, where translation students have benefited from university level MT courses, including undergraduate and postgraduate degrees dealing with post-editing issues (Çetiner & İşisağ, 2019, Stasimioti & Sosoni, 2019). This is specifically linked to the area of NMT, which has been incorporated into most MT engines, such as Google Translate and DeepL Translator (Rescigno et al., 2020). Published literature on Google Translate has mainly been concerned with the usefulness of this tool in academic contexts on the basis that it may help students learn or translate into a second or foreign language (Van Rensburg, Snyman & Lotz, 2012; Groves & Mundt, 2015). In a study conducted by Groves and Mundt (2015), poor quality texts were obtained when translating from Malay and Chinese into English through Google Translate. These findings, however, did not intend to prevent translators and students from using Google Translate, as the aim of the study was to show students that MT is still far from perfect and can be used in certain scenarios, for instance, to translate a website with the intention of only understanding.

The need for effective classroom practices has led to the examination of both the metadiscoursal features used in CLIL settings and the accuracy of the equivalents obtained with NMT. Previous research has been concerned with the translation of metadiscoursal elements in written texts rather than

in spoken texts. Williams (2010) investigated how translation students make use of metadiscoursal markers in the target language, examining three case studies where the students had to translate three texts from French into English. The findings suggested that translation using MT devices is problematic since many students failed to convey in the translated version the meaning of elements such as transition markers, attitude markers, engagement markers and hedges in the source. Farghal & Kalakh (2019) conducted detailed research into the translation from English to Arabic of metadiscoursal markers in a political speech. Other recent research has also examined the output of MT with certain groups of lexical items; for example, Popovic & Castilho (2019) look at the translation of ambiguous conjunctions with MT, whereas Müller, Rios, Voita & Senrich (2018) investigate the translation of pronouns from German into English.

In light of the above, the study presented here can help identify the extent to which MT engines are effective in providing metadiscoursal equivalents to items encountered in multilingual learning environments.

3. Research questions

This paper seeks to explore the use and frequency of the interactional features (Hyland, 2005) employed in CLIL discourse where multilingual patterns are observed and to determine the accuracy of the equivalents provided by MT engines. Thus, the research questions that will be addressed are as follows:

RQ(1): What types of boosters and hedges are most common in English, Spanish and Catalan?

RQ(2): Do they differ in type and frequency in the fields of History and Heritage and Psychology?

RQ(3): Do MT engines provide accurate and effective equivalents of metadiscoursal markers?

The recording of the academic discourse in two seminars provided the transcriptions used in the corpus and therefore facilitated the analysis of metadiscoursal devices used in spoken discourse. Pedagogically speaking, metadiscoursal devices play a significant role in learning practices since hypotheses and examples are essential to meet the linguistic and cognitive needs of CLIL participants. In this regard, this paper aims at investigating how meaning is conveyed in the target language and how translators must

pay attention to those linguistic markers in order to not jeopardize their original meaning.

The methodological approach taken in the study is provided in the section that follows. Elements such as the context, the participants as well as the analytical framework employed to analyse the results obtained are explained in detail.

4. Methodology

Qualitative and quantitative techniques were employed for the analysis of the corpus. A detailed description of items such as the context, the number of participants, the procedure and the research instruments employed has been provided as part of the methodological approach.

4.1 *The corpus*

The corpus under study was gathered at a Spanish university located in a bilingual region. According to the university's Strategy for the Coexistence and Promotion of Languages, known as the ECOPOL Plan, instruction in Spanish, Catalan and English is guaranteed in all degrees (Universitat Jaume I, 2019). As such, multilingualism in this higher education institution is the norm, linguistic flexibility being promoted through CLIL instruction. Following the guidelines established in this document, 5 % of the content of all degrees is provided in English. The coordination of each degree is responsible for the distribution of this 5 % using teaching materials or the organization of academic events in the different subjects. Thus, Spanish, Catalan, and English can coexist in the content classroom. These measures seek to foster participation in the learning environment, particularly in seminar sessions. CLIL discourse has thus been examined in this work in two fourth-year seminars delivered in the degrees of History and Heritage and Psychology.

Two group discussions, at the same Spanish university, that followed a CLIL approach were audiotaped, and a digital voice recorder was employed to transcribe the recordings. In other words, the written output produced contains features of spoken corpora, including fillers such as *ehh*, widely used in Spanish-speaking contexts to give speakers time to think. The transcription of the pedagogical practice from the History and Heritage seminar, which was 113 minutes long, amounts to 1,311 words. On the other hand, the Psychology

seminar contains 7,589 words and discourse lasted for 115 minutes. Hence, even though these samples are similar in terms of duration, notable differences can be observed with regard to the number of words per session (see table 1). The main reason for this is related to the classroom dynamics, given that in History, several YouTube videos about museum collecting were played.

Field	*Module*	*Number of words*	*Duration*
History and Heritage	Fundamentals of Artistic Heritage	1,311	113 min.
Psychology	Affective and Social Neuroscience	7,859	115 min.
Total		9,270	228min

Table 1. The corpus

As shown in table 1, the seminar sessions for analysis belong to the areas of soft and hard sciences. The genre observed in these modules is that of group discussions and similar outcomes can be found in terms of duration. Given that this work is part of a larger study, the following transcription codes were employed to identify speakers and preserve their anonymity (see table 2). The History and Heritage seminar was delivered by a content lecturer, being referred to as H2. As to the Psychology seminar, there were two lecturers responsible for instruction, thereby having P1 and P2 as the first and second moderators respectively.

The way of writing the text is based on transcription criteria for spoken texts in the well-known guidelines included in the MICASE corpus (Simpson et al., 2002). Short pauses were stressed with commas, whereas full stops indicated long pauses.

Transcription codes	
P1	CLIL teacher in the field of Psychology
P2	PhD student and CLIL teacher in the field of Psychology
H2	CLIL teacher in the field of History
S1/S2...	Students' numbers according to their contributions
[...]	Omitted text

Table 2. Transcription codes

4.1.1 *Participants*

The subjects of this study were three CLIL instructors (P1, P2 and H2) and sixty-seven fourth-year undergraduate students undertaking elective modules offered as part of the syllabus of two disciplinary fields, more specifically History and Heritage and Psychology (see table 3). The anonymity of the participants was guaranteed before they signed informed consent. Most participants were geographically related to the Valencian Community, a Spanish region where instruction in Spanish, Catalan and English is promoted at all educational stages. Hence, pedagogical practices are conducted in multilingual learning environments.

Participants	*Fundamentals and Theory of Artistic Heritage*			
	Year	*M*	*F*	*Age*
Teacher(s)	4th	-	1	30
Students		15	13	21-25
	Affective and Social Neuroscience			
	Year	*M*	*F*	*Age*
Teacher(s)	4th	-	2	24-50
Students		11	28	21-28

Table 3. Participants

Fundamentals and Theory of Artistic Heritage is an elective module that fourth-year undergraduate students can take during the first semester at the Spanish university under study. Delivered by a female lecturer with an intermediate level of English, this seminar involved twenty-eight students interested in the area of art collection, museums and exhibitions. The group was made up of fifteen males and thirteen females in their early 20s who were proficient in Spanish and Catalan. Their command of the English language, however, was not homogeneous, since basic and independent users were identified by means of a placement test. To meet the multilingual policy devised by the university, English was used as the means of communication in a CLIL seminar together with Spanish and Catalan. Within this session, group discussions took place after watching a series of YouTube videos dealing with art collection. These discussions were guided and controlled by the teacher through several questions to help students get the main ideas. Items listed in

these questionnaires included: *a) What kind of collections can you find at the British Museum?* and *b) What is the objective of the Department of the Scientific Research at the Metropolitan Museum of Art?*

Turning to the Psychology seminar, participants numbered two CLIL lecturers and thirty-nine learners. The module of Affective and Social Neuroscience is held during the second semester of the academic year and, as an elective course, students have the opportunity to further explore the contents of this subfield in English, which is the means of communication in science. The seminar was conducted by a CLIL lecturer (P1), who was in charge of the learning process through controlled practice activities which involved playing and discussing a number of YouTube videos dealing with disciplinary content. This laboratory learning environment was supported by a PhD student who teaches some theoretical and practical components of the module in Spanish (P2). The group consisted of eleven males and twenty-eight females with ages ranging from twenty-four to twenty-eight years old who were mostly from this bilingual region.

4.1.2 *Analytical approach*

The analytical framework acting as the basis for this study relates to Hyland's interpersonal model of metadiscourse (2005). Within this approach, metadiscoursal devices are classified into interactive and interactional categories so as to organize and evaluate discourse as well as to interact with the audience (Thompson, 2001). Bearing in mind that the focus of this paper is on interactional metadiscourse, boosters and hedges have been examined in this study (table 4).

The counting of samples was conducted manually to determine their frequency in each field. Moreover, MT engines were used to assess the translations of metadiscoursal devices in terms of accuracy. The findings are presented and discussed later.

Category	*Subcategory*	*Examples*
Hedges	Noun	argument, possibility, view, idea
	Verbs	seem, suggest, expect, anticipate
	Modal verbs	may, would, can, might
	Adverbs	potentially, often, normally, about
	Adjectives	probable, common, potential, any
	Phrases	in general, at least, to some extent
Boosters	Noun	evidence, fact, assertion, conclusion
	Verb	show, determine, emphasize, reveal
	Adjective	significant, clear, vast, generally
	Adverb	constantly, largely, indeed, entirely
	Phrases	in fact, for the most part, of course

Table 4. Hyland's interactional elements (2005: 49): hedges and boosters

4.2 *Research instruments*

Two free and well-known multilingual online systems were employed: Google Translate and DeepL. Both services support the translation of selected texts and documents created with Microsoft Office programmes into multiple languages. Google Translate works with the following file formats: .docx, .pdf, .pptx, and .xlsx., and DeepL processes PDF (.pdf), Word (.docx) and PowerPoint (.pptx) files.

Google Translate was launched by Google in 2006 using SMT engines, these later being replaced by NMT in 2016 (Google Translate, n. d.). This software allows for the translation of texts available in different forms and sources, such as short excerpts, websites, text files and even images. Not only does Google Translate provide users with these functions, but it can also translate and transcribe speech immediately for many languages.

DeepL Translator is an NMT software developed in 2016 with the aim of producing high-quality translated texts from and into more than 20, mainly European, languages, but also languages such as Japanese and Chinese (DeepL Translator, n. d.). As with Google Translate, the translation of text files is also available.

One difference between the systems is related to the number of languages employed for the translation of written texts. The languages used for the study are English, Spanish, and Catalan, the two formers being employed as

pivot languages (i. e. languages that may be intermediary sources to facilitate translation between two or more languages) in NMT systems (Cheng, 2019; Dabre et al., 2021). Both Google Translate and DeepL Translator software have adopted NMT but due to their inner paradigms, it is not known whether English and Spanish are still employed as pivot languages. Catalan, on the other hand, seems to be a lower resource language, and translation from and into Catalan is only provided by Google Translate (Ko et al., 2021). Other differences involve the translation of texts within images as well as text-to-speech and audio recording functions, which are not found in DeepL Translator.

5. Results and discussion

The results obtained in the small corpus used here are discussed below. The type of metadiscoursal marker, the frequency of boosters and hedges as well as the accuracy of their equivalents when supported by NMT engines were analysed and considered for the purpose of this study.

It appeared from the analysis of the data that the frequency of interactional markers is directly related to language and disciplinary domain (see figure 1). From all the hedges identified, a high number of them are encountered in English (63 %) and Spanish (30 %), whereas examples in Catalan are almost non-existent (7 %). As for the total number of boosters, the predominance of this type of discourse marker in Spanish has been observed in nearly 67 % of the cases. Their occurrence, however, is noticeably lower in English and Catalan, as only 31 % and Catalan 1 % of these markers could be identified. From these results, it can be discerned that English and Spanish are by large the most frequent languages when using hedges and boosters.

What is also evident is that the incidence of these devices greatly differs between History and Psychology seminars. As seen in table 5, 179 out of the 194 hedges identified came from the Psychology discourse, while History discourse accounted for only 15. Disciplinary differences were also observed in boosters, with Psychology containing 82 items and History the remaining 7. These findings reveal that the number of hedges exceeded boosters in both areas of expertise. This trend for a higher use of hedging devices over boosters has also been confirmed in recent research on academic articles (Livingstone, 2019; Takimoto, 2015).

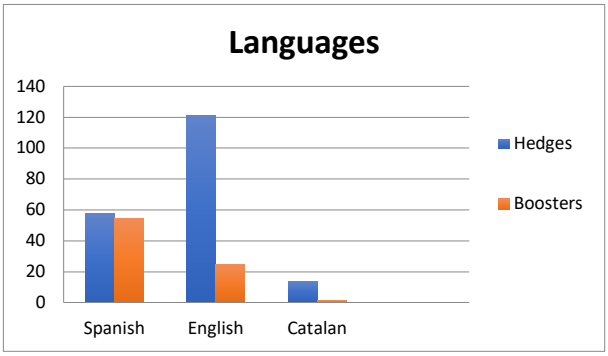

Figure 1. Languages: Boosters and hedges

Field	*Hedges*				*Boosters*			
	English	*Spanish*	*Catalan*	*Total*	*English*	*Spanish*	*Catalan*	*Total*
History	1	13	1	15	0	7	0	7
Psychology	120	46	13	179	25	48	2	82
Total	121	58	14	194	25	55	2	89

Table 5. Results: Interactional markers

5.1 *Hedges*

5.1.1 *History and Heritage vs. Psychology*

In contrast to what has been said above, it can be seen that the use of hedges in the English language is higher than in boosters, where Spanish was the predominant language. Table 6 shows that the most common hedging devices in Psychology involve the verb *to think*, followed by the modal verbs *can* and *would*.

Hedges: Psychology					
English	*Frequency*	*Spanish*	*Frequency*	*Catalan*	*Frequency*
argument	1	algo	6	algo	2
implication	1	algún/a	4	alguna/alguns	5
expect	1	algunos/as	3	no tindre clar	2
idea	10	hasta qué punto	1	poder+inf	
think	32	enterarse	1	a lo millor	2
point to	2	tener claro	2	típic	1
feel	4	normalmente	1		1
want	9	poder + inf.	5		
may	4	posible	2		
might	8	a lo mejor	2		
would	11	parecer	8		
can	25	desde mi punto de vista	2		
cannot/ can't	9	en general	1		
could	1	creo/creemos que	5		
should	2	pensamos	1		
		al menos	2		
TOTAL	**120**		**46**		**13**

Table 6. Results: Hedges in the Psychology seminar.

Examples (1) and (2) below relate to the lecturer's expressed uncertainty and/or doubt about the content being presented to encourage the students to get involved in the CLIL seminar and thus develop their thinking skills.

1. P1: What do you <u>think</u> is Neurotechnology? And what do you expect in this class according to the title? I mean, if we are talking about Neurotechnology, what do you <u>think</u>? And you can answer in the language you feel more comfortable [...] This is just like a quick brainstorming.

2. P1: I would like you to, because you point somehow these questions, to think about the advantages and disadvantages or pitfalls of all this sort of things, OK? I think there might be since I said at the beginning really nice, cool, surprising, fancy...But they are not useful. But they might be. Or they might be really unuseful. Yeah?

Concerning hedges in Spanish, tentative cognition and likelihood can be observed in the use of verbs such as *parecer* (to seem), *creer* (to think) and *poder* (can). With respect to nouns, *algo* (something) was the most frequently used hedging device. These categories are related to the mental processes of the speakers as well as to the feasibility of the ideas shared in the discussion, especially when they are used as part of personal constructions as evidenced in (3):

3. S17: Pues nosotros pensamos...pues, igual que ellas. El vídeo que más interesante nos ha parecido desde el punto de vista, claro, clínico son las gafas, por eso. Tienen, por una parte, mejoras. El componente médico de decir no pueden ver, y por otro la autoestima que ellos van a sentir. Capacidad de poder controlar ellos su vida, pues ahí ganan mucho a nivel psicológico [...].

As for Catalan, the most recurrent hedges in the disciplinary field of Psychology are *alguna* (some), *no tindre clar* (not be clear) and *poder* (can). All of them are used to express modality as well as the speaker's commitment to the proposition and its reliability. It is worth mentioning that the use of *algo* (something) and *a lo millor* (maybe) seems to be a highly negative transfer from Spanish into Catalan as they are not grammatically correct.

Hedges: History					
English	*Frequency*	*Spanish*	*Frequency*	*Catalan*	*Frequency*
any	1	algún/a, algunos	4	haver de	1
		cualquier	4		
		creo que	1		
		poder	4		
TOTAL	**1**		**13**		**1**

Table 7. Results: Hedges in the History and Heritage seminar

Turning to the field of History, table 7 above shows that the occurrence of hedges is low in English and Catalan, in fact only being used once in each (*any* in English and *haver de* (have to) in Catalan. Despite this, 13 instances were detected in Spanish, with *poder* (can), *algún/algunos* (some) and *cualquier/a* (any) being used 4 times. Evidence can be found in example (4) below.

4. H2: [...] lo que he hecho ha sido seleccionar una serie de vídeos ehh de extraídos de las páginas oficiales que cuelgan museos tipo el British, el Metropolitan, o el Getty en sus páginas oficiales de YouTube y están dedicadas a, cada uno de ellos se dedica a un aspecto. Pues algunos de ellos, pues por ejemplo, cómo se crea ese museo, cómo se construye por ejemplo el edificio del museo. Acordaros que hay algunos de ellos que se crean en edificios históricos, otros que se ubican en edificios de nueva planta [...]

These results suggest that the modal verb *can* as well as the determiner *any* are commonplace in both domains when using English. However, the presence of hedges in Psychology outnumbers their equivalents in History in Spanish and Catalan. Whereas in these languages they mainly serve as reliability markers, in English they seem to be audience-oriented.

5.1.2 *Translation*

As mentioned earlier, the corpora were translated with the aid of the MT engines Google Translate and DeepL Translator. Both engines provided correct translations to the hedges identified approximately 95 % of the time. However, these tools were characterized by frequent MT errors on the grounds that the texts translated were not written but produced in oral communication, including incorrect use of capital letters or wrong gender, among others.

It should also be noted that neither Google Translate nor DeepL Translator were able to translate the multi-word expression *tener claro* (to understand, to be clear about), although the problems for each engine are different. As seen in (6), the translation provided by Google Translate fails to convey the original meaning, which should be translated as "I don't really understand it". Google Translate's version could be interpreted as the speaker being the one who is not expressing their thoughts correctly, which is the opposite of the meaning conveyed in the source text. This error can thus be considered a mistranslation. In (7), DeepL's version is a literal translation that does not express the same meaning as the original.

5. Source text. S10: [...] Porque recuerdo que cosas así, es que no lo tengo muy claro, cosas así se podían hacer o se podían mejorar prótesis y tal, pero se le había perdido porque el cerebro tenía ahí ya como que esas zonas preparadas, lo que pasa que lo había perdido. [...]
6. Translation 1: S10: [...] Because I remember things like that, it's that I'm not very clear about it, things like that could be done or could be improved prostheses and such, but it had been lost because the brain had there already as if those areas were prepared, what happened there lost. [...] (Google Translate)
7. Translation 2: S10: [...] Because I remember that things like that, I don't have it very clear, things like that could be done or prosthesis could be improved and so on, but it had been lost because the brain already had there as those prepared areas, what happens is that it had lost it. [...] (DeepL Translator)

Nevertheless, there are cases where the MTs provide very good solutions. For example, the Spanish term *claro* was used once as an adjective (clear) and five times as an adverb (of course). Google Translate and DeepL identified these word functions and provided accurate translations.

Other mistranslations where MT engines do not convey the meaning of the metadiscoursal markers can be observed in hedging verbs like *enterarse* (to understand, to find out). Proper equivalents were not obtained given that *to find out* in (9) and *to know* in (10) do not convey the original meaning of the source text.

8. Source text. S7: El, el vídeo, bueno el objeto que nos ha parecido más interesante era el de las gafas que no nos hemos enterado muy bien, tenemos que averiguar mejor cómo funciona porque es una rayada y también estábamos hablando de que, de que... [...]
9. Translation 1. S7: He, the video, well the object that we found most interesting was the one with the glasses that we have not found out very well, we have to find out better how it works because it is a striped one and we were also talking about what, what ... [...] (Google Translate)
10. Translation 2. S7: The, the video, well the object that we found most interesting was the one with the glasses that we didn't really know about, we have to find out better how it works because it's a scratch and we were also talking about what, what [...] (DeepL Translator)

Other issues that should be considered are related to omission. As can be seen in (12), the modal verb *poder* (can) is omitted, but the meaning is main-

tained in the translated output, which seems not to be necessary in Spanish. As the output obtained from Google Translate was accurate, it is not included in the examples below.

11. Source text. P1: So basically you will have around 5 minutes so then you *can* discuss with ehh your group ah about which is the device or the, the creation, the invent, you better like or the ones because it may be a few of them, OK? [...]
12. Translation 2: P1: Así que básicamente tendrás unos 5 minutos para (omission) discutir con tu grupo sobre cuál es el dispositivo o la creación, el invento que más te gusta o los que más te gustan porque pueden ser unos pocos, ¿vale? [...] (DeepL Translator)

5.2 *Boosters*

5.2.1 *History and Heritage vs. Psychology*

As regards boosters, these markers are not as used as frequently as hedges. In table 8 below, their distribution in the three languages can be seen in the Psychology seminar. The most common boosters in English are the adverbs *always*, *especially* and *of course*, and the verb *to know*. The equivalent of this verb is also present in Catalan, although only 1 case was found. The language where boosters are used most is Spanish, the most frequently being *muy* (very), *claro* (clear) and *poco* (little).

In the following examples, equivalent boosting strategies are employed by students in both English and Catalan to signal their certainty in the truth of their personal contributions. This also occurs in Spanish when using adverbs that serve as emphasizers.

13. S7: I don't know how the representation of the objects ehh can represent on your brain.
14. S13: Ho dic perquè lo de la pròtesis n'hi ha un moment que sí que li fan en les dos cames i li estan fent una prótesis, ara, en este moment. Entonces no sé si serà lo mateix o és algo que ja n'hi havia.
15. S7: El, el vídeo, bueno el objeto que nos ha parecido más interesante era el de las gafas que no nos hemos enterado muy bien, tenemos que averiguar mejor cómo funciona [...].

Boosters: Psychology					
English	*Frequency*	*Spanish*	*Frequency*	*Catalan*	*Frequency*
of course	5	en parte	1	clar	1
prove	1	siempre	2	saber	1
especially	4	claramente	1		
always	5	obviamente	1		
know	10	realmente	5		
		muy	11		
		claro	6		
		mucho	4		
		tan/tanto	5		
		hecho	2		
		mostrar	1		
		destacar	1		
		sobre todo	1		
		poco/poquito/poquillo	7		
TOTAL	**25**		**48**		**2**

Table 8. Results: Boosters in the Psychology seminar

In the History seminar, the presence of boosters is limited to the Spanish language. A total of 7 devices are listed in table 9, with *poquito* (a little) and *mucho* (a lot) having the highest frequency and *siempre* (always) and *mostrar* (to show) having just one use each.

BOOSTERS: HISTORY					
English	*Frequency*	*Spanish*	*Frequency*	*Catalan*	*Frequency*
-	-	siempre	1	-	-
		mucho	2		
		mostrar	1		
		poquito	3		
TOTAL			7		

Table 9: Results: Boosters in the History seminar.

In the discourse, *poquito* referred to the idea of *a small extent*, as in (16) below. Conversely, *mucho* was used to add force to a statement, particularly to nouns that are omitted, which is a common practice in Spanish. As seen in (17), *mucho* emphasizes the amount of information available on museum collecting, whereas in (18) the student refers to the difficulty of the task.

16. H2: [...] También hay algún vídeo que muestra un poquito cómo se va organizando y elaborando una exposición ehh temporal.
17. H2: [...] A ver, este año la actividad en inglés que hemos planteado para que, ya os digo, para que os vayáis familiarizando con el vocabulario específico y demás...ehh va en relación al último de los temas que hemos visto. Sobre el coleccionismo hay mucho. ¿Qué estuvimos hablando la semana pasada? ¿Os acordáis? ¿No?, ¿nada?
18. S11: Pero esto nos costará mucho escribir todo esto, ¿no?

Drawing on the results obtained in this study, it can be seen that there is a substantial difference between the field of Humanities and Health Sciences. Seminar sessions are extremely practical genres where interaction is key. Given that the CLIL approach was promoted within course syllabi, it is no surprise that the use of multilingual patterns was apparent in the classroom setting. As evidenced, both the use and frequency of boosters in Spanish was higher than expected given that in both disciplines, English was meant to be the vehicular language. Disciplinary differences were also observed in the incidence of hedging devices in that in Psychology, the number of examples in English outnumbered those in Spanish and Catalan, while in History, only Spanish was employed to persuade and engage the audience. These findings suggest that the higher presence of linguistic items in languages other than English is determined by speakers' command of the target language.

5.2.2 *Translation*

As with hedging devices, there is a high percentage of accuracy in the translations of the boosters employed obtained through Google Translate and DeepL Translator (97 %), for all three languages. Certainly, the translation of boosters from Catalan into Spanish and English did not pose linguistic challenges for the MT engine. Turning to text inconsistencies, MT engines, as with hedges, omitted some boosters that were in the source text, as in (19), where *un poquito* (a little) was missing in the translation provided by DeepL Transla-

tor (21). While the main meaning of the source utterance is maintained, the nuance of *un poquito* is lost. Such omissions thus negatively affect the meaning conveyed by the speaker.

19. Source text. H2: [...] También hay algún vídeo que muestra un poquito cómo se va organizando y elaborando una exposición ehh temporal.
20. Translation 1. H2: [...] There is also a video that shows a little how a temporary exhibition is being organized and elaborated ehh (Google Translate)
21. Translation 2. H2: [...] There is also a video that shows how a temporary exhibition is organized and created (DeepL Translator).

Mistranslations were also encountered in boosting devices. In the original text (22), *hace poco* means *a while ago*. In (23) DeepL *Translator*, only *a few years ago* was provided as an equivalent even though it is not in the source text. The problem in this excerpt is therefore not only a mistranslation, but a repetition.

22. Source text. S4: A mí por ejemplo ehh más de, de, desde mi punto de vista, también del pasado que el vídeo era presente y futuro, pero pasado porque lo de los implantes cocleares, por ejemplo, hace poco, bueno, hace unos cuantos años era inviable y ahora se están viendo incluso en vídeos, ves como niños que ven a sus padres y los escuchan por primera vez [...]
23. Translation 2: S4: To me for example ehh more of, from my point of view, also from the past that the video was present and future, but past because the cochlear implant thing, for example, a few years ago, well, a few years ago it was unviable and now they are even being seen on videos, you see like children seeing their parents and hearing them for the first time [...] (DeepL Translator)

5.3 *Other relevant issues*

Other elements that were not the focus of study also caught our attention. As previously mentioned, DeepL does not offer translation from and into Catalan but, interestingly, when a text in Catalan is typed or copied in the source language space, it is identified as Spanish and translated into English. The output is not perfect, but it is relatively accurate.

24. Source text. S11: Jo tinc una pregunta. Això pots d'alguna manera, quan li fiques una pròtesi a algú, connectar o algo pa que puga menejar-lo o...?

Saps lo que vull dir? Que si va connectat al cervell d'alguna manera pa que se puga menejar la extremitat que has perdut o que no tens?

25. Translation 1 (Cat-Sp). S11: Yo tengo una pregunta. ¿Esto puedes, de alguna manera, cuando le metes una prótesis a alguien, conectar o algo pan que pueda menear-o ...? Sabes lo que quiero decir? Que si va conectado al cerebro de alguna manera pan que se pueda menear la extremidad que has perdido o que no tienes? (Google Translate)

26. Translation 2 (Cat-Eng): S11: I have a question. Can you somehow, when you put a prosthesis on someone, connect or something bread that can handle it or...? Do you know what I mean? What if it connected to your brain in some way that could lead to a limb that you lost or didn't have? (Google Translate)

27. Translation 3 (Sp-Eng): S11: *Jo tinc* a question. So, how do you find a way to prosthetize someone, to connect or something to manage or...? Do you know what I'm going to say? That if it is connected to the neck in some way to manage the extreme that you have lost or that you do not tense? (DeepL Translator)

As this is a corpus based on spoken CLIL discourse, expressions that are usual in informal language can be identified. This is the case of the term *pa* in Catalan, which not only means *bread* but is also the shortened form of *per a* (to) as in (24). Google Translate fails to recognize this trait from spoken discourse and provides us with Spanish (26) and English (27) equivalents that are not suitable for this context. Even though DeepL Translator recognizes this speech as characteristic and translates it correctly, (27), the whole excerpt is not as accurate, taking into account that the first sentence mixes Catalan and English, ***Jo tinc** a question*.

As academic discourse was retrieved from multilingual learning environments following a CLIL approach, code-switching is a common practice. Consequently, some utterances are articulated in English and others in Spanish. When adding these utterances to DeepL, one of the languages is identified automatically and no change between languages within a text is allowed.

28. Source text. P1: Bueno, so what would you highlight? <Eng> ¿Qué destacaríais del vídeo? <Sp> What thing you most it was more surprising for you or brought your attention? < Eng> ¿Qué os ha llamado más la atención? < Sp> What do you think? < Eng>.

29. Translation 2. P1: Bueno, ¿qué destacarías? <Eng> ¿Qué destacaríais del vídeo? <Sp> ¿Qué es lo que más le sorprendió o llamó su atención? <Eng>

> ¿Qué os ha llamado más la atención? <Sp> ¿Qué piensas? <Eng> (DeepL Translator).

In the example above, English is recognized by the MT engine as the main language of the source text; therefore, the excerpt is translated into Spanish. The intention was to check whether the booster *destacar* (to highlight) would be translated correctly into English. Notwithstading, as the MT engine recognised English as the source language, the translation for "¿Qué destacaríais del vídeo?" (What aspects would you highlight from the video?) was not provided.

6. Conclusion

This paper has delved into the use of interactional metadiscourse markers within 2 CLIL seminars delivered in the degrees of History and Heritage and Psychology. More specifically, the purpose of this study was to ascertain the number of hedges and boosters employed in multilingual contexts as well as the accuracy of NMT in three languages, namely English, Spanish and Catalan.

Focusing on RQ(1): What types of boosters and hedges are most common in English, Spanish and Catalan?, the highest number of hedges occurs in English, followed by Spanish, then Catalan. The verbs *think* and *can* are the most common hedges used in the target language, as opposed to in Spanish, where *parecer* (to seem) is the hedging verb that appears most often. In Catalan, however, *alguna* (some) seems to be the preferred hedging device. Participants, however, showed a preference for using Spanish when employing boosters, with *muy* (very) and *claro* (of course) being the most recurrent terms. The frequency of boosters in English and Catalan was much lower, the verb *to know* being common in both languages.

As for RQ(2): Do they differ in type and frequency in the fields of History and Heritage and Psychology?, the results revealed that the occurrence of boosters and hedges was much higher in the Psychology seminar, with *muy* (very), *to know*, *saber* (to know), *parecer* (to seem), and *alguna* (some) being the most frequent devices. In contrast, there was a low incidence of these metadiscoursal markers in the History seminar. The hedges encountered here were mostly in Spanish, with the same number of instances for *poder* (can), *cualquier* (any) and *algún* (some), whereas *any* and *haver de* (have to) were the only resources noted in English and Catalan respectively. Boosters,

on the other hand, were only identified in Spanish in the History seminar, for example, *poquito* (a little).

When it comes to RQ(3): Do MT engines provide accurate and effective equivalents of metadiscoursal markers?, it seems that both engines are fairly accurate, since over 90 % of the annotation of the texts were marked as correct. As discussed throughout this article, the results from Google Translate were more accurate than those of DeepL. The errors found in Google Translate are usually related to the mistranslation of multi-word expressions, but DeepL also fails to convey the meaning of some terms.

As for the limitations of this study, the small corpus size should be noted, which was in part the result of the limited number of lecturers who agreed to their seminars being audio-recorded. This research did not focus on all the CLIL modules delivered within a given degree, but on an interdisciplinary comparison in the fields of Humanities and Health Sciences. As such, a larger corpus collecting data from other programmes delivered within those disciplines would provide a more complete picture of CLIL metadiscoursal markers and the accuracy of MT engines when dealing with multilingual corpora. The limitations of the MT output should also be acknowledged. This study has examined texts taken from spoken corpora. To this end, the transcriptions used for MT contain many strongly oral markers, which caused some problems for the two MT systems used. In addition, triangulating the data with the proficiency levels of the participants would have increased the generalizability of the results.

Pedagogically speaking, the findings presented in this paper will be helpful for linguists, translators and CLIL educators working in multilingual areas. The production of devices such as boosters and hedges in the target language is key to engaging participants in the teaching-learning process, especially when they are used to attract and persuade the audience. In this sense, teachers and students can benefit from NMT in that they can learn and compare linguistic structures when exchanging knowledge in different areas of expertise. On the other hand, translators have the opportunity to review and improve the accuracy of NMT systems working with English, Spanish and Catalan. Further research in oral discourse should explore the use of hedges and boosters in a wide range of bi-/multilingual educational practices conducted in other areas of expertise (e. g. social sciences and technological sciences). Other studies could also consider the integration of NMT engines in the teaching-learning process.

References

Ädel, Annelie. 2010. Just to give you kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English. *Nordic Journal of English Studies* 9(2): 69-97.

Allen, Jeffrey. 2003. Post-editing. *Benjamins Translation Library 35*: 297-318.

Barefoot, Betsy O., & Fidler, Paul P. 1992. *1991 National Survey of Freshman Seminar Programming (Monograph No. 10)*. Columbia, SC: University of South Carolina, National Resource Center for The Freshman Year Experience.

Bellés-Fortuño, Begoña. 2018. Evaluative language in medical discourse: A contrastive study between English and Spanish university lectures. *Languages in Contrast* 18(2): 155-174.

Bellés-Fortuño, Begoña & Querol-Julián, Mercedes. 2010. Evaluation in research article abstracts: A cross-cultural study between Spanish and English medical discourse. In Lorés-Sanz, Rosa; Mur-Dueñas, Pilar, & Lafuente-Millán, Enrique (eds.) *Constructing Interpersonality. Multiple Perspectives on Written Academic Genres*. Newcastle Upon Tyne: Cambridge Scholars Publishing, 83-98.

Bogdanovic, Vesna & Mirovic, Ivana. 2018. Young Researchers Writing in ESL and the Use of Metadiscourse: Learning the Ropes. *Educational Sciences: Theory and Practice* 18(4), 813-830.

Bojar, Ondrej; Federmann, Christian; Fishel, M.; Graham, Yvette; Haddow, Barry; Huck, Matthias, Koehn, Philipp & Monz, Christof. 2018. Findings of the 2018 Conference on Machine Translation (WMT18). In *Proceedings of the Third Conference on Machine Translation (WMT), Volume 2: Shared Task Papers*). Bruxelles: Association for Computational Linguistics, 272-307.

Breuer, Esther & Archer, Arlene. 2016. *Multimodality in Higher Education*. Leiden: Brill.

Broggini, S. & Murphy, A. C. 2017. Metadiscourse in EMI lectures: Reflections on a small corpus of spoken academic discourse. *L'Analisi Linguistica e Letteraria XXV* (2): 75-91.

Cenoz, Jasone. 2013. Defining Multilingualism. *Annual Review of Applied Linguistics* 33: 3-18. doi: 10.1017/S026719051300007X

Çetiner, Caner & İşisağ, Korkut Uluç. 2019. Undergraduate Level Translation Students' Attitudes towards Machine Translation Post-Editing Training. *International Journal of Languages' Education and Teaching* 7(1): 110-120.

Charles, Maggie. 2013. English for academic purposes. *The handbook of English for specific purposes* 1: 137-153.

Cheng, Yong. (2019). Joint Training for Pivot-Based Neural Machine Translation. In Cheng, Y. (ed.) *Joint Training for Neural Machine Translation. Springer Theses.* Singapore: Springer. https://doi.org/10.1007/978-981-32-9748-7_4

Chu, Chenhui & Wang, Rui. 2020. A Survey of Domain Adaptation for Machine Translation. *Journal of Information Processing* 28: 413-426.

Coyle, Do; Hood, Phillip, & Marsh, David. 2010. *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.

Curry, Mary Jane & Lillis, Theresa. 2018. *Global Academic Publishing. Policies, perspectives, and pedagogies*. Clevedon, UK: Multilingual Matters.

Curzon, Leslie Basil. 2003. *Teaching in further education: An outline of principles and practice* (6th ed.). London / New York: Continuum.

Dabre, Raj, Imankulova, Aizhan, Kaneko, Masahiro, & Chakrabarty, Abhisek. 2021. Simultaneous multi-pivot neural machine translation. *arXiv preprint arXiv:2104.07410*.

Dalton-Puffer, Christiane. 2011. Content-and-language integrated learning: From practice to principles? *Annual Review of applied linguistics* 31: 182-204.

DeepL. n. d. *DeepL*. https://www.deepl.com/translator

Deroey, Katrien L. B. 2015. Marking importance in lectures: Interactive and textual orientation. *Applied Linguistics* 36(1): 51-72.

Deroey, Katrien L. B. & Taverniers, Miriam. 2012. Just remember this: Lexicogrammatical relevance markers in lectures. *English for Specific Purposes* 31(4): 221-233.

D'Angelo, Larissa. 2018. PowerPoint presentations in the classroom: Re-evaluating the genre. *Language Value* 10(1): 29-46.

Edwards, John. 2004. Bilingualism: Contexts, constraints, and identities. *Journal of Language and Social Psychology* 23(1): 135-141.

Engwall, Lars. 2016. The internationalisation of higher Education. *European Review* 24(2): 221-231.

European Commission. 2020. Education and Training: About multilingualism policy. https://ec.europa.eu/education/policies/multilingualism/about-multilingualism-policy_en [Access 20/07/2020]

European Commission. 2015. Language teaching and learning in multilingual classrooms. *Education and Training*. https://ec.europa.eu/assets/eac/languages/library/studies/multilingual-classroom_en.pdf [Access 12/04/2020]

Farghal, Mohammed & Kalakh, Bushra. 2019. Engagement in Translation: Interactional Metadiscourse Markers in American Presidential Debates. *Jordan Journal of Modern Languages and Literatures* 12(1): 103-22.

Farrokhi, Farahmand & Ashrafi, Somayeh. 2009. Textual metadiscourse resources in research articles. *Journal of English Language Teaching and Learning* 212: 39-75.

Flowerdew, John. 1994. Research of relevance to second language lecture comprehension: An overview. *Academic listening: Research perspectives*: 7-29.

Gallardo del Puerto, Francisco, Gómez-Lacabex, Esther, & García-Lecumberri, María Luisa. 2009. Testing the effectiveness of content and language integrated learning in foreign language contexts the assessment of English pronunciation. In Ruiz de Zarobe, Yolanda & Jiménez Catalán, Rosa María (eds.) *Content and Language Integrated Learning: Evidence from research in Europe*. Bristol, UK: Multilingual Matters, 63-80.

Garone, Anja, Van de Craen, Piet, & Struyven, Katrien. 2020. Multilingual nursing education: Nursing students' and teachers' interests, perceptions and expectations. *Nurse Education Today* 86, 104311: 1-6.

Google Translate. n. d. *Google Translate*. https://translate.google.com/

Graddol, David. 1999. The decline of the native speaker. *The AILA Review* 13: 57-66.
Groves, Michael & Mundt, Klauss. 2015. Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes* 37, 112-121.
Hamel, Rainer Enrique. 2013. El campo de las ciencias y la educación superior entre el monopolio del inglés y el plurilingüismo: Elementos para una política del lenguaje en América Latina. *Trabalhos em Linguística Aplicada* 52(2): 321-384.
Hassan, Hany; Aue, Anthony; Chen, Chang; Chowdhary, Vishal; Clark, Jonathan; Federmann, Christian; Huang, Xuedong; Junczys-Dowmunt, Marcin; Lewis, William; Li, Mu; Liu, Shujie; Liu, Tie-Yan; Luo, Renqian; Menezes, Arul; Qin, Tao; Seide, Frank; Tan, Xu; Tian, Fei; Lijun, Wu; Wu, Shuangzhi; Xia, Yingce; Zhang, Dongdong; Zhang, Zhirui, & Zhou, Ming. 2018. Achieving human parity on automatic chinese to english news translation. arXiv preprint arXiv:1803.05567.
Hazelkorn, Ellen. 2011. Measuring world-class excellence and the global obsession with rankings. In King, Roger; Marginson, Simon, & Naidoo, Rajani (eds.) *Handbook on globalization and higher education*. Cheltenham: Edward Elgar Publishing, 497-516.
Harry, Drew; Gordon, Eric, & Schmandt, Chris. 2012. Setting the stage for interaction: a tablet application to augment group discussion in a seminar class. In Poltrock, Steven; Simone, Carla; Grudin, Jonathan; Mark, Gloria & Riedl, John (eds.) *Proceedings of the ACM 2012 conference on Computer supported cooperative work*. New York: Association for Computing Machinery, 1071-1080.
Hultgren, Anna Kristina. 2014. English language use at the internationalised universities of Northern Europe: Is there a correlation between Englishisation and world rank? *Multilingua* 33(3-4): 389-411.
Hutchins, W. John. 1995. Machine translation: A brief history. In Koerner, Ernst Frideryk Konrad & Asher, Ronald E. (eds.) *Concise history of the language sciences*. Oxford: Pergamon Press, 431-445.
Hyland, Ken. 2017. Metadiscourse: What is it and where is it going? *Journal of Pragmatics* 113: 16-29.
Hyland, Ken. 2010. Metadiscourse: Mapping Interactions in Academic Writing. *Nordic Journal of English Studies* 9(2): 125-143.
Hyland, Ken. 2005. *Metadiscourse*. London / New York: Continuum.
Jover, Gonzalo; Fleta, Teresa, & González, Rosa. 2016. Pre-service education of primary school teachers in the context of foreign language bilingual teaching. *Bordón. Revista de pedagogía* 68(2): 121-135.
Ko, Wei-Jen; El-Kishky, Ahmed El-Kishky; Renduchintala, Adithya; Chaudhary, Vishrav; Goyal, Naman; Guzmán, Francisco; Fung, Pascale; Koehn, Philipp, & Diab, Mona (2021). Adapting high-resource NMT models to translate low-resource related languages without parallel data. *arXiv preprint arXiv:2105.15071*.
Kraker, Myra J. 2000. Classroom discourse: Teaching, learning, and learning disabilities. *Teaching and Teacher Education* 16(3): 295-313.
Kuteeva, Maria, & Mauranen, Anna. 2018. Digital academic discourse: Texts and contexts: Introduction. *Discourse, Context & Media* 24: 1-7.

Lanvers, Ursula, & Hultgren, Anna Kristina. 2018. The Englishization of European education: foreword. *European Journal of Language Policy* 10(1): 1-11.

Lasagabaster, David. 2012. El papel del inglés en el fomento del multilingüismo en la universidad. *ELIA. Estudios de Lingüística Inglesa Aplicada* 12: 13-44.

Läubli, Samuel; Castilho, Sheila; Neubig, Graham; Sennrich, Rico; Shen, Qinlan, & Toral, Antonio. 2020. A set of recommendations for assessing human–machine parity in language translation. *Journal of Artificial Intelligence Research* 67: 653-672.

Läubli, Samuel; Senrich, Rico, & Volk, Martin. 2018. Has machine translation achieved human parity? a case for document-level evaluation. arXiv preprint arXiv:1808.07048, 2018.

Lee, Joseph J. & Subtirelu, Nicholas C. 2015. Metadiscourse in the classroom: A comparative analysis of EAP lessons and university lectures. *English for Specific Purposes* 37: 52-62.

Livingstone, Kerwin Anthony. 2019. Examining the Use of Metadiscourse Markers in Academic Writing. *International Journal* 5(3): 244-254.

Llinares, Ana, & Dalton-Puffer, Christiane. 2015. The role of different tasks in CLIL students' use of evaluative language. *System* 54: 69-79.

Llinares, Ana, Morton, Tom, & Whittaker, Rachel. 2012. *The Role of Language in CLIL*. Cambridge: Cambridge University Press.

Llinares, Ana & Morton, Tom. 2010. Historical explanations as situated practice in content and language integrated learning. *Classroom Discourse* 1(1): 46-65.

Lorenzo, Francisco; Casal, Sonia, & Moore, Pam. 2010. The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied linguistics* 31(3): 418-442.

Mauranen, Anna. 2012. *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.

Mauranen, Anna. 1993. Contrastive ESP rhetoric: Metatext in Finnish-English economic texts. *English for Specific Purposes* 12: 3-22.

Martin, James R., & White, Peter R. R. (2005). *The evaluation of language: Appraisal in English*. Hampshire: Palgrave Macmillan.

McIntyre, Dan. 2009. *History of English: A resource book for students*. London: Routledge.

Molino, Alessandra. 2018. "What I'm Speaking Is Almost English...": A Corpus-Based Study of Metadiscourse in English-Medium Lectures at an Italian University. *Educational Sciences: Theory and Practice* 18(4): 935-956.

Molino, Alessandra. 2017. Repetition and rephrasing in physical sciences and engineering English-medium lectures in Italy. In Boggio, Cecilia & Molino, Alessandra (eds.) *English in Italy: Linguistic, educational and professional challenges*. Milano: FrancoAngeli, 182-202.

Morell, Teresa. 2020. EMI teacher training with a multimodal and interactive approach: A new horizon for LSP specialists. *Language Value* 12(1): 56-87.

Morell, Teresa. 2015. International paper conference presentations: A multimodal analysis to determine effectiveness. *English for Specific Purposes* 37: 137-150.

Müller, Mathias; Rios, Annette; Voita, Elena & Sennrich, Rico. 2018. A large-scale test set for the evaluation of context-aware pronoun translation in neural machine translation. In *Proceedings of the Third Conference on Machine Translation*. Belgium, Brussels: Association for Computational Linguistics, 61-72.

O'Boyle, Neil. 2015. The risks of 'university speak': relationship management and identity negotiation by mature students off campus. *International Studies in Sociology of Education* 25(2): 93-111.

O'Brien, Sharon. 2022. How to deal with errors in machine translation: Postediting. *Machine translation for everyone: Empowering users in the age of artificial intelligence* 18: 105.

Pérez-Ortiz, Juan Antonio; Forcada, Mikel L., & Sánchez-Martínez, Felipe. 2022. How neural machine translation works. *Machine translation for everyone: Empowering users in the age of artificial intelligence 18*, 141.

Popel, Martin. 2018. CUNI Transformer Neural MT System for WMT18. In *Proceedings of WMT*. Brussels, Belgium: Association for Computational Linguistics, 486-491.

Popović, Maja, & Castilho, Sheila. 2019. Are ambiguous conjunctions problematic for machine translation? In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*. Varna, Bulgaria: INCOMA Ltd, 959-966.

Rescigno, Argentina Anna; Vanmassenhove, Eva; Monti, Johanna, & Way, Andy. 2020. A Case Study of Natural Gender Phenomena in Translation. A Comparison of Google Translate, Bing Microsoft Translator and DeepL for English to Italian, French and Spanish. In *CLiC-it*.

Ruiz de Zarobe, Yolanda; Sierra, Juan Manuel, & Gallardo del Puerto, Francisco. 2011. *Content and Foreign Language Integrated Learning: Contributions to multilingualism in European contexts*. Bern: Peter Lang.

Shterionov, Dimitar; Nagle, Pat; Casanellas, Laura; Superbo, Riccardo, & O'Dowd, Tony. 2017. Empirical evaluation of NMT and PBSMT quality for large-scale translation production. In *20th Annual Conference of the European Association for Machine Translation, EAMT 2017*. Praha, 74-79.

Simpson, Rita C.; Briggs, Sarah L.; Ovens, Janine, & Swales, John M. 2002. *The Michigan corpus of academic spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.

Stasimioti, Maria & Sosoni, Vilelmini. 2019. Undergraduate Translation Students' Performance and Attitude vis-à-vis Machine Translation and Post-editing: Does Training Play a Role? In *41st Translating and the Computer Conference* (TC41) *Proceedings*. London: The International Association for Advancement in Language Technology (AsLing), 125-136.

Stavans, Anat & Hoffmann, Charlotte. 2015. *Multilingualism*. Cambridge: Cambridge University Press.

Soter, Anna O.; Wilkinson, Ian A.; Murphy, P. Karen; Rudge, Lucila; Reninger, Kristin, & Edwards, Margaret. 2008. What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research* 47(6): 372-391.

Swales, John. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Sylvén, Liss Kerstin. 2017. Motivation, second language learning and CLIL. *Applied linguistics perspectives on CLIL*, 51-66.

Takimoto, Masahiro. 2015. A corpus-based analysis of hedges and boosters in English academic articles. *Indonesian Journal of Applied Linguistics* 5(1): 95-105.

Thompson, Geoff. 2001. Interaction in academic writing: Learning to argue with the readers. *Applied Linguistics* 22: 58-78.

Universitat Jaume I. 2019. *Estratègia de convivència i promoció lingüística ECOPOL*. http://documents.uji.es/alfresco/d/d/workspace/SpacesStore/e83e1f8b-7593-4212-8015-abea76bf34d8/XX_SLTJM_ECOPOLfinal-oct19.pdf?guest=true

Van Rensburg, Alta; Snyman, Cobus, & Lotz, Susan. 2012. Applying Google Translate in a higher education environment: Translation products assessed. *Southern African linguistics and applied language studies* 30(4): 511-524.

Williams, Malcolm. 2010. Translating Metadiscourse: An Explanatory Analysis of Problems in Students' Work. *Mutatis Mutandis: Revista Latinoamericana de Traducción* 3(1): 73-90.

Zare, Javad & Tavakoli, Mansoor. 2016. The use of personal metadiscourse over monologic and dialogic modes of academic speech. *Discourse Processes* 54(2): 163-175.

Acknowledgments

Lucía Bellés-Calvera acknowledges the funding received from Universitat Jaume I (research fellowship: PREDOC/2017/36).