The use of code glosses in MA dissertations in English and in Italian: A pilot study in an EMI context

Marina Bondi*, University of Modena and Reggio Emilia, Italy
Jessica Jane Nocella, University of Modena and Reggio Emilia, Italy
(*Corresponding Author: marina.bondi@unimore.it)

Within the English Medium of Instruction (EMI) framework adopted by many European universities, students listen to lectures, present, write, and take exams in English. Research on EMI settings mostly focused on lectures, leaving scant attention to student writing and in particular to the writing of dissertations in the Italian context. This paper addresses this gap by investigating lexico-grammatical occurrences of selected metadiscursive features adopted in the final dissertations of Italian masters' students in business and economics courses. The dissertations are retrieved from the MoReThesisCorpus, a large digital repository of theses and dissertations at the University of Modena e Reggio Emilia. By paying close attention to reformulations and exemplifications, our findings reveal a developmental interconnection between spoken and written metadiscursive features.

Keywords: EMI, ELF, code glosses, masters' dissertations, metadiscursive features

Introduction

English as lingua franca (ELF), situations where English is the main language of communication, plays an important role in different contexts and specifically in academia, where non-native speakers of English (NNS) often outnumber those of native speakers (NS) (Mauranen, 2010). With the increasing number of programmes allowing student mobility across countries (e.g., Erasmus) and the growing number of joint degrees in academic settings all over the world, English Medium Instruction (EMI) is becoming increasingly popular, confirming the prominent role of English as the main language of communication and knowledge dissemination (Clark, 2018; Mauranen, 2010; Wu et al., 2020). EMI instruction involves situations of

intercultural use of English in academic contexts, which often means striving for greater explicitness in search of mutual understanding (Mauranen et al., 2010). This may imply greater effort in indicating local organization, negotiating topics, and careful use of metadiscourse (i.e., references to the discourse itself) (Mauranen, 2012), with a view to communicative efficiency, clarity, and explicitness (Wu et al., 2020). Translanguaging, which has been defined both as a systematic and more spontaneous shift from one language to another (Coyle et al., 2010; Li, 2011), has also been of great interest in EMI contexts.

Given this background, it is worth investigating how English is shaped by local and global linguistic forces, starting from the distinctiveness of the sociolinguistic contexts in which English is studied and the functional ranges and domains in which it is used. Much attention has been paid to oral language in EMI (Doiz & Lasagabaster, 2022; Jensen & Thøgersen, 2020), examining the possible effects of ELF on NNSs' language use (Clark, 2018) as well as the use of metadiscourse markers and discourse strategies adopted in both L1 and EMI lectures (Ädel 2010; Aguilar-Pérez & Khan, 2022; Arkın & Osam, 2015; Costa & Mariotti, 2017; Doiz & Lasagabaster, 2022; Molino, 2018; Wu et al., 2020). Analysis of EMI lectures in the Italian context, for example, has explored several features, such as explanatory strategies (Gotti, 2014), code switching (Gotti, 2015), lexical bundles (Molino, 2019), stance features (Solly, 2018), and metadiscursive strategies in general (Gotti, 2014; Molino, 2018), for their potential contribution to enhancing clarity.

Studies on ELF (Mauranen, 2012, 2023) have underlined the importance of metadiscourse as a key feature of intercultural communication, contributing to discourse explicitness and to comprehension needs. Because of its role in enhancing clarity, metadiscourse is often taken to be a form of 'sensitivity to the situational demands of the EMI classroom' (Molino, 2018, p. 606).

Metadiscourse puts together an array of linguistic strategies that allow writers/speakers and their respective audiences 'mutual acts of comprehension and involvement' (Hyland, 2005, p. 4). According to the interpersonal model of metadiscourse (Hyland, 2005; Hyland & Tse, 2004), there are two types of discursive resources, namely, an interactive one and an interactional one. The former is centred on the reflexive side of metadiscourse (Ädel, 2006; Mauranen, 1993, 2023) and is related to the organization of discourse and to what needs to be further elaborated in the text in order to make it clear to the reader. This includes linguistic features such as transitions (e.g., *in addition*, *but*, *thus*), frame markers (e.g., *finally*, *to conclude*), endophoric markers (e.g., *as noted above*, *see Fig. x*), evidential markers (e.g., *according to*, *as X states*), and code glosses (e.g., *namely*, *such as*) (Hyland, 2007). Interactional resources, on the other hand, are concerned with the relationship that the writer builds with the reader, which involves features such as, attitude markers (e.g., *unfortunately*, *surprisingly*), engagement markers (e.g., *consider*, *note that*),

hedges (e.g., *might*, *perhaps*, *possible*), boosters (e.g., *it is clear that*, *in fact*), and forms of self-mention (e.g., *I*, *we*, *my*, *our*) (Hyland, 2010; 2015; 2017; Hyland & Tse, 2004).

A focus on metadiscourse has played an important part in studies on academic writing (Hyland, 2017). Research has shown that the use of metadiscursive features varies across L1 and L2 writing (e.g., Kobayashi, 2020; Murillo, 2012), often paying particular attention to variation in aspects of authorial identity, stance, and audience engagement (e.g., Mur-Dueñas & Šinkūnienė, 2016). An area that certainly needs to be studied further is that of MA/MSc and PhD thesis writing. Dissertations can help trace important stages in the development of academic literacy through the formative years when students – as novices – experiment with developing their identity as academic writers and start shaping their disciplinary voices as rhetorical subjects (Paré et al., 2011). Studying the rhetorical and linguistic features of dissertations can help understand developmental perspectives, as well as disciplinary variation and variation across first, second, or additional languages.

In this context, our study aims at contributing to the knowledge of metadiscourse in an EMI academic context through a micro-diachronic qualitative case study on MA dissertations written by students enrolled in an EMI master's degree programme. We look at EMI as a site of complex language contact, much in the same way as ELF (Mauranen, 2018), even if the introduction of EMI in the Italian context is relatively recent. The case under investigation is even more recent: the master's course has been taught entirely in English since 2015/2016 and candidates in this programme have started writing within the EMI context since 2017. Against this background we set out to:

1. explore whether, and if so how, specific metadiscursive features (i.e., code glosses) in both English and Italian have changed in students' dissertations since the introduction of EMI lessons;

2. carry out a comparison of dissertations written in Italian in the same course over the same time frame. Has the EMI course in any way influenced the way students write in their L1?

For reasons of space, the present study centres on selected features of metadiscourse, looking in particular at the use of code glosses – exemplification and reformulation – which seem particularly apt to represent the writer's sensitivity to potential audience needs.

After this brief overview of studies on EMI and metadiscursive features, the rest of the paper is structured as follows. The next section presents the methodology, with details on the corpus and on the analytical framework used for this study. The findings of our analysis are then introduced by looking first at exemplification and reformulation in English, then at the

Italian corpus collected for parallel reference. The concluding remarks sum up the results across both languages and draw the implications of this pilot study while also highlighting its limitations.

Methodology

The corpus

In order to carry out our micro-diachronic case study, we examined a small corpus of dissertations of students enrolled in the master's course in international management at the Department of Economics of the University of Modena and Reggio Emilia. This programme was chosen because it allows us to compare dissertations written in English with dissertations written in Italian, as well as to compare work produced before the switch to EMI with work produced by students enrolled in EMI courses. The EMI programme offers students the possibility to write their dissertation in either language, English or Italian. Other EMI degree programmes would have offered higher numbers of dissertations available in English, but fewer dissertations in Italian (if any), hence we excluded them from the current study.

Our small sample of data was retrieved from the MoReThesisCorpus which consists of all dissertations submitted by students at the University of Modena and Reggio Emilia between 2011 and 2020. The MoReThesisCorpus that we are developing is a corpus of dissertations from the official repository of the University that will allow for comparisons across different stages of education (BA/MA/PhD) and across different languages (Bondi & Di Cristofaro, 2023). Focusing on a specific Italian university, the corpus includes a wide range of disciplines both in Italian and in English as a result of the EMI framework. We chose to use the MA in International Management programme. First, we selected all international management dissertations which were written in English (total of fourteen texts) and were available on Open Access. Then, we randomly collected another set of dissertations available in Italian for the parallel timespan. As this programme switched to an EMI framework in 2015/2016, the two corpora were further subdivided into two temporal sub-corpora with the same distribution to allow for parallel analysis: the first includes dissertations written between 2011 and 2016 (pre-EMI) and the second includes dissertations between 2017 and 2020 (EMI). Table 1 displays the structure of our two corpora highlighting the number of texts and tokens for each subset.

Table 1   Distribution of dissertations in English and Italian in the corpus

Corpus		Number of texts	Number of tokens
English corpus	Pre-EMI	7	1,231,240
	EMI	7	1,238,677
Italian corpus	Pre-EMI	7	2,081,119
	EMI	7	1,366,211

Analytical framework

In order to analyse our data quantitatively, we availed ourselves of Wordsmith Tools 7.0 (Scott, 2016) to explore interactive metadiscursive features (Hyland, 2007) adopted in dissertations. In particular, we studied code glosses – forms of reformulation and exemplification based on Hyland's (2007) model – pre- and during EMI contexts. With reformulation, we looked at how clauses are re-elaborated to facilitate understanding of the content (e.g., use of appositions, parentheses, quotes, and reformulation markers), while, with exemplification, we looked at how the content is clarified through concrete examples (Hyland, 2007) to explain general questions and to make them closer to the audience and more easily remembered. Examining concordance lines (Sinclair, 2004), we also identified discourse acts behind each code gloss marker to explore syntactic patterns and the rhetorical functions of each feature.

In particular, for exemplification, we adopted Su et al.'s (2021) framework. Table 2 below provides explanations and examples for each of the semantic features that were considered in our work. The same method was applied to the analysis of the Italian corpus.

Table 2   Exemplification framework (adapted from Su et al., 2021)

Semantic feature	Explanation	Example
Exemplified	is the argument that is being clarified through the exemplification, which can be further divided into exemplification-sub-category and relevant-study	*Efficiency-centred design refers to how firms use their activity system design to aim at achieving greater efficiency through reducing transaction costs*. For example, a local firm…
Exemplification sub-category	specifies a 'more abstract or subordinate category– used' (Su et al. 2021, p. 126)	the applicability of the rules of conflict dictated by these normative tools is not totally excluded; *specifically, only the applicability of the single dispositions of the* ….

Semantic feature	Explanation	Example
Relevant study	is used to cite authors and relevant literature	Some empirical studies *(e.g., Asiedu, 2006;Vijayakumar et al., 2010)*
Indicator	markers that introduce the exemplification	The impact on society is significant, *as for example* the reduction of pollution…
Hinge	signals a connection between parts of the text	*In fact*, for example, in the 19th century and earlier most workers also regularly worked at home,
Supporting statements	adds information to clarify and support the exemplified or the exemplification	Guanxi tags people with a relationship from others. '*In China, people have a strong tendency to divide people into two categories: those they know and have trust-based relations with (in-group) and strangers(out-group)*' (Worm, 1998: 185)

Reformulations guarantee textual cohesion and facilitate discursive progression by providing a retrogressive interpretation of the previous utterance, thus allowing speakers to explain, rephrase, reconsider, summarize, or even distance themselves from it. They expand on the information previously given and facilitate – or guide – the hearer's understanding. Although reformulations are often regarded as 'repairs' in unplanned spoken discourse, they are purposeful in writing. The identification of these processes has long represented an open question in pragmatics and led to consider different types of reformulations. These range from those that can be characterized as paraphrastic, based on a semantic equivalence between two members, e.g., *GDP* (*Gross Domestic Product*), to those entailing a change in the enunciative perspective (Garcés Gómez, 2009) in forms that rather than just expressing an equivalence, actually create the equivalence (Cuenca, 2003).

For our data analysis we classified our reformulation markers according to Hyland's (2007) categories and subcategories:

a) expansion (with subcategories: explanation and implication)

b) reduction (with subcategories: paraphrase and specification).

For purposes of comparability with other English data, we chose to adopt the list of markers used by Hyland (2005; 2007) after checking it against a small portion of the corpus. Similarly, for Italian we took advantage of the

markers identified in the literature on exemplification (Lo Baido, 2018) and reformulation (Bazzanella, 1995, 2001, 2006; Ciabarri, 2013; Fiorentini and Sansò, 2017) and checked them against a sample of the corpus.

Starting from such list of forms of reformulation and exemplification, we manually tagged the metadiscursive features present in our corpus using the 'Set' feature in Wordsmith Tool 7. This allowed us to analyse our data from both a quantitative and a qualitative perspective. Each feature's rate of occurrence (proportion) was calculated by taking the number of occurrences divided by the number of words in the same text. Given the differences in text length in the two corpora, in order to ensure comparability, the results were then normalized to 10,000 words (pttw: per ten thousand words).

Findings

Code glosses

In this section, we will explore code glosses present in both the English and the Italian corpus. Table 3 shows the normalized frequency (per 10,000 words) of the overall number of exemplification and reformulation markers in our corpora.

Table 3   Normalized frequency of code glosses in the English corpus

Code glosses	English corpus		Italian corpus	
	Pre-EMI	EMI	Pre-EMI	EMI
Exemplification	3.52	2.96	2.61	2.44
Reformulation	2.67	2.74	4.14	2.13

According to our results, in the English corpus, exemplification is the preferred feature in both periods. However, while in pre-EMI dissertations exemplification features are much more frequent than those of reformulation, a more homogeneous distribution of the two features seems to be present in EMI dissertations.

In contrast, in the Italian corpus, there seems to be a reverse trend with a clear preference for reformulation markers, which are more frequently used in the pre-EMI period. Exemplification markers, however, become more prominent than reformulation markers in the EMI context.

Overall, the preliminary quantitative overview also highlights that the frequency of these markers is quite limited. Moreover, the EMI data across languages shows greater balance between the two types of code glosses. The limited dimension of our corpus, however, suggests that a qualitative analysis of the specific markers used and of their context might be more interesting.

Exemplification in the English corpus

Table 4   Exemplification markers in the English corpus

	English corpus	
Exemplification markers	pre-EMI (normed frequency)	EMI (normed frequency)
for example	0.42	0.61
an example of	0.01	0.04
for instance	0.26	0.41
such as	1.83	1.29
e.g.	0.12	0.26
like	0.88	0.35
say	0	0
Total	3.52	2.96

Table 4 displays normalized frequencies of exemplification markers in the English corpus for both periods under observation. While pre-EMI dissertations present a higher number of exemplification markers overall (3.52), EMI dissertations show a more balanced choice of the markers used. In fact, in the pre-EMI period there is a high preference for *like* (0.88) and *such as* (1.83), and a less frequent use of other markers (ranging from 0.01 to 0.42). In EMI dissertations, instead, a more balanced choice of lexical markers (all ranging from 0.04 to 1.29) seems to be present with a slight preference for *for example* (0.61) and *such as* (1.29).

By carrying out a concordance analysis, other qualitative differences emerge. In pre-EMI dissertations, exemplification markers all fall in the same rhetorical patterns where the indicator (or exemplification marker) falls between the exemplified (when present) and its exemplification (Example 1). This structure is also present in EMI dissertations (Example 1).

Example 1

along with international players **like** Unilever, Cream Bell, Movenpick, etc.

However, while in pre-EMI dissertations this seems to be the only pattern, in EMI dissertations there seems to be more variation in terms of syntactic patterns, especially with regards to *for example*. As can be seen in the examples below, *for example* is frequently followed by a supporting statement (Example 2), in some cases introduced by the researcher/writer performing the act of exemplifying (Example 3).

Example 2

In addition <hinge>, **for example** <indicator>, 'employers can cut costs substantially: by letting their workers work from home one or more days they can reduce the amount of office space they need considerably' (Popma 2013, p. 8) <supporting statement>.

Example 3

In a lot of cases that I examined <researcher>, **for example** <indicator>, the one of Franco Madama and the other one of Alberto Rossetti <exemplification>

Moreover, both in pre-EMI and EMI dissertations, the marker *like* is used similarly as it precedes the exemplification sub-category. *Such as*, on the other hand, is not only used with the exemplification sub-category through lists or parentheses in pre-EMI dissertations, but it is also used to expand the content. This should not be surprising given the high use of *such as* in pre-EMI dissertations (Table 4), where it seems to be the main exemplification marker, thus adopting the functions of other markers (e.g., *for example*, *for instance*). The marker *e.g.*, which mostly appears in brackets in both sub-corpora, is used to present only the exemplification sub-category in EMI dissertations, while in pre-EMI it also precedes the relevant studies sub-category.

Reformulation in the English corpus

Table 5 below shows the normed frequency of the reformulation markers used in the pre-EMI and EMI period.

Table 5   Reformulation markers in pre-EMI and EMI dissertations

Reformulation markers	English corpus	
	pre-EMI (normed frequency)	EMI (normed frequency)
Parentheses	0.47	0.35
in particular, *particular*	**0.96**	**1.02**
that is	0.06	0.07
especially	**0.67**	0.4
in other words,	0.07	0.12
namely	0.09	0.14
specifically,	0.13	0.23
or x	0.04	0.14
which means, *this means*	0.18	0.27
Total	2.67	2.74

From a quantitative point of view, we notice that similarly to the exemplification results, reformulation markers are more equally distributed in EMI dissertations. In fact, besides *in particular* and *particularly* (Table 5), the other markers, *that is* and *especially*, seem to equally range from 0.07 to 0.4, respectively. In pre-EMI dissertations, instead, the most frequent markers are *in particular*, *particularly* (0.96), and *especially* (0.67), while the remaining markers are less frequent, ranging between 0.06 and 0.47.

As visible from Table 6, both periods observed seem to display a preference for reduction functions, which represent circa two thirds of both sub-corpora.

Table 6   Percentages of reformulation markers functions in pre-EMI and EMI dissertations

Functions	pre-EMI (%)	EMI (%)
Expansion	31.71	33.35
Reduction	68.29	66.65

For a clearer view of the data, Table 7 displays the percentage of subfunctions of both expansion and reduction in more detail. While specification is high in both sub-corpora (63% in pre-EMI dissertations and 58% in the EMI dissertations), explanation is higher in pre-EMI dissertations (29%). Nonetheless, in EMI dissertations we see a rise in the implication and paraphrasis subfunctions, confirming the trend towards variation, which is similar to what we had already observed in the use of a range of reformulation markers (Table 5).

Table 7   Percentages of reformulation markers' subfunctions in pre-EMI and EMI dissertations.

Subfunctions	Pre-EMI (%)	EMI (%)
Explanation	29.45	23.34
Implication	2.26	10.01
Paraphrasis	5.19	8.52
Specification	63.1	58.13

Looking at our data qualitatively, we can notice other patterns. In pre-EMI dissertations, markers that are typically used to introduce specification (e.g., *particularly*, *in particular*, *specifically*, *especially*) are preferred and are usually followed by nominal structures (Example 4). This reminds us of the nominal exemplification sub-category seen earlier (see 'Exemplification in the English Corpus' section). In EMI dissertations, on the other hand, such markers

frequently introduce clauses that not only specify the previous content, but they also expand it and add information, introducing more complex patterns (Example 5).

Example 5

such situations, refugees **in particular** women and children, become more…

Specifically, the reshoring trend is revised with an attempt to gauge its impact on consumer perception, drawing upon extant literature and providing a collection of useful insights that could help companies gain and leverage improved relationships with their customers**.**

In pre-EMI dissertations, explanations are mostly achieved through the use of parentheses. This is a more direct way to expand textual elements with a reduced use of 'propositional embellishments' (Hyland, 2007, p. 267) that guide the reader through the understanding of the text. In EMI dissertations, explanation is not only through parentheses, but also through other markers such as *what means, this means, that is to say* (Example 6). These markers are much more frequent in EMI dissertations, suggesting more variation in their use. In fact, they are not only used to paraphrase and explain previous content, but they are also frequently adopted to draw implications (Example 7). This might suggest an increasing attention to clarifying content as well as greater familiarity with a wide range of markers and patterns.

Example 6

Another relevant dimension invoked in the literature is the scope of the BMI – **that is,** how much of a BM is affected by a BMI.

Example 7

To develop the workers and increase their engagement. **This means** to leave the traditional way to control the workers with punishment or bonus.

Code glosses in the Italian corpus: a comparison

A parallel analysis of a small corpus of dissertations in Italian was also conducted for comparison, following the same criteria.

Exemplification in the Italian corpus

As already discussed, in the Italian corpus exemplification markers are less frequent than reformulation markers. Table 8 displays the frequency of each of the markers considered.

Table 8   Exemplification markers' frequency in the Italian corpus

Exemplification markers	Pre-EMI (normed frequency)	EMI (normed frequency)
***ad/per esempio* [for example]**	**1.11**	**0.51**
esempi (sono esempi...) [these are examples…]	0.15	0.04
esempio (ne è esempio...un esempio... per fare un esempio, altro esempio) [to provide an example of]	0.13	0.05
es. /e.g. [ex., e.g.]	0.11	0.06
magari [perhaps]	0.03	0.02
***come* [like]**	**0.89**	**1.08**
si pensi, si veda, si prenda, si consideri... [think of, see, consider, take…]	0.01	0.03
basti pensare/ricordare, ricordiamo, prendiamo [think of/remember]	0.004	0.01
***anche* [also]**	**0.1**	**0.39**
column and parentheses	**0.08**	**0.25**
diciamo [let's say]	0	0
mettiamo/poniamo il caso [suppose we]	0	0
per dire [so to say]	0	0
caso (facciamo il caso, prendiamo il caso..facciamo un/il caso) [let's consider]	0	0
che ne so/che so... [imagine]	0	0
faccio per dire [to say]	0	0
io prendo [let's say]	0	0
mettiamo un ipotesi/facciamo un'ipotesi [suppose]	0	0
non so [don't know]	0	0
per dirne una [to cite]	0	0
per ipotesi [in the case]	0	0
tipo [like]	0	0
vedi [see]	0	0
Total	2.61	2.44

Our data in Table 8 shows no important changes in the use of exemplification markers between the two periods. A slight decrease in the use of *ad esempio* and *per esempio* (from 1.11 to 0.51) in favour of an increase in the use of *anche* (from 0.1 to 0.39), *come* (from 0.89 to 1.08), punctuation (parentheses and column – from 0.08 to 0.25) is worth mentioning, however, suggesting more lexical and syntactical variation in EMI dissertations.

Qualitatively, results also show similar patterns in which exemplification is achieved in EMI and pre-EMI dissertations. For instance, we find indicator–exemplified–exemplification / exemplified–indicator–exemplification (Example 8). Exemplification markers also frequently fall in the exemplified–indicator–supporting statement patterns (Example 9).

Example 8

come una domanda di prodotto non sollecitata <exemplified> […] dovuta<hinge> **ad esempio** <indicator> a un fenomeno di saturazione del mercato<exemplification>. [… as an unsolicited product demand due for example to a phenomenon of domestic market saturation]

Example 9

Ad esempio <indicator>, Kapferer (1986), ha analizzato il modo in cui le insegne possono differenziarsi le une dalle altre, e ha sostenuto […] <supporting statement> [For example, Kapferer (1986) analyzed how signs can differentiate themselves from one another, and argued that…]

On the other hand, in both corpora punctuation markers (e.g., parentheses and column) are either used to narrow down a subject (explanation sub-category), or to cite authors (explanation–relevant studies).

The use of similar patterns across the two phases suggests that both lingua-cultural contexts might be abiding to writing conventions belonging to the same discourse community. The limited frequency of exemplification markers in the Italian corpus, on the other hand, seems to point to a specificity of the local academic tradition. The small corpus does not show substantial impact of the EMI context as far as exemplification is concerned.

Reformulation in the Italian corpus

Table 9 displays the normalized frequency of reformulation markers in both pre-EMI and EMI dissertations. From a quantitative point of view, we notice a decrease in the general use of reformulation markers as in pre-EMI dissertations. They are almost twice as many than in EMI dissertations. However, when looking closely, we notice that, while *ovvero* and *in particolare* are the preferred markers in both corpora, in pre-EMI dissertations they are much more frequent than in EMI ones with respectively 1.01 and 1.41 (in bold in Table 9). Similar to the English EMI corpus, in EMI dissertations in Italian,

there seems to be a more equal distribution among markers which range from 0.01 to 0.87.

Table 9   Reformulation markers in the Italian corpus

	Italian corpus	
Reformulation markers	Pre-EMI (normed frequency)	EMI (normed frequency)
voglio dire [I mean]	0	0
vale a dire [meaning]	0.06	0.01
nel senso [this means]	0.04	0.01
diciamo [let's say]	0	0
in altre parole [in other words]	0.14	0.05
anzi [actually]	0.03	0.01
parentesi [parentheses]	0.25	0.25
***ovvero* [or rather/in other words]**	**1.01**	**0.87**
anche detto [also known]	0	0
che significa [which means]	0.03	0
segnatamente [especially]	0	0
specificatamente [specifically]	0.01	0.02
***in particolare* [in particular]**	**1.41**	**0.68**
particolarmente [particularly]	0	0
cioè [that is]	0.55	0.19
ossia [that is]	0.61	0.05
Total	4.14	2.13

When looking at the percentage of the functions of markers in Table 10, we notice that in both sub-corpora there is a similar proportion between the two, with reduction representing circa 57% of the functions and expansion circa 43%.

Table 10   Percentages of reformulation markers functions in Italian pre-EMI and EMI dissertations

Functions	PRE-EMI (%)	EMI (%)
Expansion	43.69	42.11
Reduction	56.31	57.89

As it becomes clear from Table 11, such equal distribution is also present in the distribution of the subfunctions with similar percentages for

explanation (circa 40%) and specification (circa 50%) in both sub-corpora. On the other hand, implication and paraphrase, which were becoming prominent in the English corpus, seem very limited here.

Table 11   Percentages of reformulation markers subfunctions in Italian pre-EMI and EMI dissertations

Subfunctions	pre-EMI (%)	EMI (%)
Explanation	40	41.12
Implication	3.69	0.99
Paraphrase	7.14	6.91
Specification	49.17	50.99

Overall, from a quantitative point of view, the small Italian sub-corpus shows little diachronic variation in the macro-categories, but also a marked imbalance in the various subfunctions, when compared to the English corpus. Given the limitations of the corpus, however, it is difficult to say whether this might be attributed to individual preferences or to cultural rhetorical preferences.

In terms of qualitative analysis, there seems to be no significant change in the use of such reformulation markers. However, it is worth mentioning that in EMI dissertations, unlike in pre-EMI ones, specification is frequently used to describe research steps or the focus and aim of writers' work (Example 10). This might suggest a possibly greater awareness of the function of metadiscourse in highlighting the structure of research in the dissertation:

Example 10

In particolare, l'obiettivo della parte empirica della presente ricerca, è quello di determinare quali siano i fattori che influenzano maggiormente la decisione dei consumatori italiani di diventare clienti più o meno regolari di Carrefour […]

[In particular, the aim of the empirical part of the present research is to determine what factors most influence Italian consumers to become regular customers of Carrefour]

Another micro-shift between the two sub-corpora might be that of translations from English to Italian. In fact, we noticed that in pre-EMI dissertations it is frequent to encounter direct translations of English technical terms or fixed expressions into Italian (Example 11). In EMI dissertations instead, writers prefer to reformulate or explain the concept in Italian rather than providing the reader with the literal translation of such terms (Example 12).

Example 11

'Make it happen', **ovvero** 'Fai in modo che accada'

Example 12

funzionamento passivo di tipo input-output, basato, **cioè**, sul funzion-amento stimolo-risposta. [Passive input–output type functioning, that is, based on stimulus-response functioning.]

This might be a consequence of a constant exposure to English through EMI lectures, leading students to take for granted the meaning of certain technical terms. This could be potentially linked to translanguaging practices that might have been adopted during EMI lessons, leading students to spontaneously switch from one language to another.

Conclusions and limitations

This study investigated the use of code glosses in master's level dissertations written both in English and Italian by university students who had graduated from a Department of Business and Economics pre- and post-EMI instructional contexts. The results have shown that students recognized the importance of reformulation and exemplification in the development of their topic, giving both strategies substantially equal importance, even if an overall limited role.

Exemplifications turned out to be the preferred strategy when writing in English. While showing a minimally decreasing trend in the shift from pre-EMI courses to EMI ones, exemplifications manifest a clear increase in the diversity of markers and patterns used. Reformulations, on the other hand, were more stable in (lower) frequency but also showed a wider diversity of use in the EMI period. The data about dissertations in English, then, shows increasing attention to the need for clarifying content in different ways (e.g., via implications), as well as greater familiarity with a wide range of markers and patterns.

The corpus of dissertations written in Italian was analysed along the same lines, revealing different patterns. While exemplification was overall more limited, reformulation markers were characterized by a marked decrease but also by a clearer connection with key elements of structure and a reduction in the need to paraphrase English loan words, while reflecting a greater use of translanguaging.

What can be considered interesting beyond the limited value of frequency data in our small corpus is, on the one hand, the preference for exemplification when writing in EMI across languages and, on the other, the growing awareness of a range of markers and their key functions. From the point of

view of the transition to EMI, this might imply a developmental effect on lexico-grammatical diversity more than on the rhetorical organization of the text. To go back to our initial research questions about the impact of EMI on writing in English and in Italian, we can conclude that EMI dissertations in both languages are characterized by a similar balance between the two types of code glosses and an increased awareness of the language resources with a more equal distribution among markers.

The shift towards a more balanced use of the two types of code glosses, as well as the wider range of markers, can be related to the intercultural EMI context. The fact that the role of code glosses remains altogether limited could be related to the students' status as novice writers who are still developing their awareness of the conventions of academic discourse. Another important factor of course could be the local academic tradition, which is quite distant from the academic traditions of anglophone contexts in the regular use of exemplification. All these hypotheses would deserve exploration through a larger corpus and a comparison with similar dissertations from other cultural backgrounds.

Some of the limitations of this study have been partly anticipated in the methodology section. The most important is related to the size of the corpus. The corpus consists of a very small sample of texts, as we have only analysed fourteen dissertations in English and fourteen in Italian. The results should be checked against a larger corpus including more dissertations from the MoReThesisCorpus or from other universities in Italy. It would also be interesting to compare the results of this small pilot study with the results obtained from more dissertations and across disciplines. It might help us refine the description of how management students write compared to other disciplines and in different language preferences.

This study provided an opportunity to explore one of the many possibilities offered by the compilation of a corpus of L2 dissertations. The compilation of the MoReThesis Corpus at the University of Modena and Reggio Emilia intends to offer a resource for the study of academic discourse from different perspectives. As such, the MoReThesisCorpus is a corpus repository, allowing for the creation of ad-hoc sub-corpora tailored to a wide range of analytic purposes. These include all areas of English for Academic Purposes (EAP) with special attention to varieties of academic English.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

References

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.

Ädel, A. (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*, 69–97.

Aguilar-Pérez, M., & Khan, S. (2022). Metadiscourse use when shifting from L1 to EMI lecturing: Implications for teacher training. *Innovation in Language Learning and Teaching*, *16*, 1–15.

Arkın, E., & Osam N. (2015). English-medium higher education: A case study in a Turkish university context. In S. Dimova, A. K. Hultgren, & C. Jensen (Eds.), *English-medium instruction in European higher education* (pp. 177–199). Walter de Gruyter.

Bazzanella, C. (1995). I segnali discorsivi. In L. Renzi, G. Salvi, & A. Cardinaletti (Eds.), *Grande grammatica italiana di consultazione*, vol. III (pp. 225–257). Il Mulino.

Bazzanella, C. (2001). I segnali discorsivi tra parlato e scritto. In M. Dardano, A. Pelo, & A. Stefinlongo (Eds.), *Scritto e parlato. Metodi*, *testi e contesti* (pp. 79–97). Aracne.

Bazzanella, C. (2006). Discourse markers in Italian: Towards a "compositional" meaning. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 449–464). Elsevier.

Bondi, M., & Di Cristofaro M. (2023). MoReThesisCorpus: Documenting academic language as used in the theses submitted to the University of Modena and Reggio Emilia. *Iperstoria*, *21*, 9–30.

Ciabarri, F. (2013). Italian reformulation markers: A study on spoken and written language. In C. Bolly & L. Degand (Eds.), *Across the line of speech and writing variation*: *Corpora and language in use. Proceedings 2* (pp. 113–127). Presses Universitaires de Louvain.

Clark, C. (2018). The case of the non-native English speaker in EMI. In C. M. Coonan, A. Bier, & E. Ballarin (Eds.), *La didattica delle lingue nel nuovo millennio* (pp. 563–576). Edizioni Ca' Foscari.

Costa, F., & Mariotti, C. (2017). Differences in content presentation and learning outcomes in English Medium Instruction (EMI) vs Italian-Medium Instruction (IMI) contexts. In J. Valcke & R. Wilkinson (Eds.), *Integrating content and language in higher education: Perspectives on professional practice* (pp. 187–204). Peter Lang.

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

Cuenca, M. J. (2003). Two ways to reformulate: A contrastive analysis of reformulation markers. *Journal of Pragmatics*, *35*(7), 1069–1093.

Doiz, A., & Lasagabaster, D. (2022). Looking into English-medium instruction teachers' metadiscourse: An ELF perspective. *System*, *105*, 102730.

Garcés Gómez, M. P. (Ed.). (2009). *La reformulación del discurso en español en comparación con otras lenguas: Catalán, Francés, Italiano, Inglés, Alemán e Islandés.* Universidad Carlos III, Boletín Oficial del Estado.

Gotti, M. (2014). Explanatory strategies in university courses taught in ELF. *Journal of English as a Lingua Franca, 3*(2), 337–361.

Gotti, M. (2015). Code-switching and plurilingualism in English-medium education for academic and professional purposes. *Language Learning in Higher Education, 5*(1), 83–103.

Fiorentini, I., & Sansò, A. (2017). Reformulation markers and their functions: Two case studies from Italians. *Journal of Pragmatics, 120*, 54–72.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* Bloomsbury.

Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics, 28*(2), 266–285.

Hyland, K. (2010). Metadiscourse: Mapping interactions in academic writing. *Nordic Journal of English Studies, 9*(2), 125–143.

Hyland, K. (2015). Metadiscourse. In K. Tracy, T. Sandel, & C. Ilie (Eds.), *The international encyclopedia of language and social interaction* (pp. 1–11). Wiley.

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113*, 16–29.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25*(2), 156–177.

Jensen, C., & Thøgersen, J. (2020). Comprehensibility, lecture recall and attitudes in EMI. *Journal of English for Academic Purposes, 48*, 100912.

Kobayashi, Y. (2020). Rhetorical preferences in L2 writings: A contrastive analysis of metadiscourse markers. *Journal of Modern Languages, 30*(2), 1.

Li, W. (2011). Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics, 43*, 1222–1235.

Lo Baido, M. C. (2018). Categorization via exemplification: evidence from Italian. *Folia Linguistica Historica, 39*, 69–95.

Mauranen, A. (1993). *Cultural differences in academic rhetoric: A textlinguistic study.* Peter Lang.

Mauranen, A. (2010). Features of English as a lingua franca in academia. *Helsinki English Studies, 6*, 6–28.

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

Mauranen, A. (2018). Second language acquisition, world Englishes, and English as a lingua franca (ELF). *World Englishes, 37*(1), 106–119.

Mauranen, A. (2023). *Reflexively speaking: Metadiscourse in English as a Lingua Franca.* De Gruyter Mouton.

Mauranen, A., Hynninen, N., & Ranta, E. (2010). English as an academic lingua franca: The ELFA project. *English for Specific Purposes, 29*(3),183–190.

Molino, A. (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, A. (2019). Lexical bundles in EMI lectures: An exploratory study. *Textus, English Studies in Italy, 1*, 119–140.

Mur-Dueñas, P., & Šinkūnienė, J. (2016). Self-reference in research articles across Europe and Asia: A review of studies. *Brno Studies in English, 42*, 71–92.

Murillo, S. (2012). The use of reformulation markers in business management research articles: An intercultural analysis. *International Journal of Corpus Linguistics, 17*(1), 64–90.

Paré, A., Starke-Meyerring, D., & McAlpine, L. (2011). Knowledge and identity work in the supervision of doctoral student writing: Shaping rhetorical subjects. In D. Starke-Meyerring (Ed.), *Writing in knowledge societies* (pp. 215–236). Parlor Press.

Scott, M. (2016). *WordSmith Tools* (Version 7.0) [Computer Software]. Lexical analysis software. https://lexically.net/wordsmith/downl.

Solly, M. (2018). Taking stance in English as a lingua franca: Managing interpersonal relations in academic lectures / the discursive construal of trust in the dynamics of knowledge diffusion. *Journal of English as a Lingua Franca, 7(2),* 415–420.

Su, H., Zhang, Y., & Lu, X. (2021). Applying local grammars to the diachronic investigation of discourse acts in academic writing: The case of exemplification in linguistics research articles. *English for Specific Purposes, 63*, 120–133.

Wu, X., Mauranen, A., & Lei, L. (2020). Syntactic complexity in English as a lingua franca academic writing. *Journal of English for Academic Purposes, 43*, 100798.

Biographies

Marina Bondi is Professor of English Linguistics at the University of Modena and Reggio Emilia (Italy) and founding director of the CLAVIER centre (Corpus and Language Variation in English Research). Her research centres on textual, pragmatic, and phraseological aspects of academic and professional discourse across genres, discourse identities, and media.

Jessica Jane Nocella is currently a postdoctoral fellow at the Department of Studies on Language and Culture at the University of Modena and Reggio Emilia (Italy). She holds a PhD in human sciences from the University of Modena and Reggio Emilia with a thesis on evaluative language in the context of museums and Slow Art Day.