Rhona P. Lohan*

Universidad Complutense de Madrid, Spain
rlohan@ucm.es

ESIC University, Madrid, Spain
rhona.lohan@esic.university

DIVERGENT TRAJECTORIES IN STUDENT ACADEMIC BUSINESS WRITING: A COMPARATIVE LONGITUDINAL ANALYSIS OF LEXICO-PRAGMATIC RESOURCES IN EME AND NON-EME CONTEXTS

Abstract

This longitudinal study explored the differing development of lexico-pragmatic resources in student writing within English-medium education (EME) versus non-EME contexts. Conducted over two years at a private business university, the study compared an EME bachelor's cohort with an equivalent Spanish-medium group that received supplementary English language courses. Students' written discourse was analysed using pre- and post-tests (IELTS Writing Task 2 essays). Comparative measurements tracked the acquisition of semi-technical/academic vocabulary and textual cohesion, which was specifically measured through discourse markers. The findings indicate that while both groups showed growth, the nature of that development varied. EME students demonstrated greater gains in semi-technical/academic vocabulary, likely due to consistent exposure, yet showed minimal change in their use of discourse markers. Conversely, non-EME students, benefiting from explicit English for Specific and Academic Purposes (ESP/EAP) instruction, achieved stronger development in discourse markers, but exhibited limited growth in vocabulary. These results suggest that while EME facilitates vocabulary acquisition, discourse markers require targeted pedagogical intervention. The study advocates for integrated language support, confirming the need for a balanced curriculum combining incidental exposure with explicit instruction. This research offers practical implications for curriculum design in multilingual higher education settings.

Key words

English-medium education (EME), semi-technical vocabulary, discourse markers, ESP/EAP, IELTS writing, longitudinal comparative study.

* Corresponding address: Rhona P. Lohan, ESIC University, C. Valdenigrales, s/n. 28223 Pozuelo de Alarcón (Madrid), Spain.

160

1. INTRODUCTION

The increasing globalisation of higher education has positioned English as the dominant medium of instruction in diverse academic contexts worldwide. Such a shift directly affects second language (L2) learners, for whom developing academic vocabulary is essential for effective communication (Malmström et al., 2025; Nation, 2022). The present study adopts the term English-medium education in multilingual university settings (EMEMUS or EME for short), viewing it as a broad conceptualisation (Dafouz & Smit, 2020, 2023). This perspective emphasises a holistic focus on the students' educational experiences, encompassing curriculum, assessment, and the multilingual nature of the university setting. Although many key studies use the label English as a medium of instruction (EMI), this literature is incorporated into the EME perspective; however, the original term (EMI) is retained when directly referencing the source material to maintain fidelity to the published work. In EME settings, students are expected to engage with complex content through English, but this language input tends to be incidental and content-driven. This approach may not be sufficient for the systematic development of a comprehensive academic lexicon, particularly technical vocabulary or discourse markers essential for written argumentation (Reynolds & Skoufaki, 2024).

Moreover, the discipline-specific lexical demands placed on L2 learners in EME contexts are often underestimated, with limited attention given to how students acquire the specialised vocabulary needed to succeed in academic writing tasks (Tang & Curtis, 2025). Consequently, a comprehensive understanding of effective academic written discourse requires focusing on this lexico-pragmatic challenge, examining both the lexical aspect (technical vocabulary) and the pragmatic aspect (strategic use of discourse markers, or DMs). DMs are necessary for achieving textual cohesion and coherence across any text (Hyland, 2019). Their role has been well-established in the literature, particularly through their analysis as metadiscourse markers – linguistic devices that help organise discourse and guide readers through the rhetorical structure of a text (Dafouz, 2008; Hyland, 2019).

Lexical development is crucial for success in EME programmes (Wei & Macaro, 2024), as a well-developed vocabulary is key for understanding disciplinary literacies (Coxhead, 2024) and for effective participation in lectures and seminars (Lin & Morrison, 2010). The literature has examined various aspects of this challenge, revealing differences across contexts. Therefore, while some contexts show incidental lexical acquisition from content subjects (Reynolds et al., 2022), others offer explicit language instruction (McKinley & Rose, 2022; Webb & Nation, 2017), pointing to the effectiveness of EAP courses (Rose et al., 2020; Soruç et al., 2021). Moreover, the importance of ESP courses is acknowledged for their role in enhancing students' vocabulary and disciplinary language skills, which are essential for effective communication and participation in internationalised classrooms (Aguilar, 2018).

161

Academic business texts are scholarly writings focused on business-related topics, which entail the use of specific technical and academic vocabulary in a clear, organised manner. Effective text production, therefore, involves employing various lexico-pragmatic resources. Within the diversity of these resources, technical vocabulary refers to discipline-specific terms that are required for demonstrating subject knowledge (Coxhead & Nation, 2001), while discourse markers are linguistic devices used to establish textual cohesion and guide the reader through the organisation of a text (Hyland, 2019). General academic vocabulary, distinct from both, is used across all disciplines. Though academic texts comprise numerous linguistic features, the selection of technical vocabulary and discourse markers is motivated by their high degree of visibility. These lexical and pragmatic features are key to students' success in EME contexts, but they present distinct learning challenges and may require tailored strategies for effective acquisition (Nation, 2013). This study focuses on the semi-technical/academic vocabulary and discourse markers in student writing at the early stages of their academic studies.

While research has long highlighted the importance of academic vocabulary, longitudinal research on its development remains limited (Coxhead, 2024; Malmström et al., 2025; Sánchez-Pérez, 2023). Furthermore, a recognised gap exists in the study of technical vocabulary, with some researchers calling for more detailed investigation of subject-specific lexis in EME settings (Wei & Macaro, 2024). Building on this foundation, the present study is also motivated by the need for further research into the development of discourse markers, focusing specifically on how both of these underresearched areas manifest in EME contexts.

This study directly addresses this critical gap by reporting on a two-year longitudinal comparative analysis of semi-technical/academic vocabulary and discourse markers in student writing at a private business university. Holistic writing scores from the same dataset were previously used to evaluate overall writing proficiency (Lohan & Marín Palacios, in press); here, the focus shifts to lexico-pragmatic development, tracked through density and diversity measures over time. The comparative methodology involves two carefully selected cohorts: one enrolled in an EME bachelor's degree and one in an equivalent Spanish-medium bachelor's degree that includes supplementary English language courses. By including the non-EME group, this study provides an opportunity to compare the two cohorts and explore possible differences in their academic writing development in their different language-medium programmes. The two-year longitudinal design is a valuable means of capturing developmental trajectories rather than a mere snapshot, offering a more in-depth understanding of how specific features develop over time.

This paper begins with a review of the relevant literature, followed by an outline of the methodology. The results of the data analysis are then presented and discussed in relation to the research questions, with a focus on the implications for pedagogical approaches in both settings. The paper concludes by summarising the study's contributions and suggesting avenues for future research.

162

2. LITERATURE REVIEW

This review will examine the current literature on academic written language in EME settings, focusing on technical vocabulary and discourse markers. It will explore their acquisition in various learning contexts (EME and non-EME), the pedagogical approaches used to support their development, and finally, the use of IELTS Academic Writing Task 2 as an assessment tool.

2.1. Academic vocabulary and academic success

Academic vocabulary plays a vital role in the L2 learners' academic success, particularly in EME contexts. Recent studies found that a well-developed vocabulary is crucial for effective academic communication and serves as a significant predictor of academic success (Malmström et al., 2025; Masrai & Milton, 2021). For instance, students with larger vocabularies perform better in EMI courses (Masrai & Milton, 2021). A longitudinal study on Master's students reported significant gains in receptive academic vocabulary in EMI programmes; however, it also highlighted variation in students' initial lexical knowledge, suggesting many learners begin EMI with potential deficiencies (Malmström et al., 2025). Nonetheless, existing developmental research often reports findings limited either in duration or in the linguistic features tracked. For instance, a short-term longitudinal study observed lexical gains over one semester but found minimal progress in broader language skills (e.g., syntax, fluency) (Sánchez-Pérez, 2023). This restricted length of observation is the basis for the call for more detailed and extended longitudinal investigation in EME contexts (Coxhead, 2024; Sánchez-Pérez, 2023).

It has been noted that "little can be conveyed without grammar, yet without vocabulary, nothing can be conveyed" (Nation, 2022, p. 1). A qualitative understanding of vocabulary is necessary for academic success (Laufer & Nation, 1999), as lexical gaps increase cognitive load, impacting deep understanding. It is also necessary for active participation in academic settings such as lectures and seminars (Lin & Morrison, 2010). Therefore, academic vocabulary acquisition is essential for academic success in EMI contexts, where students simultaneously process new subject content and complex English (Nation, 2022; Wei & Macaro, 2024). A well-developed academic lexicon is important for effective written communication, enabling precise expression and argumentation (Coxhead, 2024).

2.2. The core linguistic components

Academic written discourse requires a command of specific linguistic components. Technical vocabulary consists of discipline-specific terminology that conveys subject knowledge (Mudraya, 2006; Othman, 2024; Wei & Macaro, 2024). These

163

terms often represent abstract or complex ideas, making them more difficult to grasp (Hellekjær, 2010; Lin & Morrison, 2010), and may not be encountered frequently enough in students' learning environments to ensure long-term retention (Evans & Morrison, 2011; Rose et al., 2020). Students may also be confused when familiar words take on specialised meanings in academic contexts (Mudraya, 2006).

Discourse markers are words or phrases used to organise speech and writing, guiding the listener or reader through the discourse (Malmström et al., 2025; Masrai & Milton, 2021; Nation, 2013, 2022). They help structure academic texts and support coherence (Biber & Conrad, 2009; Hyland, 2019). While they may seem simple, their functions are nuanced and context-dependent, presenting a challenge for learners who must master their pragmatic and rhetorical roles. These resources have been classified as metadiscourse markers, comprising two main types: textual markers, which manage discourse flow by connecting ideas and ensuring coherence (e.g., logical connectors like *therefore* and sequencers like *first*), and interpersonal markers (e.g., certainty markers like *clearly* and attitude markers like *unfortunately*), which reflect the writer's stance and engage the reader (Dafouz, 2008, pp. 97, 99). Consequently, lower proficiency EME learners may struggle to comprehend their function in structuring arguments, resorting to strategies like direct translation, which can inhibit a deeper understanding (Rose et al., 2020).

164

2.3. Lexical and pragmatic development and pedagogical approaches

Lexical and pragmatic development is shaped by the learning environment. In EME settings, students engage with disciplinary content delivered in English, which can support incidental vocabulary development. However, this content-driven approach often lacks a systematic focus on language instruction, limiting opportunities to acquire a full range of academic vocabulary needed for deeper comprehension and production (Malmström et al., 2025; Othman, 2024). The dual challenge of mastering both disciplinary content and the specific language used to express it is well documented (Dafouz, 2018; Kuteeva & Airey, 2014). In contrast, non-EME settings typically offer more structured English courses, such as ESP/EAP, but provide fewer opportunities for authentic engagement with academic discourse. In these contexts, students study disciplinary knowledge in their native language, while English is taught as a separate subject (McKinley & Rose, 2022).

As a result of these contextual factors, lexical development in EME and non-EME contexts unfolds differently. While implicit acquisition, or incidental learning through extensive exposure, occurs naturally in EME, it is often insufficient for mastering the precise academic and technical vocabulary required for university study (Reynolds et al., 2022). This is because the specialised nature of the vocabulary, alongside limited repetition in content-focused lectures, makes it difficult for long-term retention.

Research demonstrates that strategies like pre-emptive instruction from content teachers and the use of translanguaging to explain complex terms (Pun & Tai, 2021) are effective in these settings.

Given the inadequacy of relying on incidental learning alone, EME students require targeted support to build an academic lexicon. By contrast, EAP/ESP courses provide the necessary systematic instruction for lexical development (Webb & Nation, 2017). They are designed to teach not only the high-frequency general academic vocabulary (Coxhead, 2024), but also specialised, technical terms specific to a particular field (Mudraya, 2006; Wei & Macaro, 2024). For EME students with lower English proficiency, such instruction can be impactful, helping them bridge the gap between general language competence and the specialised lexical demands of their disciplines (Kamaşak & Sahan, 2023; Soruç et al., 2021). Research has shown that integrating language support into EME programmes enhances students' preparedness for academic tasks and fosters more effective engagement with disciplinary content (Costa & Mastellotto, 2022; Rose et al., 2020).

2.4. IELTS writing task 2 as an assessment tool

The IELTS exam, developed by the British Council, International Development Program of Australian Universities and Colleges (IDP), and Cambridge Assessment, is a globally recognised assessment tool for evaluating English proficiency, with results accepted by over 12,000 institutions worldwide (IELTS, 2023). In EMI research, it has been used to examine both students' academic readiness and their language development trajectories (Rogier, 2012; Rose et al., 2020; Schoepp, 2018) and to identify learners' problems in writing (Nguyen & Nguyen, 2022). The Writing Task 2 component, central to this study, has faced criticism for its limited authenticity and reliance on opinion-based responses (Uysal, 2010). Nonetheless, its structured rubric – covering task response, coherence and cohesion, lexical resource, and grammatical range – offers a consistent framework for evaluating writing proficiency.

The Lexical Resource (LR) and Coherence and Cohesion (CC) criteria are particularly relevant for studies of lexical and structural development, as they directly assess a candidate's capacity for lexical sophistication and logical argumentation. The Lexical Resource criterion is defined by the rubric as the candidate's ability to use a wide range of vocabulary with flexibility and precision. This includes the appropriate use of less common and topic-specific vocabulary, accurate word choice, and effective paraphrasing (IELTS, 2023). Meanwhile, the Coherence and Cohesion criterion assesses the organisation of information and ideas, focusing on clear paragraphing, logical sequencing, and the appropriate use of cohesive devices such as discourse markers. In this study, equal attention is given to the technical vocabulary (as the core of LR's 'topic-specific' requirement) and the use of discourse markers (as a key component of CC's 'cohesive devices'). This dual focus provides a comprehensive measure of students' ability to convey discipline-appropriate content and to structure

that content coherently within academic written discourse arguments in EME contexts.

While previous studies have explored lexical development in EME and non-EME contexts, few have examined how these environments influence changes in specific lexical features over time; this study addresses that gap through the following research questions:

RQ1: To what extent does the type of instruction (EME versus non-EME) influence the development of semi-technical/academic vocabulary usage, as observed between pre- and post-test academic writing samples over a two-year period?

RQ2: What is the impact of the type of instruction (EME versus non-EME) on the developmental trajectory of discourse marker usage (frequency, variety, and appropriateness), as observed between pre- and post-test academic writing samples over a two-year period?

3. METHODOLOGY

3.1. Participants and contexts

This study investigates the lexical and pragmatic development in EME and non-EME student groups using pre- and post-IELTS writing tests over two years. The data were collected at a private business university that takes a distinctive approach to language support, as it fully integrates EAP courses into its curriculum; each bachelor's degree includes a single mandatory EAP course as a core requirement. The university also offers a compulsory supplementary diploma programme that runs concurrently with all bachelor's degrees, providing students with additional practical subjects aligned with current market trends.

The participant pool consisted of 86 students who met the minimum B2 English proficiency level (CEFR) upon university entry. Two cohorts were identified for comparison: students in EME degrees in Marketing and Business Administration (n=52), and their non-EME counterparts in equivalent Spanish-taught programmes (n=34). All participants, regardless of degree stream, undertook the mandatory EAP course in their second year of study, which accounted for 6 ECTS credits. For non-EME students, their language provision extended beyond this EAP course; they received an additional 180 hours of ESP/EAP instruction over two years as part of their complementary diploma programmes. EME students, conversely, were immersed in English-taught content courses. Moreover, a situational analysis of this context (Lohan & Dafouz, 2024) discovered that the EME cohort encountered business vocabulary through both content-based instruction and curriculum-integrated masterclasses. These sessions, delivered during class time by invited professionals, focused on real-world case studies and likely exposed students to technical terminology relevant to their disciplinary fields.

166

Data were collected as part of a departmental initiative designed to assess academic writing development for internal quality assurance. While student numbers and degree programmes were retained for tracking individual development, all analyses in this research reference participants solely by their degree stream (EME or non-EME) and no personal identifiers were collected. The study was conducted at the programme level. Given the internal origin of the data and their robust anonymisation, ethical clearance was deemed appropriate.

3.2. Writing task and data collection

Participants completed the same academic writing task at two points: at the beginning of their first year (September 2022) and at the end of their second year (May 2024). The task was modelled on IELTS Academic Writing Task 2 and administered under timed classroom conditions (30 minutes), with all responses being handwritten. The prompt was:

> Recent research has shown that business meetings and training are increasingly taking place online.
> What are the advantages and disadvantages of this?

This prompt was selected for its relevance to students' disciplinary content and its potential to elicit structured argumentative writing. Importantly, the identical prompt was administered twice – at the beginning of the first academic year and again at the end of the second academic year. As these students were at the start of a four-year bachelor's programme when the initial test was administered, the prompt was designed to require foundational knowledge of business principles. This ensures the prompt is appropriate for first-year students with limited disciplinary expertise, yet it still encourages the use of formal vocabulary, cohesive devices, and technical vocabulary, making it suitable for the required lexico-pragmatic analysis across the two years of data collection. The use of the identical prompt at Time 1 and Time 2 is a necessary methodological choice to establish construct equivalence (Chapelle & Lee, 2021; Messick, 1989). Maintaining task invariance – consistency in topic, rhetorical purpose and audience – minimises the risk of conflating developmental change with task-related variation, thereby supporting valid longitudinal comparison (Bachmann & Palmer, 2010). This consistency is important for lexico-pragmatic analysis, as it ensures that both data sets are situated within a stable communicative context. Although concerns regarding testing effects are acknowledged, the two-year interval between administrations is considered sufficient to minimise the likelihood of recall.

Students were informed that the task would not affect their grades, and it was intended to support curriculum development and evaluate writing progress. This same dataset was used in a previous study focusing on the holistic proficiency scores

gained (Lohan & Marín Palacios, in press); the current study isolates technical vocabulary and discourse markers to explore developmental patterns in their use and cohesion.

3.3. Data analysis

Participants' IELTS Writing Task 2 essays, collected at two time points (pre- and post-intervention), were manually scored for three key lexical measures: Semi-Technical/Academic Vocabulary Density (Technical Vocabulary Density or TVD for short), Discourse Marker Total Density (DMTD), and Discourse Marker Type Density (DMTyD).

The dataset comprised a total of 172 essays, with each participant contributing one pre- and one post-intervention essay. The EME group produced 104 texts (52 participants x 2), with a combined word count of 13,181 words at pre-test and 11,350 words at post-test. The non-EME group contributed 68 texts (34 participants x 2), totalling 6,750 words at pre-test and 7,321 words at post-test. These figures provide a contextual insight into the size and scope of the two subcorpora analysed, with exact counts serving as the basis for calculating lexical density measures. The calculation for each lexical density measure was performed using the formula:

Density = Number of items / Total words in the essay

where "items" refers to the specific lexical feature being measured. For example, if an essay contained 4 semi-technical/academic vocabulary items and had a total of 229 words, the TVD was calculated as:

$4/229 = 0.0175$.

This value was rounded to 0.02, or 2% for reporting purposes in SPSS. Density scores were expressed as raw proportions rather than per 100 words. Group-level mean scores were calculated by averaging individual density scores at each time point. These means formed the basis for all statistical analyses.

The lexical resources searched for in this study were compiled from a combination of established academic sources. The semi-technical/academic vocabulary was derived primarily from Coxhead's (2000) Academic Word List (AWL). The AWL was selected because it represents the high-frequency, cross-disciplinary vocabulary that is foundational for students entering academic study, as discussed in the data collection section. To maintain measurement consistency and transparency across the study period, the exact same list was applied to the student writing collected at both points in time. To ensure the list's relevance to the discipline, the AWL foundation was supplemented with additional topic-relevant terms included to reflect the thematic focus of the writing task prompt. A list of the semi-technical vocabulary used is available in Appendix A. The reduced search scope will be addressed in the limitations section.

In turn, a pre-selected list of specific discourse markers was systematically categorised by textual function (listing, exemplification, contrast/concession, cause/effect, and conclusion). This categorisation was loosely adapted from Dafouz (2008, p. 97) and subsequently established using published IELTS writing conventions (British Council, 2025) and core academic writing principles (University of Warwick, 2020). This system enabled functional analysis. It moved the study beyond simple frequency counts to examine how participants structured their arguments and expressed logical relationships between ideas. By classifying each marker by its textual purpose, changes in usage accuracy and appropriateness were tracked across the pre- and post-test essays (see subsection 4.3.). Such categorisation facilitated a clear comparison between EME and non-EME participants. However, this selective approach meant the analysis did not include other metadiscoursal features, such as interpersonal markers or hedges. The complete classification of discourse markers by type is presented in Appendix B.

To ensure consistency and transparency in the lexical scoring process, two independent raters, all PhD students in Applied Linguistics and Education, participated in verifying the lexical counts. These raters were blind to group and time point during this stage. Each reviewed a subset of essays to verify the author's manual counts, without access to the original results or knowledge of whether the essays were from the pre- or post-test phases. The author conducted the initial counts, and the verification process, measured by Percentage Agreement, confirmed a 98% consistency rate across reviewed samples. No substantive discrepancies in the final feature counts were identified.

One-way ANOVAs were conducted on pre-test scores for each lexical measure to ensure comparability between groups at baseline. This step assessed whether any statistically significant differences existed between the EME and non-EME cohorts prior to the intervention. To evaluate development over time and between groups, statistical analyses were conducted in two stages. First, post-intervention comparisons were made using one-way ANOVAs for each lexical measure (Semi-Technical/Academic Vocabulary Density, Discourse Marker Total Density, and Discourse Marker Type Density) to identify any significant differences between the EME and non-EME cohorts. Second, a series of repeated measures ANOVAs was performed for each dependent variable, using a 2 (Group: EME, non-EME) x 2 (Time: Pre-intervention, Post-intervention) design. This longitudinal approach enabled the assessment of within-group changes over time and the interaction effects between groups and time points. All analyses were conducted using SPSS with significance set at *p*<.05.

4. RESULTS

This section presents the findings from the statistical analysis conducted on the participants' IELTS Writing Task 2 (pre- and post-tests). Results are organised to sequentially answer the research questions, following the methodology described.

4.1. Preliminary findings

To establish baseline comparability of the two groups, EME and non-EME participants, the descriptive statistics for the pre-tests of the three variables (TVD, DMTD, and DMTyD) are shown in Table 1.

Descriptives		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min.	Max.
						Lower Bound	Upper Bound		
PRETechVocab Density	Non-EME	34	.006702	.0080953	.0013883	.003877	.009526	.0000	.0364
	EME	52	.005822	.0046757	.0006484	.004520	.007124	.0000	.0222
	Total	86	.006170	.0062247	.0006712	.004835	.007504	.0000	.0364
PREDMTotal Density	Non-EME	34	.019532	.0106832	.0018322	.015804	.023259	.0000	.0417
	EME	52	.019432	.0102852	.0014263	.016569	.022296	.0000	.0500
	Total	86	.019472	.0103819	.0011195	.017246	.021697	.0000	.0500
PREDMType Density	Non-EME	34	.012531	.0011641	.0011641	.010162	.014899	.0000	.0281
	EME	52	.012808	.0008825	.0008825	.011037	.014580	.0000	.0333
	Total	86	.012699	.0007005	.0007005	.011306	.014091	.0000	.0333

Table 1. Descriptive statistics for pre-test lexical measures by group

As shown in Table 1, the mean scores for the EME group (Technical Vocabulary: *M* =.006, *SD* =.005; Discourse Marker Total: *M* =.019, *SD* =.010; Discourse Marker Type: *M* =.013, *SD* =.009) and the non-EME group (Technical Vocabulary: *M* =.007, *SD* =.008; Discourse Marker Total: *M* =.019, *SD* =.011; Discourse Marker Type: *M* =.012, *SD* =.001) were notably similar across the three pre-test measures. These means were calculated by averaging individual lexical density scores within each group, as described in subsection 3.3.

One-way ANOVAs were then conducted to formally assess the baseline comparability. The results are shown in Table 2.

For Technical Vocabulary Density, the difference between groups was not statistically significant, F = .408, p = .525. Similarly, Discourse Marker Total Density showed no significant group difference, F = .002, p = .966. For the Discourse Marker Type Density, no significant group differences were observed, F = 0.037, p = 0.848. These preliminary analyses confirm that both the EME and non-EME cohorts began the study at comparable levels of lexical proficiency, hence suitable for the main analysis.

		Sum of Squares	Df	Mean Square	F	Sig.
PRETechVocab Density	Between Groups	.000	1	.000	.408	**.525**
	Within Groups	.003	84	.000		
	Total	.003	85			

PREDMTotal Density	Between Groups	.000	1	.000	.002	**.966**
	Within Groups	.009	84	.000		
	Total	.009	85			
PREDMType Density	Between Groups	.000	1	.000	.037	**.848**
	Within Groups	.004	84	.000		
	Total	.004	85			

Table 2. Summary of one-way ANOVAs for pre-test lexical measures by group

As detailed in the Methodology section, after two academic years of instruction in their respective English or Spanish-medium programmes, participants completed a second writing assessment (Test 2), an IELTS Writing Part 2 task. The same prompt was used for both assessments to maintain consistency in task demands and to enable a focused analysis of lexico-pragmatic development. The following section presents the findings of this post-intervention assessment.

4.2. Addressing RQ1: Development of lexical features (semi-technical/academic vocabulary density)

RQ1 examined the influence of instructional practice on the development of technical vocabulary density (TVD). To determine whether participants' lexical development changed significantly from pre- to post-intervention and whether this change differed between EME and non-EME groups, a 2 (Group: EME, non-EME) x 2 (Time: Pre-intervention, Post-intervention) repeated measures ANOVA was conducted. Assumption checks were performed before analysis. Normality was evaluated using the Shapiro-Wilk test, which indicated a violation for TVD (W = .806, p < .001), although other lexical measures met the assumption. Sphericity was not violated (Mauchly's W = 1.000), and Greenhouse-Geisser and Huynh-Feldt corrections were reported for completeness. Homogeneity of variances was assumed based on consistent variance estimates and group sizes. Despite the normality violation, the repeated-measures ANOVA was retained due to its robustness to moderate deviations in samples of this size. To validate the parametric findings, a Wilcoxon signed-rank test was also performed, confirming a significant increase in TVD from pre- to post-intervention (z = 5.247, p < .001). The results are presented in Table 3.

Measure	Effect	*F*-value	Df	*p*-value
Technical Vocabulary Density (TVD)	Group	0.0302	(1.85)	0.8625
	Time	332.531	(1.85)	< .001
	Group x Time	123.644	(1.85)	0.0007

Table 3. Summary of 2x2 repeated measures ANOVA results for technical vocabulary density (TVD)
Note: Degrees of freedom for all tests were (1,85).

Table 3 reveals a highly significant Group x Time interaction for TVD ($F = 123.644$, $p < 0.001$). This interaction demonstrates that the developmental trajectory for TVD differed between the EME and non-EME groups. An overall increase in TVD was observed across all participants ($F = 332.531$, $p < 0.001$), but the interaction showed that the EME group displayed a more substantial gain in TVD compared to the non-EME group.

Furthermore, the post-test scores confirmed this outcome. Given the significant Group x Time interaction, follow-up simple effects analyses were conducted to examine group differences at the post-intervention stage. Table 4 provides the descriptive statistics for the post-intervention assessment, and Table 5 summarises the post-test one-way ANOVAs.

Descriptives		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min.	Max.
						Lower Bound	Upper Bound		
POSTTechVocab Density	Non-EME	34	.008382	.0080377	.0013784	.005578	.011187	.0000	.0314
	EME	52	.016027	.0097905	.0013577	.013301	.018753	.0000	.0494
	Total	86	.013005	.0098351	.0010605	.010896	.015113	.0000	.0494
POSTDMTotal Density	Non-EME	34	.027719	.0111614	.0019142	.023825	.031614	.0000	.0498
	EME	52	.018128	.0091267	.0012656	.015587	.020669	.0084	.0357
	Total	86	.021920	.0109815	.0011842	.019565	.024274	.0000	.0498
POSTDMType Density	Non-EME	34	.015804	.0054007	.0009262	.013920	.017689	.0084	.0278
	EME	52	.012305	.0063495	.0008805	.010537	.014072	.0000	.0286
	Total	86	.013688	.0062029	.0006689	.012358	.015018	.0000	.0286

Table 4. Descriptive statistics for post-test lexical measures by group

		Sum of Squares	Df	Mean Square	F	Sig.
PRETechVocab Density	Between Groups	.001	1	.001	14.375	**<.001**
	Within Groups	.007	84	.000		
	Total	.008	85			
PREDMTotal Density	Between Groups	.002	1	.002	19.006	**<.001**
	Within Groups	.008	84	.000		
	Total	.010	85			
PREDMType Density	Between Groups	.000	1	.000	7.007	**.010**
	Within Groups	.003	84	.000		
	Total	.003	85			

Table 5. Summary of one-way ANOVAs for post-test lexical measures by group

As shown in Table 5, a statistically significant difference was found for TVD ($F = 14.375$, $p < .001$). Examination of the descriptive statistics (Table 4) indicated that the EME group ($M = 0.016$) had a significantly higher technical vocabulary density compared to the non-EME group ($M = 0.008$) at post-intervention.

To test the practical significance and magnitude of these observed differences, effect sizes (Eta-squared, Epsilon-squared, and Omega-squared) and their 95% confidence intervals were calculated and are shown in Table 6.

Measure		Point Estimate	95% Confidence Interval	
			Lower	Upper
POSTTechVocab Density	Eta-squared	.146	0.034	0.283
	Epsilon-squared	.136	0.022	0.274
	Omega-squared Fixed-effect	.135	0.022	0.272
	Omega-squared Random-effect	.135	0.022	0.272
POSTDMTotal Density	Eta-squared	.185	0.057	0.324
	Epsilon-squared	.175	0.046	0.316
	Omega-squared Fixed-effect	.173	0.045	0.313
	Omega-squared Random-effect	.173	0.045	0.313
POSTDMType Density	Eta-squared	.077	0.005	0.200
	Epsilon-squared	.066	-.007	0.191
	Omega-squared Fixed-effect	.065	-.007	0.189
	Omega-squared Random-effect	.065	-.007	0.189

Table 6. Effect sizes for one-way ANOVAs (EME vs. non-EME group)

The Eta-squared value of 0.146 for TVD suggests a large effect, implying a substantial proportion of variance in technical vocabulary use is explained by group membership at the post-intervention stage.

4.3. Addressing RQ2: Development of pragmatic features (discourse marker usage)

RQ2 investigated the impact of instructional practice on the developmental trajectory of discourse marker usage, measured by total density (DMTD) and type density (DMTyD). The key findings originate from the 2 x 2 repeated measures ANOVA, which are presented in Table 7.

Measure	Effect	*F*-value	df	*p*-value
Discourse Marker Total Density (DMTD)	Group	38.791	(1.85)	0.0522
	Time	40.889	(1.85)	0.0463
	Group x Time	146.942	(1.85)	0.0002
Discourse Marker Type Density (DMTyD)	Group	0.2225	(1.85)	0.6384
	Time	1.557	(1.85)	0.2156
	Group x Time	54.217	(1.85)	0.0223

Table 7. Summary of 2x2 repeated measures ANOVA results for discourse marker total density and discourse marker type density measures

Note: Degrees of freedom for all tests were (1,85).
Note: Assumption checks confirmed that both DMTD and DMTyD met the requirements for repeated-measures ANOVA. The Shapiro-Wilk test indicated normality for both measures (W = .986 and W = .980, respectively), and Mauchly's test showed no violation of sphericity (W = 1.000). Greenhouse-Geisser and Huynh-Feldt corrections were reported but not required.

As shown in Table 7, highly significant Group x Time interactions were found for both DMTD (F = 146.942, $p < 0.0002$) and DMTyD (F = 54.217, $p < 0.0223$). These

interactions indicate that the developmental gains were concentrated in the non-EME group. The Group x Time interaction for DMTD showed that the non-EME group experienced a notable increase in total density of discourse markers, while the EME group increased only slightly. Similarly, the interaction for DMTyD revealed that the non-EME group demonstrated an increase in the variety of discourse markers used, whereas the EME group showed a slight decrease over time. These contrasting trajectories are illustrated in Figure 1.
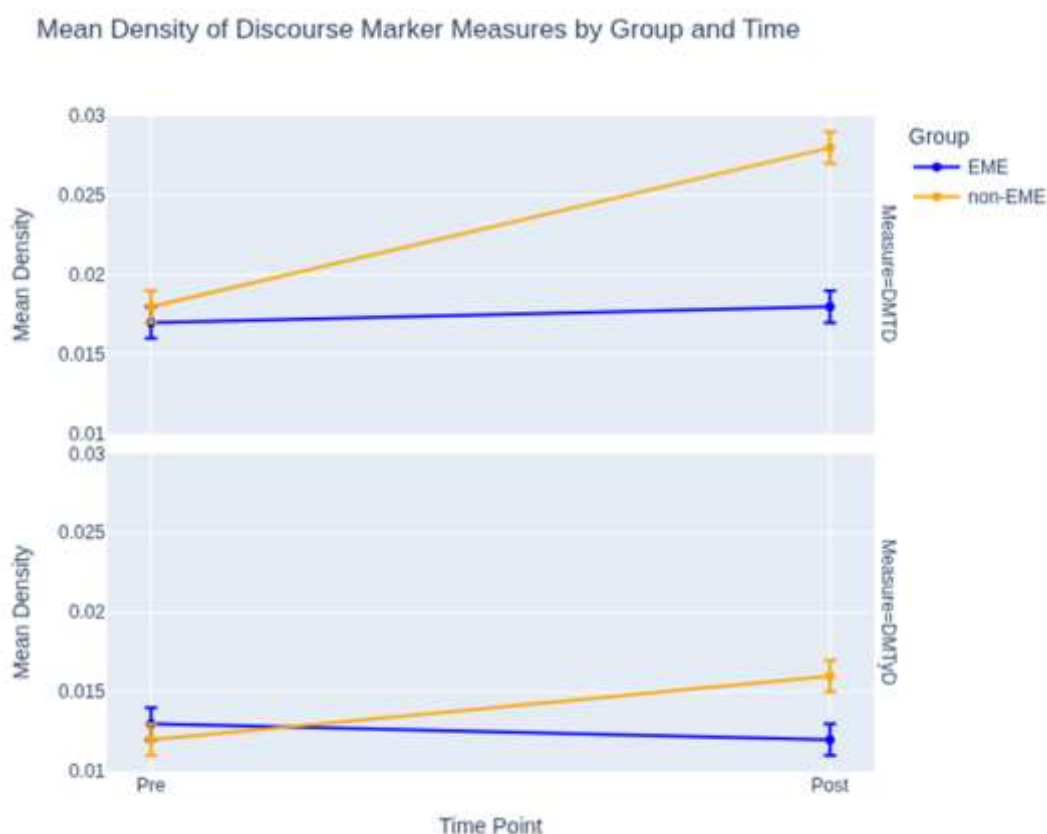
Figure 1. Mean density of discourse marker total density (DMTD) and discourse marker type density (DMTyD) measures by group and time
Note: This figure illustrates the mean density of DMTD and DMTyD at pre-and post-intervention for the EME group (English – blue line) and the non-EME group (orange line). Error bars represent standard errors of the mean.

As illustrated in Figure 1, a general increase in DMDT was found from pre- to post-intervention ($F = 40.889$, $p = .0463$). The observed interaction ($F = 146.942$, $p < .0002$) suggested that the non-EME group (orange line) exhibited a notable increase in DMTD, while the EME group (blue line) increased slightly. For DMTyD, neither the main effect of Group ($F = 0.2225$, $p = 0.6384$) nor the main effect of Time ($F = 1.557$, $p = 0.2156$) reached statistical thresholds. However, a clear Group x Time interaction was found ($F = 54.217$, $p = .0223$). In this case, the non-EME group

(orange line) showed an increase in DMTyD, whereas the EME group (blue line) experienced a slight decrease over time.

The post-test scores further confirmed these outcomes. As detailed in the post-test ANOVA results (Table 5, presented in subsection 4.2.), a significant difference for DMTD was observed ($F = 19.006$, $p < 0.001$), with the non-EME group ($M = 0.028$) outperforming the EME group ($M = 0.018$), as shown in Table 4, subsection 4.2. Similarly, DMTyD exhibited a significant difference ($F = 7.007$, $p = 0.010$) with the non-EME group displaying a higher mean density of discourse marker types ($M = 0.016$) than the EME group ($M = 0.012$), also documented in Table 4. The effect sizes, summarised in Table 6, provide context on the magnitude of these observed differences. The Eta-squared values for DMTD (0.185) and DMTyD (0.077) imply large and medium effects, respectively.

Overall, the results confirm distinct developmental trajectories for lexical and pragmatic features across the EME and non-EME groups, reflecting varied academic discourse development over the two-year period. These findings will be discussed in the next section.

5. DISCUSSION OF RESULTS

This study sets out to explore how instructional practice shapes students' academic written discourse. With both groups starting at comparable levels (see subsection 4.1. Table 2), the findings offer insights into the role of the language medium of education and feature-specific development in academic language proficiency. This discussion section will interpret the results in relation to the two primary research questions in a complementary manner, focusing on the differences across settings.

175

5.1. Differential development: Evidence from incidental (EME) and explicit acquisition (non-EME)

In response to the research questions, the results demonstrate that instructional settings influence academic written discourse in distinct ways. These differences, as evidenced by the significant Group x Time interactions (Tables 3 and 7), reflect how varying pedagogical approaches and language exposure conditions contribute to the development of specific lexical and pragmatic features. The divergent developmental patterns observed across the two groups are central to understanding academic language acquisition in this context, influenced by differing instructional practices. The results reveal that while both groups demonstrated growth, the nature of that development varied unevenly across lexical and pragmatic features.

The development of technical vocabulary was not uniform: while both groups experienced an increase over time, the EME cohort displayed substantial gains

compared to the non-EME. This suggests that this type of vocabulary, despite the noted difficulty of technical terms (Hellekjær, 2010; Lin & Morrison, 2010), is sensitive to consistent exposure present in an immersion environment. The EME students' greater gains in technical vocabulary, as shown in Table 3, subsection 4.2., are likely due to the cumulative exposure to discipline-specific language embedded in their coursework, which includes a key component of master classes with leading professionals (Lohan & Dafouz, 2024; see Methodology for full context). This pedagogical approach ensures regular exposure to authentic specialised discourse, which supports the idea that immersion facilitates the incidental acquisition of semi-specialised vocabulary, particularly when students are consistently engaging with content in English. This finding aligns with recent research that views incidental acquisition as a byproduct of EME settings (Reynolds et al., 2022). In these environments, students engage with specialised terminology across various academic tasks, which enables them to acquire discipline-specific lexis more effectively. Furthermore, the present study's longitudinal data, showing gains over the two-year period, confirm previous findings (Coxhead & Boutorwick, 2018; Sánchez-Pérez, 2023) that demonstrate EME students make progress in the area of technical vocabulary over time.

On the other hand, the development of discourse markers was notably more variable. The non-EME group experienced an increase in both total (DMTD) and type (DMTyD) density of discourse markers, which is consistent with the effectiveness of systematic instruction (Webb & Nation, 2017). However, the EME group showed only a slight increase in total markers and an actual slight decrease in marker types over the two-year period (Table 7, subsection 4.3.). This stagnation in discourse marker development echoes the findings from a previous study on the same data (Lohan & Marín Palacios, in press), which noted that a large majority of the EME students' holistic IELTS writing scores also stagnated at band 7.0. This lack of progress in the use of discourse markers – a key component of the Coherence and Cohesion criterion on the IELTS writing rubric (IELTS, 2023) and essential for constructing coherent arguments – corroborates findings that attribute stagnation to repeated written language practice without sufficient feedback or targeted instruction (Sánchez-Pérez, 2023; Wilkinson**,** 2013).

This development in pragmatic features provided evidence for the effectiveness of explicit instruction. The non-EME group's stronger performance in discourse marker usage (Figure 1, subsection 4.3.), both in terms of frequency and appropriateness, strongly supports this conclusion. The most frequent and notable development was the correct pairing of "*On the one hand*" and "*On the other hand*", which replaced earlier incorrect forms such as "*on one side*". This shift reflects a stronger grasp of contrastive structure and academic convention. Participants also showed increased control over listing markers, including "*Moreover*", "*Firstly*", and "*Secondly*", which helped organise their ideas more clearly and logically. These markers contributed to better paragraph cohesion and sequencing. In terms of exemplification, phrases like "*For example*" and "*Such as*" were used to support

claims with relevant evidence, enhancing clarity and persuasiveness. Contrast and concession markers such as "*However*", "*While*" and "*Although*" were used more accurately, allowing students to present balanced arguments and acknowledge alternative perspectives. Cause and effect relationships were expressed more effectively through markers like "*Therefore*" and "*Due to*", indicating improved ability to link ideas logically. Finally, conclusion markers such as "*In conclusion*", "*To sum up*", and "*Overall*" were used to signal closure and reinforce key points, contributing to more cohesive and purposeful essay endings.

These examples illustrate how targeted instruction in ESP/EAP not only increases the frequency of discourse marker usage but also enhances their functional accuracy. This development aligns with the structured learning environment described in the methodology section, where the non-EME group accumulated 240 hours of targeted ESP/EAP instruction over the two years. This structured learning environment, with its exercise in academic writing and discourse, aligns with research highlighting the effectiveness of explicit attention to genre, register, and discourse conventions (Hyland, 2019; Roothooft et al., 2025). This provides evidence that helps to explain the earlier finding (Lohan & Marín Palacios, in press) that these students achieved notable improvements in their holistic writing scores. Conversely, the EME group, having only received one 60-hour EAP course, showed little to no growth in this area. This disparity supports the argument that while lexical features may be acquired incidentally in EME contexts, cohesive devices such as discourse markers are less likely to be acquired in passing and therefore benefit from targeted teaching.

177

These distinct patterns of growth, confirmed by the statistically significant Group x Time interactions in Table 3 and Table 7, indicate that lexical and pragmatic development are not a monolithic process. Instead, different features appear to be sensitive to different instructional conditions. These findings reinforce the idea that while content-based instruction may promote gains in technical vocabulary, it may not be sufficient for developing other key aspects of academic written language, such as cohesive and rhetorical functions of discourse markers.

5.2. Pedagogical implications

The findings have pedagogical implications, suggesting a need for a more balanced approach to academic written discourse development in both EME and non-EME contexts.

For EME programmes, the results highlight that while incidental learning effectively develops semi-technical/academic vocabulary, it is insufficient for developing more subtle, discourse-level features. EME programmes may therefore benefit from integrating targeted language support that explicitly focuses on discourse markers or cohesion. This would address the lack of systematic language development that is often absent in immersion settings (Malmström et al., 2025;

Othman, 2024). Integrating such language support, as research has shown, can significantly enhance students' preparedness for academic tasks and ensure more effective engagement with disciplinary content (Costa & Mastellotto, 2022; Rose et al., 2020).

Conversely, for non-EME settings, the findings demonstrate the benefits of systematic instruction in developing discourse markers but also point to a potential gap in their training. Given that these students acquire discipline-specific knowledge in their first language and receive less authentic academic exposure in English, ESP courses could be improved by more intentionally incorporating discipline-specific vocabulary and providing learners with the relevant senses of academic terms (Huang & Chiang, 2025). While EAP courses typically focus on general academic vocabulary (Coxhead, 2024), incorporating specialised, technical terms (Mudraya, 2006; Wei & Macaro, 2024) would better prepare students to meet the specific lexical demands of their fields.

6. CONCLUSION

Conducted at a single private business university, this two-year longitudinal study set out to provide a direct comparison of lexico-pragmatic development in students from EME and non-EME contexts, revealing that the instructional practices have a distinct impact on both lexical and pragmatic development. The study demonstrates the distinct patterns of language gains for EME and non-EME participants, highlighting areas of strength and areas needing improvement in each cohort's development. The findings from this research show that while both groups experienced growth, the nature of that development varied. Thus, while the EME students showed greater gains in semi-technical/academic vocabulary, likely due to consistent exposure and incidental acquisition in their business content courses, there was stagnation in the use of discourse markers. The non-EME group, on the other hand, exhibited stronger performance in the use of discourse markers, which could be a benefit of their extensive, explicit instruction in EAP/ESP courses, but showed little to no improvement in their technical vocabulary. These results illustrate that while technical vocabulary is sensitive to immersion, discourse markers are less likely to be acquired incidentally and benefit from targeted instruction.

Overall, these results underscore that no single instructional approach is sufficient. The pedagogical implications of this study suggest a dual focus may be beneficial: EME settings could integrate targeted language support to promote the development of necessary rhetorical tools, while non-EME programmes would benefit from enriching EAP/ESP courses by incorporating more of the specialised, technical vocabulary required by their respective disciplines. This tailored approach offers a promising way to enhance student preparedness for the lexical and rhetorical demands of advanced academic writing.

While the study offers insights into lexical and pragmatic development over the first two years of undergraduate study, limitations should be acknowledged. The research was conducted within a single institutional context, which, although allowing for controlled comparison between EME and non-EME cohorts, may limit the broader applicability of the findings. Future studies should investigate these phenomena across a wider range of institutions and disciplinary settings, which would help validate and extend these results.

Furthermore, a significant limitation lies in the restricted scope of the linguistic analysis. The methodological necessity of achieving a robust and measurable comparison, alongside the practical constraints of time and resources inherent in a multi-cohort, longitudinal study, required the analysis to focus on a pre-selected, confined list of semi-technical/academic vocabulary and a targeted set of discourse markers. Consequently, while the findings offer reliable data on these key features, they do not account for the students' full range of language production, excluding features such as highly technical discipline-specific lexis and various non-textual metadiscoursal features.

A further limitation concerns the reuse of the same writing prompt at both time points. Although this was a deliberate methodological choice to ensure construct consistency, it may introduce some familiarity effects, as participants could have recalled aspects of the task from the initial administration. Likewise, while lexical scoring was conducted blind, inter-raters were aware that essays responded to the same prompt, which may have introduced subtle expectancy bias. Nonetheless, the two-year interval and anonymised scoring procedures were designed to minimise these risks, thereby strengthening the internal validity of the observed patterns in lexico-pragmatic development.

Moreover, the differing instructional environments introduce a confounding variable into the design. While the EME group benefits from extensive daily exposure through content courses (plus 60 hours of EAP), the non-EME group relies solely on 240 EAP/ESP hours for their instruction in English. This disparity in the nature of exposure (immersive versus explicit) is a key contextual factor and an interpretive constraint when considering the Group x Time interaction effects observed.

The two-year longitudinal design provides a solid foundation for capturing early patterns of language growth, particularly in relation to technical vocabulary and discourse markers usage. However, lexical and pragmatic development is a gradual and context-sensitive process; a longer timeframe, especially one that includes later academic experiences such as study abroad or ERASMUS programmes, could offer a more comprehensive picture of students' lexical repertoires. Future research could employ qualitative methods to explore in greater depth the participants' lived experiences and perceptions of vocabulary acquisition across the two different instructional environments.

179

Acknowledgement

References

Aguilar, M. (2018). Integrating intercultural competence in ESP and EMI: From theory to practice. *ESP Today*, *6*(1), 25–43. https://doi.org/10.18485/esptoday.2018.6.1.2

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world.* Oxford University Press.

Biber, D., & Conrad, S. (2009). *Register, genre and style*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511814358

British Council. (2025). *LearnEnglish teens: Use discourse markers*. https://learnenglishteens.britishcouncil.org/skills/use-discourse-markers

Chapelle, C. A., & Lee, H. (2021). Understanding argument-based validity in language testing. In C. A. Chapelle & E. Voss (Eds.), *Validity argument in language testing: Case studies of validation research* (19–44). Cambridge University Press.

Costa, F., & Mastellotto, L. (2022). The role of English for specific purposes (ESP) in supporting the linguistic dimension in English-medium instruction (EMI). *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, *5*(2), 37–52. https://doi.org/10.5565/rev/clil.91

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. https://doi.org/10.2307/3587951

Coxhead, A. (2024). Current issues and future research on teaching and learning of academic vocabulary in EMI contexts. *International Journal of TESOL Studies*, *6*(2), 105–108. https://doi.org/10.58304/ijts.20240207

Coxhead, A., & Boutorwick, T. J. (2018). Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, maths and science. *TESOL Quarterly*, *52*(3), 588–610. https://doi.org/10.1002/tesq.450

Coxhead, A., & Nation, I. S. P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (252–267). Cambridge University Press.

Dafouz, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics*, *40*(1), 95–113. https://doi.org/10.1016/j.pragma.2007.10.003

Dafouz, E. (2018). English-medium instruction in multilingual university settings: An opportunity for developing language awareness. In P. Garrett & J. M. Cots (Eds.), *The Routledge handbook of language awareness* (170–185). Routledge.

Dafouz, E., & Smit, U. (2020). *ROAD-MAPPING English medium education in the internationalised university*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-23463-8

180

Dafouz, E., & Smit, U. (Eds). (2023). *Researching English-medium higher education: Diverse applications and critical evaluations of the ROAD-MAPPING framework* (1st ed.). Routledge. https://doi.org/10.4324/9781003193852

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes*, *30*(3), 198–208. https://doi.org/10.1016/j.esp.2011.01.001

Hellekjær, G. O. (2010). Lecture comprehension in English-medium higher education. *Hermes: Journal of Language and Communication in Business*, *23*(45), 11–34. https://doi.org/10.7146/hjlcb.v23i45.97343

Huang, P.-Y., & Chiang, Y.-T. (2025). Word senses of academic vocabulary in medical texts and dictionaries. *ESP Today*, *13*(1), 156–174. https://doi.org/10.18485/esptoday.2025.13.1.8

Hyland, K. (2019). *Metadiscourse: Exploring interaction in writing* (2nd ed.). Bloomsbury Academic.

IELTS. (2023, December 7). *IELTS remains the world's most trusted test of English*. https://ielts.org/news-and-insights/ielts-remains-the-worlds-most-trusted-test-of-english

Kamaşak, R., & Sahan, K. (2023). Academic success in English medium courses: Exploring student challenges, opinions, language proficiency and L2 use. *RELC Journal*, *55*(3), 705–720. https://doi.org/10.1177/00336882231167611

Kuteeva, M., & Airey, J. (2014). Disciplinary differences in the use of English in higher education: Reflections on recent language policy developments. *Higher Education*, *67*(5), 533–549. https://doi.org/10.1007/s10734-013-9660-6

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33–51. https://doi.org/10.1177/026553229901600103

Lin, L. H., F. & Morrison, B. (2010). The impact of the medium of instruction in Hong Kong secondary schools on tertiary students' vocabulary. *Journal of English for Academic Purposes*, *9*(4), 255–266. https://doi.org/10.1016/j.jeap.2010.09.002

Lohan, R. P., & Dafouz, E. (2024). A situated analysis of English-medium education in a private business university: Insights from the ROAD-MAPPING framework. *Journal of English for Academic Purposes*, *72*, Article 101457. https://doi.org/10.1016/j.jeap.2024.101457

Lohan, R. P., & Marín Palacios, C. (in press). Academic writing progress: A two-year comparative study of EME and non-EME undergraduate business students.

Malmström, H., Pecorari, D., & Shaw, P. (2025). Development of academic vocabulary knowledge during English-medium instruction. *Ibérica*, *49*, 157–180. https://doi.org/10.17398/2340-2784.49.157

Masrai, A., & Milton, J. (2021). Vocabulary knowledge and academic achievement revisited: General and academic vocabulary as determinant factors. *Southern African Linguistics and Applied Language Studies*, *39*(3), 282–294. https://doi.org/10.2989/16073614.2021.1942097

McKinley, J., & Rose, H. (2022). English language teaching and English-medium instruction: Putting research into practice. *Journal of English-Medium Instruction*, *1*(1), 85–104. https://doi.org/10.1075/jemi.21026.mck

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., 13–103). Macmillan; American Council on Education.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*(2), 235–256. https://doi.org/10.1016/j.esp.2005.05.002

181

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139858656

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009093873

Nguyen, N. H., & Nguyen, K. D. (2022). Vietnamese learners' performance in the IELTS writing task 2: Problems, causes, and suggestions. *International Journal of TESOL & Education*, *2*(1), 170–189. https://doi.org/10.54855/ijte.222111

Othman, J. (2024). Academic vocabulary learning in EMI classrooms: Challenges and strategies. *Arab World English Journal*, *15*(2), 3–18. https://doi.org/10.24093/awej/vol15no2.1

Pun, J. K. H., & Tai, K. W. H. (2021). Doing science through translanguaging: A study of translanguaging practices in secondary English as a medium of instruction science laboratory sessions. *International Journal of Science Education*, *43*(7), 1112–1139. https://doi.org/10.1080/09500693.2021.1902015

Reynolds, B. L., & Skoufaki, S. (2024). Introduction to the special issue on teaching and learning of academic vocabulary in EMI contexts [Editorial]. *International Journal of TESOL Studies*, *6*(2), 1–3. https://doi.org/10.58304/ijts.20240201

Reynolds, B. L., Xie, X., & Pham, Q. H. (2022). Incidental vocabulary acquisition from listening to English teacher education lectures: A case study from Macau higher education. *Frontiers in Psychology*, *13*, Article 993445. https://doi.org/10.3389/fpsyg.2022.993445

Rogier, D. (2012). *The effects of English-medium instruction on language proficiency of students enrolled in higher education in the UAE* [Doctoral dissertation, University of Exeter]. ProQuest Dissertations and Theses Global.

Roothooft, H., Breeze, R., & Meyer, M. (2025). English writing competence and EMI performance: Student and expert perceptions of academic writing in EMI. *ESP Today*, *13*(1), 2–22. https://doi.org/10.18485/esptoday.2025.13.1.1

Rose, H., Curle, S., Aizawa, I., & Thompson, G. (2020). What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation. *Studies in Higher Education*, *45*(11), 2149–2161. https://doi.org/10.1080/03075079.2019.1590690

Sánchez-Pérez, M. (2023). The impact of EMI on participant English writing proficiency in a Spanish undergraduate engineering context. *Studies in Second Language Learning and Teaching*, *13*(2), 373–397. https://doi.org/10.14746/ssllt.38279

Schoepp, K. (2018). Predictive validity of the IELTS in English as a medium of instruction. *Higher Education Quarterly*, *72*(4), 271–285. https://doi.org/10.1111/hequ.12163

Soruç, A., Altay, M., Curle, S., & Yuksel, D. (2021). Students' academic language-related challenges in English medium instruction: The role of English proficiency and language gain. *System*, *103*, Article 102651. https://doi.org/10.1016/J.SYSTEM.2021.102651

Tang (唐宇), Y., & Curtis, A. (2025). English medium instruction academic success (EMIAS) model: A grounded theory-informed systematic review in non-anglophone contexts. *Higher Education Research & Development*, 1–19. https://doi.org/10.1080/07294360.2025.2527396

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

Wei, M., & Macaro, E. (2024). Vocabulary learning in EMI courses: A longitudinal study of Chinese university students over a term. *System*, *126*, Article 103479. https://doi.org/10.1016/j.system.2024.103479

Wilkinson, R. (2013). English-medium instruction at a Dutch University: Challenges and pitfalls. In A. Doiz, D. Lasagabaster, & J. Sierra (Eds.), *English-medium instruction at universities* (3–26). Multilingual Matters.

Wu, M. Y., Steinkrauss, R., & Lowie, W. (2023). The reliability of single task assessment in longitudinal L2 writing research. *Journal of Second Language Writing*, *59*, Article 100950. https://doi.org/10.1016/j.jslw.2022.100950

University of Warwick. (2020, July 15). *Academic English skills: Discourse markers.* https://warwick.ac.uk/fac/soc/al/globalpad-rip/openhouse/academicenglishskills/grammar/discourse/

Uysal, H. H. (2010). A critical review of the IELTS writing test. *ELT Journal*, *64*(3), 314–320. https://doi.org/10.1093/elt/ccp026

RHONA PATRICIA LOHAN is a PhD candidate in Applied Linguistics at Universidad Complutense de Madrid and a lecturer in English for Academic Purposes and Scientific Research at ESIC University, Madrid. Her research interests focus on English-medium education (EME), academic writing, English for specific and academic purposes (ESP/EAP), and sociolinguistics. This article is part of her longitudinal comparative study examining the impact EME has on English proficiency. Findings from this study have been published in the *Journal of English for Academic Purposes*, with additional papers from the same project forthcoming. ORCID: 0000-0002-7590-9548

Appendix A

Semi-technical/academic vocabulary by theme

DIGITAL/ONLINE WORK:

Terms
platform
webinar
asynchronous
synchronous
connectivity
video conference/call
screen-(sharing)
breakout rooms
virtual
technology
teleworking
digitalisation
remote

COMMUNICATION:

Terms
interaction
collaboration

engagement
feedback
participation
clarity

PRODUCTIVITY:

Terms
efficiency
flexibility
convenience
productivity
multitasking
time management
pay attention

TRAINING/EDUCATION:

Terms
online training
workshop
tutorial
self-paced
instructor-led

CHALLENGES/DISADVANTAGES:

184

Terms
isolation
technical issues
distractions
fatigue
time zone differences
stress
challenges

MISCELLANEOUS:

Terms
commuting
comfort zone
cut costs/expenses
attend

Appendix B
Discourse markers by function

LISTING:
firstly/secondly
first of all
to begin with
in addition
moreover

furthermore
on the one hand/on the other hand

EXEMPLIFICATION:
for example
for instance
such as
to illustrate this

CONTRAST/CONCESSION:
however
nevertheless
nonetheless
in contrast
on the contrary
despite this
although
whereas
while

CAUSE/EFFECT:
therefore
as a result
consequently
this leads to
because of
due to
owing to

185

CONCLUSION:
in conclusion
to sum up
overall
to summarize
all things considered