ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

LINGUE, LETTERATURE E CULTURE MODERNE:

DIVERSITÀ E INCLUSIONE

Ciclo 36

Settore Concorsuale: 10/ANGL-01 - ANGLISTICA E ANGLOAMERICANISTICA

Settore Scientifico Disciplinare: ANGL-01/C - LINGUA, TRADUZIONE E LINGUISTICA INGLESE

EMBODIED LECTURING IN ENGINEERING IN ENGLISH-MEDIUM INSTRUCTION (EMI): EXPLORING THE INTERACTION BETWEEN GESTURES, (DIS)FLUENCIES, AND PRAGMATIC CHALLENGES

Presentata da: Mariangela Picciuolo

Coordinatore Dottorato

Serena Baiesi

Supervisore

Jane Helen Johnson

Co-supervisore

Sabrina Fusari

Esame finale anno 2025

Table of Contents

Abstract

This study investigates the communicative challenges faced by an Italian first language (L1) lecturer delivering engineering courses through English as a Medium of Instruction (EMI) in an international master's program in Italy. Using a case study approach, the research explores how variations in speech rate, disfluencies, and the lecturer's use of verbal and non-verbal strategies influence the teaching and learning experience. Data were triangulated from a video-recorded lecture, student feedback, and the lecturer's reflections, with the analysis centered on student perspectives to address their pressing challenges in EMI classrooms.

The findings highlight significant variations in the lecturer's speech rate, particularly during the explanation of complex technical content. Higher speech rates are associated with straightforward explanations, while slower rates occur during more conceptually challenging segments. Additionally, the study examines the relationship between speech rate and the lecturer's pragmatic functions to understand the causes behind these fluctuations. Disfluencies were analyzed to differentiate between pauses caused by communication breakdowns and those used deliberately as communicative strategies to engage students or give them time to process complex material in their second language (L2). Gestures were also scrutinized to assess their role in either facilitating understanding or compensating for challenges in lexical retrieval and complex explanations.

Although the lecturer's technical expertise supports content mastery, rapid speech during explanations, issues with lexical retrieval, and occasional misalignment between gestures and speech contribute to students' challenges in understanding vocabulary, particularly in STEM fields. While the study reinforces that both students and lecturers cooperate effectively to achieve communicative goals in English as a

Lingua Franca in Academic (ELFA) contexts, it also emphasizes the need to further empower EMI lecturers by providing research-based evidence to improve specific areas of communication.

From a methodological perspective, this study expands the scope of EMI research beyond traditional focuses on English for Specific Academic Purposes (ESAP) and English Language Teaching (ELT) by integrating Second Language Acquisition (SLA) and Conversation Analysis (CA) approaches. It also contributes to the growing interest in applying a multimodal lens to the analysis of EMI. Furthermore, it emphasizes the importance of viewing verbal and so-called non-verbal resources, such as gestures, as integral communicative tools. This study further provides evidence that gestures are not mere accessories to speech but meaningful tools that may reveal challenges lecturers and students face in EMI settings, which may not be expressed through words. Given the embodied nature of teaching, the findings reinforce the value of studying gestures as playing a key role in either facilitating or hindering comprehension. The results support the need for EMI-specific training programs that enhance communicative strategies tailored to the disciplinary needs of both international and domestic students, particularly in technical fields like engineering.

Keywords: English as a Medium of Instruction (EMI); Engineering Education; EMI Lecturers; Speech Rate; Disfluencies; Pragmatic Functions; Communicative Strategies; Gestures.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Jane H. Johnson, for her unwavering guidance, encouragement, and support throughout the entire journey of writing this thesis. Her insightful advice and constructive feedback have been invaluable in shaping this work.

I am also sincerely thankful to my reviewers for their patience and diligence in reviewing my thesis. Their thoughtful comments and suggestions greatly contributed to improving the quality of this research.

A special thank you goes to the lecturer who kindly granted me consent to use his lectures as a resource. His inspiring teaching ignited many of the ideas explored in this thesis, and I am grateful for the influence he has had on my academic development. I would also like to extend my appreciation to the colleagues at the DICAM (Civil, Chemical, Environmental, and Materials Engineering) department who attended my training course and contributed to creating a stimulating and collaborative learning environment.

To my family, I owe an immense debt of gratitude for their unwavering support and understanding throughout this process. A heartfelt thank you goes to my lovely sister, Margherita, for always being there for me.

To my wonderful partner, Giulio, I am deeply grateful for his unwavering love, patience, and encouragement, which sustained me during the most challenging times. His belief in me and constant reassurance gave me the strength to persevere.

I would also like to thank my friends, whose support, encouragement, and moments of joy helped me stay grounded and motivated. Your laughter, companionship, and belief in my abilities lightened the toughest days and reminded me of the importance of balance and connection.

Lastly, I cannot forget my beloved cat, *Bo*, whose nightly companionship and tireless energy kept me awake during long writing sessions. Your presence was a constant comfort throughout this journey.

This thesis is as much a product of your collective support as it is of my own efforts, and for that, I am forever grateful.

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Background and Context

Globalization has significantly reshaped the landscape of higher education (HE), particularly in STEM disciplines. English has become the dominant language for academic publications, international conferences, and collaborative research. Consequently, universities worldwide have adopted English-Medium Instruction (EMI) as "'killing two birds with one stone' to achieve internationalization goals and improve English proficiency of domestic students" (McKinley & Rose, 2022, p. 9). In Europe, the Bologna Process has supported this shift, promoting internationalization, mobility, and standardization across higher education systems.

In Italy, many universities have introduced EMI engineering programs at both undergraduate and postgraduate levels to attract international students and align their curricula with global standards. However, the success of these programs relies not only on the linguistic proficiency of the lecturers but also on their ability to clearly and accurately convey complex concepts in English.

Non-native English-speaking (NNES) lecturers teaching engineering subjects in EMI settings face several linguistic challenges, which stem primarily from the specialized vocabulary (McDonough, 2010), the explanation of abstract concepts (Arden-Close, 1993; Swales, 1995), and the necessity for precision in technical communication, impacting all language domains.

As regards vocabulary, engineering involves highly specialized terminology, and precision is essential. NNES students may struggle with the extensive vocabulary required, particularly when technical terms lack direct translations in their native language. Additionally, the core disciplinary values inherent in engineering, along with the associated discourse, often involve arguments that connect scientific phenomena to specific real-world contexts, reflecting the practical and applied nature of the field (Archer, 2008; as cited in Parkinson, 2012, p. 157). These contexts may differ from those familiar to the students. However, content lecturers often appear "unaware of the linguistic implications of their teaching and of their students' needs" (Francomacaro, 2011, p. 67).

In addition, research has shown that the discourse of science and engineering is characterized by the frequent use of nominalizations, where highly complex abstract information is compressed into a single word, such as *curvature* in "curvature of material surfaces" (Pueyo & Val, 1996, p. 258). The abstractness of technical terms in science and engineering (Arden-Close, 1993) requires EMI lecturers to employ pedagogical tools to clarify these abstract concepts by connecting them to familiar experiences. Research has shown that EMI lecturers in STEM fields often rely on communicative strategies such as analogy (Kunioshi et al., 2016) and storytelling (Nesi & Alsop, 2021) to achieve this. However, these strategies can sometimes be misleading, as the effectiveness of an analogy, as well as storytelling, strongly depends on students' prior knowledge of the subject. Participants with limited knowledge of the source domain from which the analogy is drawn are significantly more likely to develop misconceptions (Wilbers & Duit, 2006). This is particularly likely in EMI settings, where lecturers must deal with a diverse cohort of students.

Technical communication also requires precision, and in this respect, pronunciation also plays a role in ensuring clarity. Mispronunciations of technical terms – common among lecturers whose L1 have different phonetic structures – can lead to misunderstandings. In engineering, where terms like "duct" and "ductile" or "stress" and "strain" carry significant meaning, these errors can be particularly problematic. As a result, lecturers may need to rely on non-verbal resources such as visuals, diagrams, and graphs, which are nevertheless commonly used in scientific instruction. As Poe et al. (2010, p. 115) note, "visuals comprise, on average, 26 percent of the surface area of the twentieth-century research article, with the Cartesian graph being the predominant visual found in research articles today," highlighting the importance for instructing engineering students in mastering this feature of the academic genre.

Therefore, to understand the dynamics of EMI in technical fields, it is essential to analyze both verbal and non-verbal strategies employed by NNES lecturers, as these non-verbal cues are paramount in supporting their verbal explanations.

1.2 Problem statement

Non-verbal strategies play a crucial role in enhancing the delivery of content, but verbal language has its own unique role, complementing non-verbal communication. The effective use of language remains equally significant, especially in fields where technical accuracy and clarity are paramount, as it ensures that complex technical concepts are conveyed with both precision and accessibility.

In this respect, Wellington and Osborne (2001) argue that "paying more attention to language is one of the most important acts that can be done to improve the quality of science education" (p. 1), aligning with the views of many other science education researchers: language in science is crucial. Similarly, Lemke (1990, p.1) expands on the importance of language in science education by explaining that:

> Learning science means learning to talk science. It also means learning to use this specialized conceptual language in reading and writing, in reasoning and problem-solving, and in guiding practical action in the laboratory and in daily life. It means learning to communicate in the language of science and act as a member of the community of people who do so. 'Talking science' means observing, describing, comparing, classifying, analyzing, discussing, hypothesizing, theorizing, questioning, challenging, arguing, designing experiments, following procedures, judging, evaluating, deciding, concluding, generalizing, reporting, writing, lecturing, and teaching in and through the language of science.

As the authors point out, content lecturers in science fields often appear unaware of the linguistic implications of their teaching. There are notable exceptions, however.

As early as the 1950s, the Massachusetts Institute of Technology (MIT) integrated communication instruction into various core science and engineering laboratory subjects (Poe et al., 2010) through the Writing Across the Curriculum Program, "where communication instructors negotiate with subject instructors to design curricular interventions in communication that seek to improve student learning of the subject matter" (Poe et al., 2010, foreword, viii). Although Poe et al. (2010) make few mentions of international students attending such courses at MIT, and their research primarily focuses on NES lecturers interacting with NES students, their study was nevertheless driven by the following research questions:

- How do students learn the persuasive devices that professional scientists use when communicating data to other scientists?
- What challenges do students encounter as they learn to use visual evidence in scientific communication?
- What role does faculty feedback play in the development of this professional skill? (Poe et al., 2010, p. 112)

To English-Medium Instruction (EMI) researchers and practitioners, particularly those interested in EMI in the field of engineering, these questions are likely to sound quite familiar, as is the growing call for cooperation between content specialists and language experts – a call that has largely gone unheard to this day (see, e.g., Lasagabaster, 2008).

Talking (teaching) science is even more challenging for non-native English-speaking (NNES) lecturers teaching content subjects in English as a Lingua Franca (ELF) in academic institutions, where English has become a language of *necessity* (as opposed to the language of *choice* in ELF settings; see Jenkins, 2009). This shift has been driven by the internationalization of Higher Education (HE), with one of its most immediate outcomes being the introduction of English-Medium Instruction (EMI) at universities.

In this respect, the globalization of HE, particularly in technical fields like engineering, has made English the predominant instructional language in non-

Anglophone countries. EMI has become a key strategy for universities aiming to attract international students, enhance global recognition, and prepare graduates for an international workforce. The Bologna Declaration (1999) and the creation of the European Higher Education Area (EHEA) have accelerated the adoption of EMI, especially in engineering, where courses are increasingly taught to multicultural and multilingual student cohorts. Nonetheless, research has shown that engineering EMI students have generally a lower mastery of the English language compared to students from different disciplinary EMI fields (Wächter & Maiworm, 2014).

Similarly, lecturers in EMI international contexts face linguistic, communicative, and cultural challenges that may hinder communication in the classroom. These challenges are particularly amplified when teaching complex technical content, where precision and clarity are critical. In this respect, previous research (e.g., Airey, 2020) has shown that engineering lecturers often do not express a need for training in this area, as they consider themselves content experts already familiar with the technical jargon. As Curle et al. (2020, p. 53) pointed out,

> Content lecturers using EMI may be familiar with the subject-specific vocabulary of their disciplines and aim to develop students' literacy in their specific subject fields; however, they might not be aware of other subject-specific features the texts in their disciplines display, e.g. how to develop an argument, present evidence and use quotes (Nesi & Gardner, 2012; Dafouz, 2018; Block & Moncada-Comas, 2019).

The present study focuses on spoken English and *overt* communication challenges within EMI, specifically in the field of engineering. Engineering instruction in English-taught programs (ETP) poses unique linguistic and conceptual challenges due to its specialized vocabulary (McDonough, 2010), abstract concepts (Arden-Close, 1993; Swales, 1995), and the need for precision in technical communication. These challenges are further intensified in EMI settings, where both NNES lecturers and students must communicate in an L2 and bring differing assumptions regarding disciplinary values, academic purposes, lecture organization, interaction patterns, and assessment methods, to name a few.

Indeed, research has increasingly shown that communication challenges in EMI, particularly in the sciences, are often subtle and not immediately apparent to either lecturers or students. These challenges include not only the linguistic aspects of technical communication but also paralinguistic and non-verbal elements such as speech rate, disfluencies, and gestures, which play a crucial role in ensuring effective communication. The growing availability of training programs for EMI lecturers in Europe – where research on EMI has been ongoing for 20 years (see, e.g., Deroey, 2023, for an overview) – has led to increased awareness among content lecturers of the challenges associated with teaching in EMI courses. Nonetheless, as Macaro (2019) aptly points out,

> Content specialists are busy people. Learning about language issues will be an additional burden to carrying out their research in their own field. Some may be on the verge of making discoveries of huge importance to the planet and to the human race. I personally would fully understand if these academics said to me that 'EMI training' has to be firmly framed by that perspective (Macaro, 2019, p. 274).

Similarly, Airey (2020, p. 343) recognizes that "[i]t is difficult enough to convince content lecturers that they need to develop pedagogical content knowledge in their L1 teaching". Airey (2020) further argues that,

> One of the earliest conclusions from my own work is that EMI – however, one defines it – simply exacerbates communicative issues that already exist in monolingual L1 settings (Airey and Linder 2006). The problem is that content lecturers tend to underestimate the role of languages and other semiotic resources in the teaching and learning of their discipline.

As a physicist, Airey (2020, p. 343) recently argued that one way "to get content lecturers to reflect on the linguistic goals they have for their students" is to convince them that they are "disciplinary teachers". This implies recognizing that challenges in mastering disciplinary discourse in an L2 – and even more so in teaching it to students from diverse linguistic, cultural, social, and academic backgrounds – differ across disciplines. Additionally, it involves acknowledging that difficulties may arise in less visible or inaudible areas, such as what speakers cannot say or choose not to articulate (Björkman, 2008).

1.3 Research Objectives

The primary objective of this study is to explore the communicative challenges faced by a non-native English-speaking (NNES) lecturer in an EMI engineering course, focusing on speech rate, disfluencies, gestures, and the lecturer's and students' perceptions of these teaching practices. The following research questions guide the investigation:

RQ1: What are the lecturer's perceptions regarding their own teaching performance in EMI, and how do they perceive students' needs and major challenges?

This question explores the lecturer's self-reflection on their teaching practices and challenges within the EMI context, providing insights into how they adjust to students' learning needs.

RQ2: What are students' perceptions of their lecturer's teaching performance in EMI engineering classes, and what challenges do they report?

This question focuses on gathering students' views on the effectiveness of the lecturer's teaching strategies and the challenges they face in comprehending EMI instruction.

RQ3: What are the key misalignments between the lecturer's perceptions of their teaching performance and the students' reported challenges in the EMI context?

This question examines the discrepancies between the lecturer's self-assessment and the challenges reported by students, with the aim of identifying areas where the lecturer's strategies may not fully align with students' linguistic and cognitive needs.

RQ4: During which pedagogical phases does the lecturer tend to alter speech rate (e.g., speed up or slow down), and what pragmatic functions correspond to these shifts?

This question explores the relationship between the lecturer's speech rate and the pragmatic functions they perform, identifying patterns in how speech rate shifts according to different phases of the lecture.

RQ5: How do slower speech rates correspond to pauses, and are these pauses markers of (dis)fluencies (DFs)? How are these markers distributed throughout the lecture, and with which pragmatic functions do they tend to occur most frequently?

This question investigates the connection between slower speech, disfluency markers, and pragmatic functions, mapping where disfluencies occur and analyzing the corresponding pragmatic functions.

RQ6: How are disfluency markers (DFs) and communicative strategies (CSs) connected, and what repair strategies do the lecturer and students use to address disfluencies?

This question examines how the lecturer and students manage disfluencies, particularly focusing on the relationship between disfluency markers and the communicative strategies used to maintain communication flow during the lecture.

RQ7: Can the lecturer's gestures reveal more about linguistic challenges, such as difficulties in word formulation or conceptualization? How do gestures interact with the lecturer's verbal explanations?

This question explores the role of gestures in overcoming linguistic challenges, investigating how gestures support or interact with the lecturer's verbal explanations, particularly when facing difficulties in word retrieval or conceptualization.

1.4 Significance of the Study

This study investigates the communicative challenges faced by an Italian first language (L1) lecturer delivering engineering courses through English as a Medium of Instruction (EMI) in an international master's program in Italy. Using a case study approach, the research explores how variations in speech rate, disfluencies, and the lecturer's use of verbal and non-verbal strategies influence the teaching and learning experience. Data were triangulated from a video-recorded lecture, student feedback, and the lecturer's reflections, with the analysis centered on student perspectives to address their pressing challenges in EMI classrooms.

From a methodological perspective, this study expands the scope of EMI research beyond traditional focuses on English for Specific Academic Purposes (ESAP) and English Language Teaching (ELT) by incorporating insights from Second Language Acquisition (SLA) and Conversation Analysis (CA) to explore how communication strategies influence comprehension and engagement in EMI contexts. It also contributes to the growing interest in applying a multimodal lens to the analysis of EMI. Furthermore, it emphasizes the importance of viewing verbal and so-called non-verbal resources, such as gestures, as integral communicative tools. This study further provides evidence that gestures are not mere accessories to speech but meaningful tools that may reveal challenges lecturers and students face in EMI settings, which may not be expressed through words. Given the embodied nature of teaching, the findings reinforce the value of studying gestures as playing a key role in either facilitating or hindering comprehension.

1.5 Overview of the Thesis Structure

This thesis is structured as follows: Chapter 2 reviews the relevant literature on EMI in STEM fields, focusing on linguistic and pedagogical challenges. Chapter 3 presents the methodology employed, including both qualitative and quantitative analyses of speech rate, disfluencies, communicative strategies, pragmatic functions and gesture analysis. Subsequent chapters discuss the findings (Chapter 4) and their implications for EMI training and support programs (Chapter 5).

Chapter 2. Literature Review

2.1 English as a Medium of Instruction in Engineering

The increasing globalization of higher education has led to the widespread adoption of EMI in STEM fields, particularly engineering, across non-English-speaking countries. In a large-scale study, Wächter and Maiworm (2014) found that, according to the International Standard Classification of Education (ISCED) 2011, English-Taught Programs (ETPs) across Europe are most commonly offered in the fields of social sciences, business, and law (35%), followed by sciences (23%), with engineering, manufacturing, and construction representing the third largest share (18%). These three fields alone account for 76% of all EMI programs offered, highlighting the importance of addressing language and communication challenges in these areas – particularly in engineering, where the English proficiency of both foreign and domestic students is considerably lower than that of students in other disciplines (Wächter & Maiworm, 2014). Furthermore, in a 2010 review of ESP materials, McDonough identified over 20 professional fields, including engineering, where English is essential for effective communication.

Engineering instruction in ETP presents distinct linguistic and conceptual challenges that set it apart from other STEM disciplines. These challenges stem primarily from specialized vocabulary (McDonough, 2010), abstract concepts (Arden-Close, 1993; Swales, 1995), and the necessity for precision in technical communication, impacting all language domains. These linguistic demands are closely tied to the core disciplinary values inherent in the engineering field, such as "the design process and problem-solving within predefined specifications" (Parkinson, 2012, p. 157). Consequently, the associated discourse often involves arguments that connect scientific phenomena to specific real-world contexts, reflecting the practical and applied nature of engineering (Archer, 2008; as cited in Parkinson, 2012, p. 157). For instance, Artemeva (1998; as cited in Parkinson, 2012, p. 157) observed that the contrasting values between a North American engineering company and a Russian engineering company led to differing perspectives on rhetorical purpose, audience, and organizational structure. These differences were evident in the sentence and paragraph organization, thematic structure, and even the content of periodic progress reports. While North American engineers focused on finding practical solutions to technical problems, Russian engineers identified more with the role of scientists. Similarly, Italian engineering practices emphasize the integration of experiential knowledge with systemic frameworks, reflecting the need to balance project-specific and contextual demands with theoretical principles (Secundo et al., 2015). This pragmatic approach, shaped by task volatility and project-oriented workflows, aligns with the broader cultural and institutional contexts in which Italian engineers operate.

This dynamic is particularly relevant in EMI academic settings, where students from diverse social, linguistic, cultural, and academic backgrounds may bring distinct disciplinary values that sometimes conflict with those of the institution offering the EMI degree program. Italian engineering programs, for example, are not developed in isolation but are influenced by the cultural and institutional environment in which they exist. They are rooted in bridging the gap between theory and practice, emphasizing the importance of adapting communication strategies to align with varied expectations and knowledge frameworks.

ESP research in the field of engineering has increasingly focussed on the need for L2 engineers to acquire specialized vocabulary (McDonough, 2010). In this context, Ward (1999, 2001, 2009) explored the challenges faced by undergraduate engineering students in Thailand when reading English-language textbooks, with a particular focus on vocabulary. In his 1999 research, Ward used a corpus-based approach to investigate the vocabulary size necessary for EAP engineering students. He questioned whether a specialized engineering corpus could offer similar reading efficiency benefits comparable to those associated with the commonly accepted threshold of 3,000 word families – i.e., all related forms of a headword, such as *use*, *used*, *uses*, *using*, *usable*, *usefulness*, *useless*, *user*, and *users* (Ward, 1999, p. 310) – which is generally considered necessary for effective reading comprehension in a second language.

Ward constructed a corpus from engineering textbooks used at his university and compared it with general word list corpora. He concluded that engineering students should focus on the specific words and terms commonly used in engineering texts. He found that a specialized vocabulary of 2,000 word families could cover up to 95 percent of a foundation-level engineering textbook from his corpus, suggesting that this specialized vocabulary can encompass a significant portion of the language needed for their specific academic field. In addition, research has also shown that the discourse of science and engineering is characterized by the frequent occurrence of nominalizations, where highly complex abstract information is compressed into a single word, e.g., *curvature of material surfaces* (Pueyo & Val, 1996, p. 258). This characteristic of engineering discourse adds to the challenges faced by L2 students, requiring them to decode dense and abstract language structures effectively.

In his 2001 survey-based study, Ward examined the coping strategies engineering students employed when facing vocabulary challenges in textbook reading. The findings revealed that students often bypass difficult vocabulary by concentrating on the applications and examples provided in the textbooks (Ward, 2001, p. 151).

Ward's multi-year research culminated in the development of an English word list specifically tailored for lower proficiency engineering undergraduates (Ward, 2009). This final list contained 299 word types, covering 16.4 percent of a corpus of engineering textbooks. The top ten words on Ward's list included *system*, *shown*, *equation*, *example*, *value*, *design*, *used*, *section*, *flow*, and *given*.

Furthermore, engineering lexis varies considerably across genres (e.g., research articles, lab reports, design reports) as well as across different technical and engineering disciplines (Braine, 1995), meaning that the vocabulary students are expected to use can differ significantly depending on the specific field (Hyland & Tse, 2007). For instance, Hyland and Tse (2007) found that the same words could behave differently across subject areas, with *process* occurring more frequently as a noun in science and engineering, while it appeared more often as a verb in the social sciences. In this context, further research (Yang, 1986; Liu & Nesi, 1999; Mudraya, 2006; Ward, 2007; Roesler, 2021) has also shown that many polysemous vocabulary items have both general-usage meanings and meanings that are associated with specific disciplines.

Mudraya (2006) distinguished between technical words and sub-technical words in the field of engineering. Technical words, such as *urethane*, have no exact synonym, resist semantic change, and have a narrow range of use. In contrast, sub-technical words, such as *resistance*, *channel*, or *tension*, have both engineering and non-engineering meanings. In this regard, Liu and Nesi (1999), in collaboration with engineering lecturers, developed vocabulary tests based on their technical and sub-technical wordlists. Their study focused on NNES overseas students, particularly those from Hong Kong and Malaysia, who were enrolled in MSc engineering programmes at Warwick University. These students were tested at the end of their year abroad, after spending twelve months in the UK. The results indicated that many technical terms remained unrecognizable to the students, while sub-technical words were more readily recognized, even by the end of their degree programs.

The authors suggested that their findings could guide the design of future EAP syllabi, which often aim to address multiple subject specialisms, and proposed that a more targeted focus on specific vocabulary areas might be more effective. Furthermore, they considered that students' struggles with technical terminology might stem from inadequate listening skills, which are essential for understanding oral definitions in lectures, as well as insufficient strategies for interpreting glosses and definitions in subject textbooks.

Addressing these challenges, research has also explored strategies to improve reading comprehension. For instance, Pritchard and Nasr (2004) found that exposing undergraduate engineering students in Egypt to authentic texts within their discipline helped them retrieve relevant information using both textual and contextual clues, which in turn yielded positive outcomes.

Hall et al.'s (1986, p. 152) study was similarly influenced by the difficulties faced by English for Science and Technology (EST) students at a university in Thailand when reading discipline-specific texts. They noted that students struggled particularly with the organization of ideas in textbooks. In response, they developed a course approach that emphasized the connections between the content within the texts and external information – what they termed "macro-cohesion" – as well as the links at the sentence level, referred to as "micro-cohesion."

A considerable body of research has shown that engineering students also undertake writing tasks which are specific to their discipline (see, e.g., Hyland 2009). While, for instance, writing in the humanities and social sciences require students to analyze and synthesize from multiple sources, in science and technology, students are mainly tasked with activity-based tasks, such as describing procedures, defining objects, and planning solutions.

As regards listening comprehension in the oral genre of the academic lecture in engineering, Olsen and Huckin (1990) found that NNES students particularly struggled with comprehending the rhetorical problem–solution structure of the lecture. Students were asked to provide immediate recall summaries after viewing a

16-minute videotaped lecture on a topic in mechanical engineering. Summaries were then discussed between the researchers and the content lecturer. It was found that although students were asked to view a very short excerpt of the lecture, most of them failed to understand the main points and how they fitted together. As the authors argue:

> A “successful” summary therefore was defined as one 1. which identified the problem-solution structure of the lecture, 2. which identified the relation of theory to tests of theory, and 3. which described the effect of debonding on the propagation of cracks and the need to trade off adhesive strength for tensile strength.

The following figure (Figure 1) illustrates the structure of the lecture analyzed, as outlined by Olsen and Huckin (1990, p. 35).

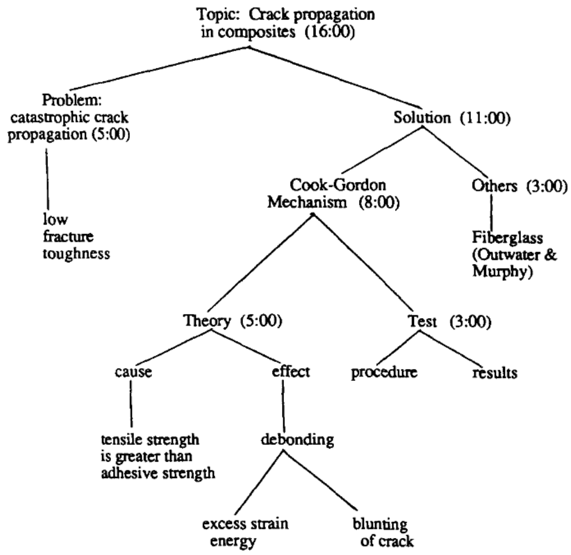

Figure 1. Overall structure of a lecture in mechanical engineering analyzed by Olsen and Huckin (1990, p. 35) in their experiment to identify students’ challenges in lecture comprehension.

They noticed that successful students used a “point-driven” strategy while the unsuccessful ones used an “information-driven” strategy. As the authors argue (p. 41):

> listeners using an information-driven strategy simply try to absorb facts; they are more concerned with information per se than with the speaker’s intentions or goals. By contrast, listeners using a point-driven strategy take a broader view, a more context sensitive view of the interaction between speaker and listener, where context includes the speaker’s presumed intention, the usual goals of the particular genre, the larger situation of which the discourse is a part, the potential role of relevant issues in the larger context, and even cultural effects between speakers of different cultures or subcultures.

Taking a narrower linguistic perspective, Olsen and Huckin (1990, p.42) recognize that the lecture investigated was "clearly point-driven" as the lecturer made use of metadiscourse markers – such as *The real problem is*, *That's all I'm implying here*, *So I'm just indicating here*, *The whole idea here*, and *The key thing is* – as well as prosodic markers to give intonational emphasis on key words and phrases – like *actually*, *even though*, and *debond* – along with visual cueing. However, these cues were apparently lost on most of the students tested. As the authors suggest, while it's crucial in science and engineering to develop broad problem-solving abilities, these skills are not frequently emphasized in student education. This situation appears to be largely influenced by a type of "disciplinary "cultural" conditioning" associated with the methods used in science and engineering education, which seems to contribute significantly to engineering students' difficulty in understanding the lecture. Therefore, Olsen and Huckin (1990, p. 44) argue that "if we are to teach students (both native and nonnative) to understand and communicate more effectively, we may need to help them see the larger goals, agendas, and contexts in their fields, as well as the organization of their discourse".

In line with Olsen and Huckin's (1990) study, Csomay and Wu (2020) recently compared university classroom language in an anglophone U.S. context and an EMI Singapore context by adopting a Multidimensional (MD) analysis approach (Biber, 1988). The U.S. corpus – a subset of the Michigan Corpus of Academic Spoken English (MICASE; Simpson-Vlach & Leicher, 2006) – and the EMI Singapore corpus included lecture and seminar recordings from three disciplines: humanities, natural science, and engineering. The results of the comparison revealed systematic variations between the two contexts. For instance, at the beginning of humanities classes, lecturers in the U.S. context typically use language associated with a contextual, directive orientation, whereas lecturers in Singapore tend to use language associated with a conceptual, informational focus. The authors suggest that these differences may stem from variations in academic culture and pedagogical approaches, which lead to differences in discourse organization.

Olsen and Huckin (1990, p. 42) also mentioned visual cueing as one of the discourse features that makes the lecturers discourse “clearly point-driven”. In this context, ESP research has long emphasized the importance of graphs and diagrams as integral components of science and technology discourse. Myers (2003) observed that visual representations – such as diagrams, images, flow charts, graphs, and micrographs – are central to the social practices of engineering and science disciplines. However, these are not easily accessible forms of language; rather, they are conventional depictions that require disciplinary knowledge to comprehend. Therefore, science and engineering students must learn to “read” these visual forms. Consequently, they need to be taught how different modes – visual, written, and oral – interact. In a similar vein, Airey and Linder (2008, 2009) suggest that educators should motivate students to ask questions either during or after class, distribute lecture materials beforehand, and complement their verbal explanations with more visual aids.

2.1.1 Linguistic and Pedagogical Challenges for EMI Lecturers in Engineering

Research on EMI has consistently highlighted that the needs of NNES lecturers and students vary across disciplines (Evans & Morrison, 2011; Kuteeva & Airey, 2014; Chan, 2015). In the sciences, where English is more widely used, its use is acknowledged as a “pragmatic reality” (Bolton & Kuteeva, 2012, p. 444).

In the literature, few studies have focused exclusively and explicitly on the engineering EMI field. Most research, particularly survey-based studies, has taken a comparative approach, examining lecturers’ perceptions across two or more disciplines. We will therefore focus on these studies, while contrasting their findings with the broader literature on challenges faced by EMI lecturers in general.

Klaassen's (2001) survey-based study is one of the first to investigate the complexities and challenges of delivering engineering education in English at international universities. The research found that EMI lecturers faced *linguistic*, *pedagogical*, and *cultural* challenges in the design and delivery of EM engineering education to a diverse, international student body (Klaassen, 2008).

Regarding *language-related challenges*, earlier studies have suggested that in "hard core EMI" subjects (Macaro, 2020, p. 264), where mathematical codes and formulas are frequently used, language is perceived to play a lesser role in teaching (Dearden & Macaro, 2016; Macaro et al., 2016; Macaro, Curle et al., 2018; as cited in Curle et al., 2020, p. 29). These findings align with Francomacaro's (2011) study, which investigated Italian EMI lecturers' perceptions of their English language competence. In this study, lecturers reported feeling confident in their English skills and did not believe they faced issues in interacting with students or evaluating their progress. However, Francomacaro (2011) also noted that content lecturers seemed "unaware of the linguistic implications of their teaching and of their students' needs" (p. 67). This observation is supported by several studies conducted in the Italian context, where lecturers from various fields, including engineering, expressed concerns about their English proficiency (Pulcini & Campagna, 2015; Campagna, 2016; Picciuolo & Johnson, 2020) and identified their language skills as a major challenge in teaching EMI programs (Guarda & Helm, 2016).

In the Spanish context, Aguilar's (2015) survey-based study identified three main themes in the responses provided by the engineering lecturers surveyed, two of which relates to language, while the third relates to the multicultural dimension of EMI. As regards language, lecturers in this study were unwilling to provide linguistic feedback to their students, as they mainly considered themselves as content lecturer. This finding is in line with broader literature in EMI in other disciplines (see e.g., Airey, 2012; Macaro et al., 2016). Nonetheless, in Aguilar's (2008) study, lecturers also reported providing students with glossaries of terms. This practice, on the one hand, demonstrates their awareness of the language demands faced by their students, and on the other hand, highlights EMI lecturers' particular focus on key technical vocabulary. However, in the Italian context, Picciuolo and Johnson (2020) found that engineering lecturers did not express a need for training in this area, as they considered themselves content experts already familiar with the technical jargon. As Curle et al. (2020, p. 53) pointed out,

> Content lecturers using EMI may be familiar with the subject-specific vocabulary of their disciplines and aim to develop students' literacy in their specific subject fields; however, they might not be aware of other subject-specific features the texts in their disciplines display, e.g. how to develop an argument, present evidence and use quotes (Nesi & Gardner, 2012; Dafouz, 2018; Block & Moncada-Comas, 2019).

In this respect, Klaassen and Räsänen (2006) rightly noted that most NNES teachers believe their English skills are adequate, given their active participation in international communities and conferences. Conversely, Othaman and Saat (2009) found that pre-service science teachers in Malaysia expressed major concerns about their students' poor levels of English and anticipated needing to change their pedagogy as a result, particularly when it comes to "explaining concepts in English" (p. 311). Similarly, in Dearden and Macaro's (2016) study, science teachers were confident about their level of English, believing it to be sufficiently high to handle the teaching of complex constructs. This confidence was partly due to the belief that teaching science and math was easier and required little language, as "[i]n Maths, you are saved by the formulae, and the formulae are true or false in any language", and "[i]n Science, it's probably easier because the number of words you have to use in English is lower" (Dearden & Macaro, 2016, pp. 471-472).

As Dearden and Macaro (2016, pp. 478-479) comment, from the lecturers' replies,

> there is not only an urgent need for research into subject-specific language requirements but clearly a need for teacher professional development given that research in these subjects (Othman & Saat, 2009; Probyn, 2006), albeit in different contexts, has already begun to show that complex language and indeed carefully scaffolded interactive

> learning plays a critical role in understanding content thoroughly. Unlike the views expressed by at least one science teacher, science requires a great deal of language in order to put across concept definition and explanation (Rollnick, 2000; Yassin, Tek, Alimon, Baharom, & Ying, 2010), and it has its own specific genre which is second nature to the teacher but which has to be learnt by the student in order for them to become part of this community of practice (Wingate, 2015).

Nonetheless, Aguilar (2008) also found that, despite being confident in their mastery of the English language, lecturers expressed concerns about their fluency when lecturing in English. Fluency, often linked to speech rate, can reflect how comfortably and naturally a speaker conveys information in a second language.

In this context, Thøgersen & Airey (2011) analyzed the speech rate of a Danish science lecturer teaching in EMI. Using two different metrics, namely "syllables per second" and "mean length of run" – which measures the average number of syllables spoken between pauses – they found that the lecturer's pace did indeed slow down when delivering a lesson in English compared to Danish (L1). However, when later questioned about his speech, the lecturer denied intentionally reducing his speech rate or even being aware of the decrease. Given that the lecturer studied was "a highly experienced lecturer who teaches in English daily" (p. 210), it seems likely that he was unconsciously adjusting his pace to accommodate students' lower proficiency in the L2. Additionally, the researchers suggested that the slower speech rate might also be due to the lecturer's more formal style when teaching in EMI, which is similar to written English. However, despite the slower pace, the researchers found that the amount of content delivered remained consistent. Hincks (2010) also observed a decrease in speaking rate in ELF presentations compared to native (Swedish) presentations. However, she also found a corresponding reduction in information content, suggesting that the slower pace in ELF may indicate the difficulties some speakers face when switching to English.

More recently, Querol-Julián and Amondarain-Garrido (2024) found that in their analysis of lecturer-student interactions during online EMI lectures, the lecturer's periods of "eloquent" silence (Ephratt, 2008; as cited in Querol-Julián & Amondarain-Garrido, 2024, p. 75) – i.e., deliberate pauses in speech for purposes such as emphasis or prompting students' responses – were longer and more frequent in the EMI lesson compared to the Spanish L1 lesson. Students in the EMI classes also exhibited longer periods of silence compared to those in the L1 classes; however, these were not as frequent or extended as the lecturer's pauses. The authors suggest that the lecturer's extended wait times in the EMI lesson may reflect a need to give students more time to process information and formulate their contributions in the additional language. Furthermore, the authors found a correlation between students' silence and the function performed in the interaction episode, with students tending to remain silent for longer during regulative episodes – those involving the organizational aspects of teaching, usually monologic – than during instructional ones, which involve content delivery and are generally more dialogic (Dalton-Puffer, 2007, p. 29; as cited in Querol-Julián & Amondarain-Garrido, 2024, p. 80). Conversely, the lecturer's silence duration remained consistent regardless of the function performed in the episode. As the authors suggest, further research is needed to determine whether the challenges faced by both the lecturer and the students were due to the virtual context, the language of instruction, or both.

Vinke (1995) offered another possible explanation for the slower speech rate observed in EMI lecturers. He noted that experienced engineering lecturers at the Dutch university studied faced several linguistic challenges related to vocabulary, clarity, accuracy of expression, and redundancy. This issue, first identified by Chaudron (1982; as cited in Arden-Close, 1993, p. 256), pertains to the over-elaboration of vocabulary meanings through increased redundancy. Such redundancy may confuse NNES listeners, as they often struggle to determine whether repeated information is being rephrased or if new information is being introduced.

According to Vinke, it is the lecturers' perception of these limitations that, by increasing their workload in terms of preparation time and mental energy (see also Curle et al., 2020, p. 55), may in turn reduce their speech rate, expressiveness, and ability to manage these aspects of teaching, which is likely to have a negative effect on students' learning. These problems affected less experienced lecturers to an even greater extent.

In this respect, Aguilar (2008) observed that, despite lecturers generally having a positive attitude towards the implementation of EMI at their institution – recognizing the positive effect teaching in English had on their fluency and the higher international composition of their classes – some lecturers also expressed discomfort regarding the additional burden that teaching their content in English placed on them. This finding aligns with most research on EMI (see Curle et al., 2020), which has shown that EMI lecturers often teach in English either because "they had been nominated, they had studied abroad, they were proficient English speakers," or "they had simply volunteered" (p. 33). In this context, it is important to repeat that while in "ELF interactions, English is used as the common language of *choice* among speakers who come from different linguacultural backgrounds (Jenkins, 2009)" (Kaur, 2014, p. 215, emphasis added), in EMI, English often functions as a language of *necessity*, driven by education policymakers.

A second area that EMI research has focused on relates to the *pedagogical strategies* needed to convey highly technical content. Although research has yet to establish what constitutes effective EMI pedagogy (Curle et al., 2020, p. 42), several studies on interaction and classroom discourse in EMI have identified recurring challenges faced by lecturers when teaching in EMI, as well as the strategies they – consciously or not – put into practice to overcome these challenges.

In the fields of sciences and engineering, Vinke (1995) found that agricultural science lecturers felt that EMI limited the variety of pedagogical tasks or activities they could offer to students. Similarly, Airey (2011) reported that Swedish lecturers mentioned challenges in incorporating humour (in line with Helm & Guarda, 2015), and providing local context through EMI, particularly when students came from diverse cultural backgrounds. In this respect, Alsop et al. (2013, p. 7) pointed out that:

> Although engineering lecturers around the world may use a common language to deliver the same kind of syllabus for the same broad purpose, engineering lectures are likely to remain both context- and culture-specific. Lectures of all kinds often include pragmatic elements that serve to entertain, instruct, and make key information more memorable. The way in which these features are presented varies from place to place, however, and cultural differences may represent a challenge both to those who attend lectures and to those who deliver them.

Research has also shown that EMI lecturers, to varying degrees of awareness, implement strategies to overcome these pedagogical challenges. In the field of sciences and engineering, for example, Airey et al. (2017) found that a student-centred approach in an EMI program – by facilitating student engagement and active participation in learning – is likely to compensate for any limitations in the lecturer's English proficiency.

Following Mariotti (2021, p. 42), language-related scaffolding techniques include lecturers

> asking several types of questions (Crawford Camiciottoli, 2005; Thompson, 1998), expressing stance through the use of the inclusive "we" (Fortanet-Gómez, 2004; Hansen & Jensen, 1994), producing speech rich in rephrasing and examples (Flowerdew & Miller, 1997), and using discourse markers and signposting (Chaudron & Richards, 1986; Flowerdew & Miller, 1997). Young (1994), in particular, stresses the importance of making information clearly accessible to students and underlines the relevance of redundancy and explicitness for academic didactic spoken discourse.

Additionally, EMI research has shown that lecturers employ various language-related strategies, including "defamiliarizing categories, such as preemptive focus on form (mainly typographical), input enhancement, codeswitching, and humor" (Costa, 2017; as cited in Costa, 2021). Costa (2012) was the first to highlight that despite differing views on the use of the L1 in EMI classrooms – with faculty and students often seeing the L1 as a valuable tool for understanding content, while academic staff may oppose its use due to concerns about excluding international students and violating official policies (see Curle et al., 2020, p. 11) – strategies such as preemptive focus on form (FonF), including code-switching (see e.g. Sahan, 2020) and translanguaging (see e.g., Thai, 2021), along with language-related episodes (LREs; Basturkmen & Hong, 2021), are commonly employed in EMI settings.

In disciplinary contexts, LREs involve brief interruptions in the primary discussion of conceptual content to focus on language, providing valuable opportunities for lecturers to emphasize key vocabulary within the discipline's register, correct students' use or understanding of this vocabulary, and help students become aware of disciplinary language usage or confirm their comprehension (Basturkmen & Hong, 2021, p. 26). Interestingly, Basturkmen and Hong (2021) found that in mathematics classes, LREs tended to address lexical choice – i.e., which word should be used – more than the meaning of the lexical item or its form, such as grammar.

LREs can involve code-switching or translanguaging. Basturkmen and Hong (2021, p. 30) provide an example where the "[l]ecturer draws attention to the meaning of "solvency". The Lecturer firstly provides a meaning ("the company's ability to pay the long-term debt or meet the obligations") and secondly provides the Korean [L1] translation".

Following Cicillini (2023, p. 115), code-switching – i.e., the practice of shifting from an L2/FL to the L1 of the speaker in conversations – can perform several functions when used in instructional settings:

- Facilitate interaction. Lecturers often code-switch for strengthening cooperation between speakers of different languages (Cogo, 2009) and for encouraging students to interact, for example, by specifying one or more interlocutors in an additional language (Klimpfinger, 2007).
- Overcome language problems. A great deal of translanguaging occurs in group work and discussions to encourage communication, especially when the students' proficiency levels are different (Kuteeva, 2020). It may also serve as scaffolding when language problems emerge in class (Adamson & Fujimoto-Adamson, 2021) and as a

way to appeal for assistance, for example, when speakers struggle to recall a certain term (Klimpfinger, 2007).

- Explain concepts. Switching to another language may serve the need to introduce or translate items of vocabulary (Costa, 2012; Tarnopolsky & Goodman, 2012), explain technical terminology (Tarnopolsky & Goodman, 2012), and clarify difficult concepts (Haroon, 2005; Jiang et al., 2019) by using a specific language which is believed to be more effective.
- Explain cultural references. Code-switching is also a way to signal culture, either implicitly (unintentional) or explicitly (intentional) (Cogo, 2009; Klimpfinger, 2007). The speaker's L1 may be useful to rephrase idiomatic expressions, jokes, and anecdotes (Méndez García & Pavón Vázquez, 2012), even though these may not always be understood by the students who do not share the same L1.

In this respect, research has also shown that lecturers are aware of the meaning-making potential of code-switching, which is commonly referred to as translanguaging when used deliberately as a mediation strategy during teaching (see Canagarajah, 2011; García & Leiva, 2014; as cited in Molino, 2023, p. 77).

EMI research has been increasingly influenced by the "multimodal turn in HE" (Lim, 2024, p. 9), which suggests that the scaffolding strategies lecturers use in EMI classes include semiotic resources such as "technologies (Zhao & van Leeuwen, 2014) as well as embodied resources such as speech, writing, posture, gaze (Taylor, 2016), gestures (Lim, 2021b), and positioning and movement (Lim et al., 2012) to design learning experiences for students" (Lim, 2024, p. 12; see also Lin, 2016; Norte & Morell, 2024). Lim (2021) explored the relationship between language and gesture within classroom discourse from a systemic functional perspective, focusing on "intersemiosis" (p. 36). Intersemiosis refers to the interaction between different semiotic modes that cohesively co-occur and semantically complement each other to convey meaning. The study provides insights into the dynamic interplay between language and gesture, showing that the two teachers investigated deliberately made specific semiotic choices to "mitigate the hierarchical distance between the teacher and students" (p. 55), contributing to what Lim refers to as "structured informality" (p. 34). More recently, Moncada-Comas and Sabaté-Dalmau (2024, p. 98) investigated a lecturer-students' interaction in an EMI engineering class by focusing on the lecturer's deployment of multimodal strategies to assist students, particularly those with lower English competence. As the authors point out, a multimodal approach to investigate EMI classroom interaction allows the researcher to understand how lecturers and students:

1. connect knowledge of the world, personal experience and theoretical concepts in a readier manner, particularly with gesture, which "can support thinking processes of students" (Arzarello et al., 2009, p. 107), sustaining and complementing speech;
2. foster a collaborative view of knowledge generation and acquisition (Williams, 2020), particularly to elaborate on, negotiate and extend verbal messages.
3. allow for a less challenging classroom participation which enhances "the possibility of comprehension on the part of students" (McCafferty & Stam, 2008, p. 17); and
4. promote student-to-student/teacher-student rapport, particularly in EMI classrooms where English is a foreign language (Barnett, 1983), including the "expression of emotions" (Kress et al., 2001, p. 74).

By taking a "critical sociolinguistic ethnographic approach" (p. 121), Moncada-Comas and Sabaté-Dalmau (2024) identified three key moves in a single lecturer's multimodal discourse: a *gesture* move, a *gesture-board* move, and a *waiting-time* move. These moves were strategically performed to achieve specific communicative purposes.

> the pedagogical functions of "giving instructions", "monitoring" negotiation of meaning and both content and specialist EFL terminology attainment; "eliciting" participation and collective problem-solving resolution (here, though, as an attempt); and, finally, "reviewing" concepts and notions by means of the emphatic repetition of the same information through different multimodal displays.

As the authors observed (p. 122), the lecturer taught in a small classroom, which allowed her to "make visual contact more frequently with each one of the students" demonstrating her awareness of the importance of engaging students through a wider range of communicative resources.

In engineering, gestures are crucial for illustrating abstract concepts, directing attention, and emphasizing important information. Iconic gestures, which visually represent the shape or action being described, are particularly beneficial for conveying spatial or procedural information (Gullberg, 2006a).

Importantly, EMI research has increasingly focused on pedagogy at the intersection of language – understood as encompassing both verbal and non-verbal resources. As a result, support programs for EMI lecturers emphasize the need for training in "language-aware pedagogical practices in EMI contexts" (Curle et al., 2020, p. 62). In fact, as Molino et al. (2022, p. 111; as cited in Costa, 2023, p. 1) rightly argue, through language, not only do EMI lecturers

> define terms, explain concepts, and give examples, but they also redress misconceptions, guide students through discourse, make sure that learners focus on what is important, and establish meaningful interpersonal relations with them to facilitate the co-construction of meanings.

In this respect, Guarda and Helm (2016, p. 13) rightly observed that "it is the disruption caused by the introduction of a foreign language for teaching and learning that can offer opportunities for reflection and innovation in pedagogy" (2016, p. 13).

Finally, it remains uncertain to what extent the intercultural dimension – identified as the third challenging area for EMI lecturers – can be perceived as an opportunity (Huang & Fang, 2022), particularly for "hard-core" EMI lecturers (Macaro, 2020, p. 264), who teach subjects heavily reliant on mathematical codes and formulas. Curle et al. (2020, p. 35) define this intercultural dimension as "non language related", referring to the challenge of teaching to "student groups with diverse cultural backgrounds, which includes diverse expectations of an academic context as students might have little understanding of the local education system" (p. 35). Nonetheless, research has shown that "[i]nput on the course also explicitly engages students with aspects of disciplinary culture" (Hafner & Miller, 2018, p. 253). The disciplinary culture – i.e., norms and practices shared by members of a group, e.g. professionals, university, with its subspecialism according to the field (Becher, 1994) – affects in turn "(2) textual features of the genre, such as its rhetorical organization into moves and steps; and (3) linguistic features of the genres, especially the way that particular genre moves are realized and how writers use language to engage with their readers" (p. 55).

Research indicates that students enrolled in EMI university programs may feel distanced from their L1 culture (e.g., Kırkgöz, 2009). Similarly, studies have shown that EMI lecturers encounter diverse learning cultures and feel pressured to accommodate or explicitly instruct students on assessment expectations (Mair, 2021), the type of interaction expected – such as questioning (Chang, 2011) – and "appropriate" classroom behaviour, like note-taking (Breeze et al., 2024). In this respect, previous research has highlighted the pervasiveness and significance of evaluation in lectures (see e.g., Deroey & Tavernier, 2011, p. 11). Evaluation refers to "the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about" (Thompson & Hunston, 2000, p. 5). Through such evaluations, lecturers socialize students into the disciplinary knowledge of their academic communities by shaping students' interpretation of statements. As early as the 1990s, Lemke discussed the "unwritten rules of classroom dialogue" and "behaviour" in the science classroom, which are closely linked to the values of the disciplinary community. But there is also another way culture impact on classroom language in ELF interaction. Conducting classroom observations and interviews with lecturers, Arden-Close (1993) examined the language problems that arose when NES chemistry lecturers (American and English) taught Omani students at Sultan Qaboos University in the Sultanate of Oman. He identified four major areas of language problems in these science lectures, three of which have already been mentioned – namely, abstractness, redundancy, and polysemous technical words. The fourth issue relates to the challenge faced by lecturers who do not share the same cultural background as their students: "finding a common range of reference", or, in other words, "finding common areas of experience and knowledge which those words represent" (Arden-Close, 1993, p. 258). As the author comments:

> They had very limited knowledge of the students' culture (Arab culture), limited knowledge of the students' background education in secondary schools and they spent a very limited time with the students. Because of this lack of knowledge, lecturers found it difficult to assess what was difficult for the students.

2.1.2 Students' Views on EMI: Challenges and Experiences

Over the past two decades, research into EMI has expanded significantly, but much of the focus has been on lecturers' perspectives and experiences. It is only in recent years that attention has begun to shift toward understanding students' viewpoints (Bagni, 2021). Similar to research on lecturers, early studies investigating students' perspectives on EMI have primarily employed survey-based approaches. As highlighted in Doiz et al.'s (2019) review, these studies indicate that students generally have a positive perception of EMI. However, limited English proficiency has been recognized as a significant obstacle to fully engaging with the learning experience, especially in non-European contexts (Belhiah & Elhami, 2015; Evans & Morrison, 2011; Kim et al., 2014; Kim & Yoon, 2018), as well as among European STEM students (Wächter & Maiworm, 2014). However, Wächter and Maiworm (2014, p.105) have pointed out that

> institutions have, over time, gotten used to the imperfection of communication in an international classroom, of which they might have originally had too high an expectation. The newly found problem in the view of the Programme Directors these days is rather the "heterogeneity" in the command of English of the students and the difficulties to manage such heterogeneity in the classroom.

Nonetheless, even in these situations, students often consider EMI courses advantageous for improving their English skills, indicating an awareness of the importance of language learning alongside subject knowledge, with the most commonly cited improvements being in general language proficiency, communicative competence (Aguilar & Rodríguez, 2011; Clark, 2017; Tatzl, 2011), as well as the acquisition of field-specific terminology (Ackerley, 2017; Arnó-Macià &

Mancho-Barés, 2015). However, students across various disciplines have also reported challenges, particularly in understanding lectures, developing oral skills, and coping with the increased workload compared to courses taught in their L1 (Ackerley, 2017; Airey, 2009; Tatzl, 2011). Many studies also highlight that students express concerns about the simplification of lecture content (Tatzl, 2011), slower teaching pace (Aguilar & Rodriguez, 2011), and the linguistic competence of lecturers, including their effectiveness in facilitating learning in an L2 (Aguilar & Rodriguez, 2011; Costa & Mariotti, 2017). Additionally, some studies emphasize that student explicitly expressed the need for supportive tools (Aguilar & Rodríguez, 2011; Klaassen, 2001).

In this regard, McKinley and Rose (2022) identify three key areas where support is particularly needed. Firstly, as insufficient English language skills have been highlighted as a major barrier for students learning content in a non-native language (e.g., Hellekjær, 2010; Wong & Wu, 2011; Costa & Coleman, 2012; Hu, Li & Lei, 2014; Belhiah & Elhami, 2015; Macaro, 2018; as cited in McKinley & Rose, 2022, p.2), research has subsequently explored which specific language skills present the greatest difficulties for EMI students. In terms of listening comprehension, Valcke and Pavón (2015) discovered that a teacher's pronunciation significantly affects students' understanding, with many students reporting difficulties comprehending lecturers' accents, which hinders their ability to effectively follow lectures (see also Tange, 2010). Even in countries with high levels of English proficiency, like Norway, comprehension in EMI environments tends to be lower compared to instruction in the native language, partly due to unfamiliar accents and the cognitive demands of processing content in an L2 (Hellekjær, 2010; Hua, 2020; Johnson & Picciuolo, 2023). Moreover, research by Kornder and Mennen (2021) indicates that students' linguistic backgrounds – whether they are monolingual or bilingual – play a significant role in how they perceive NNES lecturers' speech. Students' prior exposure to English-taught courses and their familiarity with non-native English accents also impact their evaluation of teaching quality (Jensen, 2013). For instance, international students are generally less critical of NNES lecturers, while more experienced students show greater tolerance and encounter fewer comprehension issues, regardless of their L1 (Clark, 2017). However, Costa and Mair (2022) found that international students in Italy often struggle more with local accents, whereas Italian native speakers are more lenient.

Students' prior educational background has also been shown to affect EMI students' needs, particularly as regard their productive and receptive vocabulary (Evans & Morrison, 2011; Aizawa & Rose, 2020).

Despite this, students generally prioritize clarity and pronunciation over accent when evaluating lecturers. In this context, rapid speech and strong accents are commonly perceived as obstacles, whereas clear articulation and a moderate speech rate enhance comprehension in EMI settings (Costa & Mariotti, 2021; Dubow et al., 2021). As noted by Dubow et al. (2021), students frequently remarked on how factors such as interaction, lesson pace, speech rate, explanations, and examples positively or negatively influenced their learning.

Additional factors that impact students' lecture comprehension in EMI involve the frequent use of technical vocabulary (Chan, 2015). In this respect, following Macaro (2008, p.265):

> If we turn to the research literature on EMI which has investigated students' self-reported challenges in understanding EMI classes we find that almost invariably at the top of the list comes "understanding vocabulary".

EMI students are inevitably exposed to a broad range of low-frequency words – including both technical and sub-technical vocabulary – in a second language that they need to become familiar with. Building on Macaro's (2008, p. 267) work, research in SLA has established that a listener must recognize at least 95% of the vocabulary in a spoken discourse to fully comprehend it, with the remaining 5% potentially inferred through language learner/user strategies. However, in the EMI classroom, the ability to infer the meaning of unfamiliar words or phrases is strongly influenced by students' prior knowledge of the topic under discussion. Additionally, students may need to selectively focus on either the content or specific aspects of language form – e.g. discourse markers, visual cues. Furthermore, in EMI, the required "lexical coverage" – the extent of vocabulary knowledge needed – may vary across different academic subjects, both in terms of quantity and the nature of the vocabulary involved.

Beyond vocabulary, further factors impacting students' lecture comprehension include the impromptu and informal lecture delivery styles that are more challenging to grasp in an L2 (Evans & Morrison, 2011), and the cognitive burden of listening in an L2, which can result in decreased focus (Hua, 2020).

Regarding speaking skills, research has shown that NNES students often struggle to convey content knowledge in English, which presents a significant challenge in EMI contexts (Kırkgöz, 2009). This issue is particularly concerning because students frequently feel they have not fully grasped a concept until they can express it verbally (Ball & Lindsay, 2013). This implies that if EMI students engage with material passively, their understanding may remain superficial (McKinley & Rose, 2022). Furthermore, students often find it difficult to articulate content in English and to speak in front of their peers, negatively impacting classroom interaction in EMI settings (Kırkgöz, 2009; Kim et al., 2014; Kim & Yoon, 2018). Indeed, studies indicate that EMI courses typically involve less interaction than those conducted in the students' and teachers' first language (Lo & Macaro, 2012; Pun & Macaro, 2019; Thøgersen & Airey, 2011). While this is partly due to EMI students' lower English proficiency and the difficulty in accessing the discipline-specific language required for their fields of study (Dafouz et al., 2018), Klaassen (2001) observed that cultural differences also play a significant role in how students interact in an EMI classroom. For example, students from cultures that place a high value on authority may be less inclined to ask questions or participate in discussions, exacerbating the challenges posed by language barriers.

Writing is another skill that many students struggle with in EMI contexts. Many of these challenges arise from unfamiliarity with academic discourses, genres, and referencing conventions, which can differ significantly between disciplines (e.g., Abouzeid, 2021; Eriksson, 2018; Pessoa et al., 2014; Evans & Morrison, 2011; as cited in McKinley & Rose, 2022, p. 3), as well as with practical skills, such as taking notes in English (Andrade, 2006). Recent research (Block & Mancho-Barés, 2021; Mancho-Bares et al., 2022; Rodriguez Melchor & Walsh, 2022; Sahan & Sahan, 2022; as cited in Gronchi, 2023, p. 5) has shown that in both oral and written exams, content lecturers consistently evaluate language aspects, paying attention to phonology, morpho-syntax, fluency, accuracy, coherence, and register. Consequently, students' difficulties in speaking and writing can have a substantial impact on their academic performance.

Students' unfamiliarity with new and specialized vocabulary also impacts their reading comprehension in EMI contexts (Andrade, 2006; Kırkgöz, 2005; Uchihara & Harada, 2018), which can, in turn, affect their ability to complete assigned readings for EMI courses (Tatzl, 2011). Chan (2014) found that EMI students often rely heavily on coping strategies such as dictionary use and the mental translation of English terms, which can slow down the reading process and may be ineffective.

In line with these findings, Klaassen (2001) discovered that engineering EMI students often struggle with the technical vocabulary and complex language structures required in their courses. This can result in misunderstandings, reduced class participation, and challenges in effectively expressing technical concepts. As a result, Klaassen (2001) found that students may rely heavily on lecture notes or other written materials to compensate for difficulties in understanding spoken English, which can impact their overall engagement and learning process. Similarly, in the field of engineering, Miller (2009) investigated students' preferences regarding lecturers' use of language and teaching techniques. When it came to the use of English in lectures, students emphasized the importance of simplification, clear pronunciation, minimizing lengthy explanations in English when introducing new theoretical concepts, and reducing the use of narratives. Additionally, students expressed a preference for a teaching approach that did not depend solely on language, underscoring the value of visual aids and the use of body language. Similarly, Klaassen (2001) discovered that key factors deemed particularly important by students included explaining new terminology, presenting information in multiple ways, using clear examples, as well as exhibiting liveliness, employing effective gestures, and maintaining eye contact.

2.1.3 Research on EMI Engineering Classroom Discourse

Although EMI has gained global traction, research on EMI classroom discourse is still in its "infant stage" compared to the study of lecturer and student perspectives on EMI (Querol-Julián & Amondarain-Garrido, 2024, p. 71).

In this context, Jablonkai (2021) highlighted the scarcity of corpus linguistic studies focused on analysing discourse in EMI educational settings. Nonetheless, corpus analysis has become a powerful tool for identifying common linguistic patterns in EMI engineering contexts. For instance, Björkman (2008) analyzed group-work interactions and lectures in an engineering course at a Swedish university. Her corpus-based study showed that engineering ELF lecturers frequently minimized redundancy and prioritized functional communication, often overlooking standard forms. However, despite the relatively high frequency of non-standard forms – e.g. double comparatives "*more wider*"; incorrect word formation "*discriminization*"; not marking plural forms "*many hydrogen*"; deviance from standard question word order "*Why it is black?*"; pre- and post- location (in Mauranen, 2010b, it is referred to as left-/right-topic dislocation) “*And the* ***nano-particles*** *they are in the surface area*” or “*Well it is not so emission-free* ***hydropower***” (examples are drawn from Björkman, 2008) which are nonetheless common features in ELF – it was found that communication was not impeded.

More recently, the EMIBO corpus (Johnson & Picciuolo, 2022) is an ongoing collection of transcribed master’s degree lectures delivered in English by Italian lecturers, encompassing various disciplines (including engineering) and lecture formats. All the lecturers were NNES with self-declared English language proficiency levels ranging from B1 to C1 on the Common European Framework of Reference (CEFR), with the majority reporting a C1 level. Additionally, the lecturers had varying years of experience teaching in English, and most had not received any specific teacher training. The corpus is continuously growing as new recordings are made and transcriptions are added. Currently, it contains 21 full lecture sessions from 14 different lecturers in engineering and economics, totaling 36 hours of lectures and just over 200,000 words. Both lecturer and student interactions were annotated using a simplified annotation system based on Jefferson (2004), which included notation for micropauses (.) and overlapping talk [ ]. Discourse disfluencies, such as false starts and hesitations, were lexicalized (e.g., uhm, erm). Similarly, non-verbal actions like laughing or coughing were indicated using angle brackets (e.g., <laugh>). Standard punctuation marks (comma, full stop, question mark) were also used. The corpus is divided into two parts: one includes transcripts from face-to-face (F2F) lectures recorded in audio and video, while the other consists of transcripts from online lectures, which also include written interactions from the chat.

Studies employing the EMIBO corpus have uncovered notable differences between Physical and Social Sciences in terms of lecture styles and interaction patterns. For instance, the use of specialized vocabulary varies significantly between disciplines, with the Physical Sciences corpus showing a predominance of nominalization and

and lexical bundles, as well as quantity bundles (in line with e.g., Biber & Baribieri, 2007).

Further studies have explored the use of personal pronouns in EMI lectures to assess inclusivity. Johnson and Picciuolo (2022) observed that the pronoun "you" was the most frequently used by lecturers across disciplines, indicating an attempt to engage students. However, it often functioned as an impersonal indexical, suggesting that lecturers may not be fully promoting bidirectional dialogic interaction, as noted by Dafouz et al. (2007).

The research has also highlighted differences in questioning strategies. Johnson and Picciuolo (2020) found that while Social Sciences lecturers were more adept at fostering interaction through audience-oriented questions, Physical Sciences lecturers tended to ask content-oriented questions, which were less effective in encouraging student engagement (Crawford Camiciottoli, 2008).

Moreover, the comparison of lecture modes revealed differences in the use of spatial deixis. Picciuolo (2023a) compared the frequency and function of lexical spatial deictic references in the discourse of EMI engineering lecturers across different teaching modalities, finding that the proximal locative adverb "here" was more common in face-to-face classes, anchoring students in the physical space. In contrast, "this" was more frequently used in online lectures, suggesting that in an online environment, engineering lecturers were particularly concerned and attentive about directing students' attention and felt a greater pressure to anchor them to on-screen objects. Similarly, Picciuolo (2023b) found that in online settings, engineering lecturers interact with visuals more frequently than in face-to-face lectures, as even domain-specific sentences and vocabulary become more visually salient through highlighting and pointing. By doing so, lecturers tend to guide students' attention to some of the most critical lecture content, which is particularly crucial in the hard sciences, characterized by high lexical coverage (Dang, 2018).

Since "[t]he function of lectures is to instruct, by presenting information in such a way that a coherent body of information is presented, readily understood, and remembered" (Chaudron & Richards, 1986: 14)" (in Deroey & Tavernier, 2011, p. 10) much research has been devoted to understanding how lecture content is structured when presented to students.

Young (1995) first examined the macro-structures of the monologic discourse of university ESL and EFL lecturers, in the attempt to offer "a more realistic representation of the schematic patterning of lectures to facilitate students' processing of information transmitted in this mode" (p. 159). Motivated by her own experience as ESL teacher at a Canadian university and aware of the difficulty experienced by incoming students in "processing spoken academic discourse" (p. 159) as well as of the fact that "so much of what students are required to learn is transmitted through [expository] lectures" (p. 159), Young (1995) built a corpus of seven two-hour university lectures delivered in English by both NES and NNES lecturers teaching at a Western European university in three different fields – Pure and Applied Sciences, Business and Economics, and Social Sciences. Adopting SF approach – according to which "situational factors generate linguistic choices" (p. 161), Young (1995) identified the macro-structure of this language variety – the monologic discourse at tertiary level – but also the micro-features that make up this structure. At the macro-level, spoken academic monologic lecture are characterized by phases – i.e., "strands of discourse that [...] recur and are interspersed with others resulting in an interweaving of threads as the discourse progresses" (p. 165) – such that "discourse is composed of different topics which are introduced, described, summarized, returned to and are interspersed with other subtopics which are themselves announced, discussed and exemplified" (p. 165). Therefore, Young (1995) first identified the distinct configurations of language choices which reveal the characteristics of each strand. Thus, six phases were identified, as shown in the table below (Table 1).

Macro-structure: metadiscoursal phases (*occurring across disciplines and levels*)	**Function**	**Micro-features (lexico-grammatical cues)**
Discourse Structuring	The speakers identify topics that are about to be covered to facilitate processing by the students. The addressors, in predicting content, ease the burden of comprehension of new information.	- **Verbal group** + nominal group ("**give** an example") - Inclusive **pronouns** ("we", "you") - ***Wh-*** **rhetorical questions** (i.e., students are not expected to answer) alerting students about what is to come) - **Commands** (e.g. "Now uh let's look at this uh simple code again (*lecturer points to diagram and draws box around material*)"; "so today we're going to start to look at a box")
Conclusion	The lecturers identify and classify what has already been discussed (e.g. key terms and ideas of the theories) to ensure that the information is grasped by the students. Its frequency is determined by the number of new points made in any particular discourse.	- Relational processes: verbal group + key terms, e.g., "And that's another general statement that you can make about codes and decoding." - Indicative declarative mood
Evaluation	The lecturers evaluate information which is about to be or has already been transmitted.	- Reiteration of key terms + **judgement** e.g. "So this (*lecturer points to information on the board*) is in effect a **more efficient** code than this one." ". . . **obviously** error detection is a very important function . . ."
Interaction	Lecturers maintain contact with their audience by posing and answering questions	- Polar interrogative questions
Theory / Content	Lecturer present theories, models, and definitions to students.	- Interspersed with metadiscoursal phases and examples
Example	The lecturers illustrate theoretical concepts through concrete examples familiar to students in the audience. Generally, more frequent than the Theory/Content	

Table *1*. Young's (1994) phasal analysis of the macro-structure of academic lectures and its syntactic realization. Examples are taken from the original study.

Drawing from schema theory, which refers to the relationship "between an individual's psychological conception of a form and his or her ability to comprehend and utilize it (Crookes, 1986, p. 59)" (Young, 1994, p. 174), Young emphasized the pedagogical implications of her phasal analysis of university lectures, aiming to facilitate students' processing of information by acquainting them with the correct schematic patterning of lectures. Young particularly emphasizes the usefulness of this approach in assisting "[f]oreign students, particularly those from non-Western cultures, whose educational and cultural backgrounds may differ widely and whose schemata correspondingly may also differ" (p. 174).

Drawing from the British Academic Spoken English (BASE) corpus, Deroey and Tavernier (2011) analyzed the discourse functions and their linguistic realization in 12 lectures across four broad disciplinary groups (Arts and Humanities, Social Studies, Physical Sciences (PS), and Life and Medical Sciences). These lectures were compared according to study levels, degree of interactivity, and class size (small: fewer than 40 students, medium: 40-50 students, and large: more than 50 students). Each function is instantiated by larger stretches of discourse, often made by several utterances, which in turn may contain embedded functions, a in the example provided below:

> (1) **so** <u>for example</u> you could take lithium metal plus **what should we say** **<u>you</u>** could take er ethyl bromide [...] (pslct003) (Deroey & Tavernier, 2011, p. 4)

In this example, *so* serves as a transition marker helping to guide the listener from a previous point to the specific example being introduced. Then, the hesitation device *what should we say* functions to manage lecture delivery. Finally, the personal pronoun *you* serves to establish interactivity. As the authors point out (p. 4), the same stretch of discourse can serve several functions simultaneously, as in:

> (4) I'm not feeling too good so I hope I survive this lecture (ah035)

where it can be interpreted either "as an appeal for silence and cooperation (thus managing the audience) or as creating rapport and so establishing interaction" (p.4).

Thus, the authors acknowledge that some level of subjective interpretation was unavoidable in the pragmatic coding of discourse. They also highlight that their functional analysis was conducted solely on the transcripts, without referring to the accompanying audio or video recordings or consulting the lecturers and students. They correctly argue that this is a common challenge in corpus linguistic research, as it means that important information, such as the lecturers' intentions, the students' background knowledge, and non-verbal communication or prosody, might be overlooked or could influence the researchers' interpretation of the results.

In the following table (Table 2) an overview of lecture functions and subfunctions found in Deroey and Tavernier (2011) is presented, with particular attention to the functional and linguistic patterns found in PS.

Functions	Subfunctions	Linguistic features	E.g.
INFORMING (c.f. Content phase in Young, 2004)	**Describing**	- present tense - lexis reflecting the subject field (in PS: things, models, processes and procedures)	**the lithium starting material** and **the lithium product** <u>are</u> both sensitive to water and oxygen (pslct003)
	Recounting	- past tenses - time indications	Almost absent in PS
	Reporting	- *source* + <u>communicative verb</u> + **reporting focus**. (In PS the reporting focus generally referred to experimental research, models and theorems).	*the theorem* <u>says</u> that **if we take that as the critical region in other words if we decide H-nought is false when the P-value is less than alpha that is precisely a significance test with significance level alpha** (pslct036)
	Interpreting	- verbs (e.g. suggest, notice)	what this narrator is illustrating asking you to **notice** is that [...] (ahlct009)
	Demonstrating	- personal pronoun *you* + directives - deictics (e.g. *this*, *here*, pointing to what was being demonstrated)	Especially common in PS everything is calculated under the assumption of the null hypothesis okay so Q Q this is the one-minus alpha quantile of T that's what Q is so what is the chance now that a random variable by cha-,

			by chance will give you a value greater than or equal to the one-minus alpha quantile (pslct036)
ELABORATING	**Exemplifying**	- exemplification markers (*such as*, *for example*, *for instance* - potentially ambiguous cues such as discourse markers (*so*, *you know*)	[...] **so if I have** a hundred kilometer hundred kilometer grid we could work this out (pslct027)
	Reformulating	multiple linguistic realizations, often not signaled by overt cues (*I mean* or lexical repetitions), but rather stemming from a more implicit shared disciplinary knowledge	**what I mean** by a reductionist view of human behaviour is trying to explain all human behaviour by means of a single explanation (ahlct035) the kidney starts to swell become **edematous** okay starts to swell (lslct011)
EVALUATING	**Indicating disciplinary/ personal values**	- "existing accounts of lexico-grammatical evaluative expressions are not exhaustive and the identification of evaluation often depends on the context" that is "the values of the disciplinary communities" (p. 11-12) - In Ps, evaluations and guide students either on how to use models and methods or on how to assess information based on established knowledge	when you do this reaction in a solvent **the solvent must not provide** any concentration of protons (pslct003)
	Organizing function		I **just** want to include a point there (sslct031)
	Build /Maintain rapport		now what does that tell you does it look even **vaguely familiar** to anyone [laughter] **no I've probably got it wrong** i thought it was something like the equation of relativity (lslct017)
ORGANIZING DISCOURSE	**Orientating**	Discourse orientating cues signaling upcoming discourse	**I'm going to go through** some of the different hierarchy of models the the whole range of models that we can use

			in meteorology (pslct027)
	Structuring	Discourse orientating cues signaling unfolding discourse (order of points)	**I'll talk about** models and their complexity **then I'm going to talk about** some of the waves in the atmosphere (pslct 027)
	Relating	Indicated by prospective and retrospective markers	**we're going to come back** to waves **again** (pslct027) [...] those parties which took that second route **that I mentioned earlier** (sslct031)
INTERACTING	**Regulating interaction** [i.e., eliciting student contributions or providing feedback (rare and generally lecturer-regulated): - check and improve comprehension - involve the audience in the text production - manage the class]		
	Involving the audience	Interactive devices: - pronouns referring to the listeners or including them in the same group as the speaker (*you*, *we*) - content-oriented questions - references to students' experiences	if **we** do an ultrasonogram of the aorta **we** can see here this is an example of a very large aorta and in the middle here **you**'ve got colours (lslct017) now why would Augustus sanction such overt references within literature [...] well as I've said there is a diplomatic element here (ahlct006) **those of you who might be members of a trade union or a political party** [...] will know [...] (sslct031)
	Establishing a relationship with the audience	- increasing the distance in the balance of power	I'll stop in a few minutes and we'll have a short break **but I want to talk first before I stop** a little bit about this issue of

		- decreasing the distance in the balance of power (through e.g. colloquial language, asides, humor, including self-deprecation)	tissue matching (lslct011) as I'm sure **most of you're aware** er s photosynthesis can be restricted by carbon dioxide (lslct040) it can be the actual bone marrow which is taken from the donor's bone **and I have had it I have done it and I tell you it is very painful don't recommend it except for very close friends** (lslct011)
MANAGING THE CLASS	**Managing organizational matters**	Common at the beginning of the lectures	these lecture notes will go up onto the web er within the very next few days (lslct011)
	Managing delivery	Managing the message, the physical environment (e.g. the equipment) and timing. - language commenting on their actions (e.g. *write*, *stop*) - interjections - evaluative language signaling problems (e.g. *oh sorry*, *I'm afraid*)	I'll **write** it out in full "cause then you can see what's happening (pslct003)
	Managing the audience	It constitutes a form of interaction and/or audience management. - directives (generally less direct and more polite)	**what I would like you to do** is try and find out er er try and find out an equation for the vorticity (pslct027)

Table 2. Overview of lecture functions and subfunctions in Deroey and Tavernier (2011). Examples are taken from the original study.

Interestingly, Deroey & Tavernier (2011) observed that in PS, while demonstrating (a subfunction of Informing) was particularly common, recounting (another subfunction of Informing) was almost absent. Additionally, they noted that evaluations by PS lecturers were quite common, serving to socialize students into their disciplinary community. Finally, the authors also noticed that "the frequent absence of explicit lexico-grammatical clues often hampers the functional analysis of lecture discourse and that contextual clues are thus paramount in such analyses" (p. 18).

In contrast to Deroey & Tavernier's (2011) study, which focused on NES lecturers across three subject fields, Kunioshi et al. (2014, 2016) conducted a corpus-based analysis of the pedagogical functions identified specifically in science and engineering lectures. They built their own specialized corpus - OnCAL (Online Corpus of Academic Lectures)[1] – by drawing from transcriptions of lectures on basic science courses (Physics, Chemistry, Biology, and Mathematics) from MIT OpenCourseware (MIT OCW)[2] and more specialized engineering lectures (Information Science, Advanced Mathematics, and Robotics) from Stanford Engineering Everywhere (SEE)[3]. As of January 2014, 430 lecture transcriptions were compiled, corresponding to 3.5 million words, and a total lecture time of 395 hours. Interestingly, the OnCAL corpus was designed to provide both content lecturers and students with "insights into how to better deliver or listen to lectures, respectively, by becoming more aware of the linguistic possibilities through which each pedagogical function may be realized" (p. 4). In a sense, it represents the first data-driven approach to training EMI lecturers and students. Notably, Kunioshi is himself an engineering lecturer. As the authors state:

> The OnCAL interface allows users to easily search for words or expressions and see how they are used in lectures. The interface also allows users to discover other functionalities in an intuitive way, for example, restrict searches to one specific field of study or to a set of undergraduate courses. Links to the recordings of the lectures, which are available at MIT OCW and SEE, are provided to allow users to check pronunciation, rhythm, gestures of the teacher, or how the spoken mode is combined with the use of the blackboard. (Kunioshi et al., 2014, p. 1)

[1] http://www.oncal.sci.waseda.ac.jp

[2] http://ocw.mit.edu/index.htm

[3] http://see.stanford.edu/

Drawing on Dalton-Puffer's (2007) classification of pedagogical functions, the authors identified how lecturers realize these functions through spoken language alone, though they acknowledge that lecturers also perform pedagogical functions through other non-verbal modes. An overview of the pedagogical functions identified by Kunioshi et al. (2014) in science and engineering lectures is provided in the table below (Table 3).

Function	Description	E.g.
Class Management	Announcing, framing, summarizing class content	And you have a material that is an insulator. **And I think you will appreciate why that is in a moment**, but I want to bring Boltzmann's law to bear on the issue of electronic structure in extended networks **like we are talking about today.**
Scientific Facts	Describing relevant discoveries	**And Maxwell – who was credited for** this extra term that he added to Ampere's Law, the displacement current term, **was able to predict** that electromagnetic waves should exist, **he predicted** the existence of radio waves, **which were later discovered** by Hertz, and **that was a great victory for the theory**.
LinkToPrevContent	Linking ideas for promoting continuity along sessions	Keep in mind, **as we've said before**, that during the interphase of the cell cycle, chromosomes are essentially invisible, but during the metaphase of mitosis they become condensed, and on that occasion, individuals noticed a 9–22 translocation.
Examples|Alternatives	Giving examples or alternatives	And then rising up again to theta is equal to pi over two. Do you see that? **This is another way of displaying this property.** This is at constant R – and varying theta.
Using Visuals	Math formulas, graphs, pictures in explanations	I have here a very special flute, open on both sides, **and here you see** the two holes. We will first close the two. That gives us the lowest frequency ...
Cause|Effect	Explaining cause-effect relationship	So, let's work that out of why it's this. Let's take a voltage on – well, that's a wire, the capacitor; there's a common voltage y on all of these things, okay. **As a result**, a current flows here, here and here and the total current that flows out of the capacitor is minus the sum of these currents.

Conditions	Stating conditions of validity	What that means is you have pure acetic acid. And then you dissolve it in water and bring it up to a total volume such that the concentration was 0.1 molar, **assuming that** none of it had been ionized yet.
Analogy	Using analogy to explain a concept	What did you call contours curves that formed that pattern? A saddle point. You called this a saddle point because **it was like** the center of a saddle. It is like a mountain pass. Here you are going up the mountain, say, and here you are going down, the way the contour line is going down. And this is sort of a min and max point. A maximum if you go in that direction and a minimum if you go in that direction, say. Without the arrows on it, **it is like** a saddle point.
Thought Experiment	Using thought experiments to explain new content	Gravity is a conservative force. It's very clear. **Suppose that** I do the work – that I go from A to B in some very strange way. Then it is very clear that the work that I would have done would be + mgh, because my force, of course, is exactly in the opposite direction as gravity.
Emphasis	Giving emphasis/calling attention to the topic	And, by the definition of standard cell potentials that I gave you over there, you can see that what we are getting now, because the electron flow is reversed, our sign is reversed, and so our E zero for the reaction Ag + plus an electron going to silver is equal to 0.8 volts positive. **Notice that** the zinc-zinc2 + couple was negative with respect to the standard hydrogen electrode, but because the electron flow is reversed for silver plus silver redox couple, we now have a positive potential relative to the standard hydrogen electrode.
Elicit Reply	Question to initiate interaction, elicit thinking, or check comprehension	

Table *3.* List of pedagogical functions identified in science and engineering lectures (Kunioshi et al., 2014, p. 4). Note that the function Using Visuals was renamed to "Visuals" in Kunioshi et al., 2016 (p. 297). Examples are provided from Kunioshi et al. (2016).

In their classification of pedagogical functions for science and engineering lectures, Kunioshi et al. (2016) highlight the importance of "Condition" and "Analogy." The "Condition" function is crucial because factors influencing scientific phenomena can behave differently under varying conditions, making it essential for both educators and students to explain and understand these differences through classroom discussions to achieve a deeper understanding. “Analogy” is another important pedagogical tool, used to clarify abstract concepts by connecting them to familiar experiences. Although analogies can sometimes be misleading due to their inherent limitations – and despite some inconsistencies in the literature – there is general agreement that the effectiveness of an analogy strongly depends on participants’ prior knowledge of the subject. Participants with limited knowledge of the source domain from which the analogy is drawn are significantly more likely to develop misconceptions (Wilbers and Duit, 2006). Nonetheless, when used appropriately, analogies can be highly effective in the classroom. It is important to note that, in line with Deroey and Tavernier (2011), the authors also emphasize that pedagogical functions are not mutually exclusive, meaning that a single utterance can fulfil multiple pedagogical functions simultaneously.

As Kunioshi et al. (2016) observed, after the initial release of OnCAL in 2012, a workshop was conducted to introduce the tool to Japanese science and engineering teachers preparing to teach in English. It was noted that when using the corpus, the teachers tended to search for technical terms rather than focusing on discourse markers like “the reason why”. The authors (p. 296) comment on this: “It became obvious that such users were not aware of the importance of investigating rhetorical ways of explaining content when preparing for teaching through English”. As a result, OnCAL was later updated to include instances of NNES lecturers as well as to allow users to view example sentences related to various linguistic functions. For instance, lecturers might employ the pedagogical function “ScientificFacts” to discuss significant historical events related to the development of a scientific concept. These narratives are commonly used in engineering to illustrate a phenomenon being explained or to detail the evolution of a theory (see e.g., Alsop, et al. 2013).

It is worth noting that EMI lecturers in the workshop expressed a particular curiosity about seeing more examples of "Visuals" where teachers read or explained formulas written on the blackboard. This request stemmed from their concern about how to effectively explain the details of a formula in English. Similarly, while the lecturers appreciated the OnCAL tool and found it useful, they suggested that it would be even more helpful if the search results were linked to the corresponding video segments. This would allow them to check pronunciation and observe the context, including teacher gestures, movements, and the type of classroom activity.

The Engineering Lecture Corpus (ELC)[4] (Alsop & Nesi, 2012) is a growing collection of transcripts of EMI engineering lectures delivered at three universities, namely Coventry University in the UK, Universiti Teknologi Malaysia (UTM), and Auckland University of Technology (AUT) allowing for comparison in the use of discourse features across cultural subcorpora. While the OnCAL copurs focuses on the pedagogical functions – the educational and instructional roles that different parts of a lecture serve, the ELC corpus focuses on the pragmatic aspects of lecture content – how language is used to achieve specific communication goals. Drawing from the MICASE pragmatic tagset (Maynard & Leicher, 2007) and pragmatic inventory (Simpson-Vlach & Leicher, 2006), the ELC has been refined over time (2009-2014), indicating an evolving understanding of how these pragmatic elements function in academic lectures. Their final set of pragmatic annotation tags includes five elemental categories, with various sub-categories attributed, as shown in the table below (Table 4, taken from Alsop, 2016).

---

[4]https://www.coventry.ac.uk/research/research-directories/current-projects/2015/engineering-lecture-corpus-elc/

Element	Attribute
Prayer	(only occurs in the Malaysian sub-corpus)
Housekeeping	Assigning homework; logistics/announcements; returning or going over homework or an exam; reviewing for an exam.
Explaining	defining, reasoning, translating, equating
Summary	review previous/current lecture content
	preview current/future lecture content
Story	Anecdote, exemplum, narrative, recount, story-likes/scenarios
Humour	Bawdy, black, disparagement, irony/sarcasm, jokes, mock threat/playful, teasing, wordplay

Table *4*. Overview of elements and attributes within ELC (from Alsop, 2016, p. 70)

As Alsop (2016) observed, the *explaining* element is extremely widespread in ELC (which is supported by its presence in the MICASE tagset). This is not surprising, given that explaining is one of the primary purposes of lectures, along with presenting new information (Alsop, 2016). Nonetheless, Alsop (2016) leave this element out of her study, while focusing in particular on Summary, and Humour. Findings from Alsop's (2016) study showed that summary was the most common element in terms of token count and instances and recurred throughout lectures, particularly the attribute "preview current lecture content", although it tended to be brief, showing the shortest average token count of all attributes. Conversely, the attribute "review previous lecture content" showed the longest average word count per instance of all attributes. The attribute "humour" was found to be the least common element, with the "playful" being the most common element. Nonetheless. discrepancies were found across sub-corpus – bawdry was absent in the Malaysian sub-corpus, disparaging and irony occurred more frequently in the UK corpus, self-deprecating was common in the New Zealand sub-corpus, and joke occurring only in one lecture – thus showing that cultural context influences how commonly occurring discourse features are linguistically expressed in engineering lectures.

Further studies from the ELC corpus (Alsop & Nesi, 2012; Alsop et al., 2013; Nesi & Alsop, 2021; Picavet et al., 2023) have investigated the "Story" attribute in more detail, revealing that narratives frequently occur in engineering EMI lectures. This suggests that the inclusion of stories is a strategy used by EMI engineering lecturers to enhance student engagement and support the acquisition and retention of lecture content.

More recently, Nesi and Alsop (2021) identified narrative sequences lacking a consistent progression in past tense, in line with Easton's (2016) "character-driven story" (in Nesi & Alsop 2021, p. 3). Nesi and Alsop (2021) refer to these instances as "scenarios" which they describe as hypothetical extraneous events told by the lecturer which have some analogy with the lecture content, and are generally introduced by "content conditionals" (Nesi & Alsop 2021, p. 9), that is *if-clauses* and multi-word units such as *let's say*, *imagine*, *suppose*, *you know if*. In their data, scenarios were found to be more frequently used to explain complex technical terms, as a "vocabulary elaboration" technique (first observed by Chaudron, 1982; as cited in Nesi & Alsop, 2021, p. 16) as in the following example from ELC where the EMI Malaysian lecturer explains the concept of *reversible process*:

> That means if I am you know if I wear nice perfume today and somebody smell you can just smell it and come back to me and I can reverse it back I don't think so (ELC 2018). (Alsop & Nesi 2021, p. 7)

Nesi and Alsop (2021, p. 7-8, emphasis added) emphasize the importance of investigating stories in lecture discourse, aiming to increase learners' awareness of their use and to support EMI lecturers who want to incorporate scenarios into their teaching. When lecturing in a less familiar language, there may be a tendency to stick to textbook material and avoid imaginative discussions (what Nesi and Alsop refer to as "hypothetical scenarios") that are often left for non-academic settings. However, similar to traditional storytelling, the ability to present scenarios is a valuable skill for lecturers, as it can aid students in grasping complex concepts and engaging with course content in a manner that aligns with the discipline.

More recently, Picavet et al. (2023) identified and categorized stories in EMI lecturers' discourse from a corpus of 60 hours of engineering lectures delivered in France to international MA students by L2-English lecturers of diverse nationalities. The study found that EMI lecturers rarely incorporated storytelling into their lectures, mostly using it in the form of "scenarios" introduced with compelling imperatives like *imagine* or *suppose*, and by sharing personal or professional anecdotes. However, the researchers noted that EMI lecturers made efforts to engage their audience through various strategies, such as using interactive language and employing direct-address personal pronouns. Some lecturers also displayed divergent forms of ELF in their discourse. Despite potential language proficiency challenges, Picavet and his colleagues argue that the effectiveness of storytelling in lectures does not rely solely on perfect language mastery. They advocate for incorporating storytelling into EMI lecturer support initiatives as an essential part of the EMI lecturer skill set.

The studies mentioned so far, those relying on the OnCAL corpus and those from the ELC corpus, take two different perspectives on the communicative functions of EMI engineering lectures. The first focuses more on what is being taught (pedagogical function), while the second examines how it is being communicated (pragmatic function). This divergence is not surprising, given that Professor Kunioshi is an engineering lecturer himself. As such, it is likely that he is more interested in investigating the instructional roles that different parts of a lecture serve, using a metalanguage closer to that of content lecturers (e.g., "analogy", "condition"). Conversely, studies from the ELC emphasize how language is used to achieve these specific instructional goals, analyzing its lexico-grammatical realizations. Nonetheless, there are clear overlaps between these approaches. A less evident juxtaposition concerns what Kunioshi et al. (2016) refer to as "analogy" and what Alsop (2016) refer to as "story-likes", and then as "scenarios" in Nesi and Alsop, (2021). Story-likes or scenarios are "stories based on hypotheses or predictions of future events [which] tend to be unresolved and focus on scientific judgment, often involve[ing] an analogy" (Alsop, 2016, p. 328). These events are not situated in the past, and generally use hypotheses or predictions of future events to make general claims of relevance to engineering, as in the example of "vocabulary elaboration" provided above. Alsop (2016, p. 270) explicitly mentioned that story-likes often consist of analogies in order to explain engineering concepts, as in the following example:

> **resistance** occurs when the electrons are moving through a conductor and they bash into an atom when the electron hits an atom it gives off energy it causes there to be power loss and the more collisions that occur the greater the resistance so you can compare a conductor to a crowded room if you have a crowded room and you want to walk across a crowded room it's very likely you're going to bump into somebody and the more crowded the room is the more collisions will occur **and that's like high resistance** you see **high resistance** is when the room is really crowded **and you get heaps of collisions a low resistance** is when there's not many people around and you can walk through with only very few collisions *so tha- that's a sort of um er* a very non physical physics people would hate what I've just said but I think it gives you an idea of what resistance actually is *it's um* **it's the power loss that occurs due to collisions between electrons**. (Alsop, 2016, p. 270; emphasis added)

In this example of a "story-like" structure, the lecturer introduces the target concept (marked in **bold**) – i.e., "resistance" – followed by a definition. The definition is then re-elaborated through the introduction of an analogy (underlined), which uses a source domain – e.g., a party setting, assumed to be familiar to the students. The definition and analogies are interspersed and occur three times within the same utterance.

These instances are also interspersed with an aside – introduced by a combination of disfluencies (interruption and repetition, followed by two filled pauses) – consisting of a self-deprecating element (marked in red), followed by a "comment on genre" (Ädel, 2023; marked in green). This is then followed by a metadiscourse unit aimed at managing the message (Ädel, 2023; marked in blue). A further disfluency marks the end of the aside, with the lecturer returning to the target definition of the concept of resistance.

Given the purpose of the study presented in this thesis, the taxonomy used to classify pedagogical functions within the corpus largely draws from the pragmatic function taxonomy developed by Alsop (2016) and Nesi and Alsop (2021). However, by integrating these two taxonomies, a more comprehensive framework for analyzing engineering academic lectures can be achieved. The specific taxonomy employed in this study is detailed in the Research Methodology section.

2.2 (Dis)fluencies in Speech and Communicative Strategies

Disfluency (henceforth DF) is a broad concept that has been extensively studied over the past sixty years by speech pathologists, linguists, phoneticians, and psycholinguists (for a comprehensive review see Shriberg, 1994; Graziano & Gullberg, 2018; Kosmala, 2021). A foundational definition characterizes DF as "any deviation from ideal speech delivery" (Ferreira & Bailey, 2004, p. 234; cited in Kosmala, 2021, p. 70), where "ideal" denotes speech free from disruptions (e.g. interruptions, pauses) that break the speech signal, interrupting its smooth flow. Initially regarded as "noisy or irregular events" (Shriberg, 1994, p. 1), research dating back to the late 1950s began introducing formal categorizations of DFs (DFs). Since then, various terms have been proposed to classify DFs, though these terms are often applied to what appear to be the same phenomena. Building on Kosmala's (2021, p. 74) review, in SLA research, fluency is typically linked to proficiency levels, the temporal characteristics of speech, and the native-like pace of L1 speakers. DF (DF), on the other hand, is often seen as an indication of communication difficulties arising from language-related challenges. In Psycholinguistics, DF is a broader term used to describe non-fluent speech resulting from various cognitive processes such as planning, lexical retrieval, and self-repair. However, in fields like Conversation Analysis (CA) and interactional linguistics, the term DF is generally avoided in favor of terms like repair, recasts, pauses, and fillers such as "uhm." Although these terms do not fully capture the range of markers relevant to these phenomena, the diverse definitions reflect varying perspectives on fluency and DF. This mirrors the way non-native speech is often compared to native-like speech, resulting in a deficit-oriented view of L2 speech. CA and interactional linguistics acknowledge that fluency and DF are not solely linked to ideal speech delivery and temporal aspects. Instead, they are seen as strategies through which speakers use gestures and DFs to maintain conversational flow, functioning as part of broader problem-solving and communication strategies (see e.g., Peltonen, 2019).

Same-turn DFs – instances where a speaker produces speech disruptions within a single conversational turn – typically labelled in the literature (e.g. Shriberg, 1994; Lickley, 2015) are summarized in the table below (Table 5), with examples provided from the corpus analyzed in this study.

DF type	Description	Examples
Filled pauses	(e.g., "um," "uh" – here transcribed as "erm")	So this is: a **erm** a picture of the profile,
Unfilled pauses	(silent pauses – here transcribed with (.))	because otherwise: there is a **(.)** a" an an impact on **(.)** the fluvial ecosystem
Repetitions	(of words, syllables, or sounds)	the finish **the finish** date would be May 27,
Interruptions	(i.e., truncated words)	and therefore you may **s'**, you may wonder
False starts	(i.e., when utterances are aborted)	I get the figure. **We got to** (.) **Okay, this is** the figure.
Lengthening	(elongation of sounds, here transcribed with colon ":")	because **otherwise:** there is a (.) a" an an impact on (.) the fluvial ecosystem
Combinations	Sequences involving at least two different types of DFs occurring consecutively (see Kosmala, 2021; Graziano & Gullberg 2018).	And then you have **this: (.) erm** flashing loose way

Table 5. Summary of the DFs described in Shriberg (1994)

2.2.1 Filled Pauses

In contrast to classification systems that have questioned whether filled pauses should be considered linguistic elements, Shriberg (1994) identifies them as such. Supporting this perspective, earlier research (see Shriberg 1994, p. 24) has demonstrated that: 1) filled pauses exhibit variation across languages, with their vowel quality changing in predictable ways according to the vowel inventory of each language; 2) the intonation of filled pauses occurring within a clause aligns predictably with the intonation patterns of surrounding words; and 3) the forms "um" and "uh" differ in terms of the length of the subsequent silent pause, indicating that speakers may choose between these two forms based on the amount of time they require to pause.

2.2.2 Unfilled pauses

Unfilled (or silent) pauses frequently occur at clause boundaries or following discourse markers, and it has been proposed that this happens due to the cognitive effort involved in planning the upcoming clause (see e.g., Graziano & Gullberg 2018). Conversely, intra-clausal occurrences of unfilled pauses are seen as self-interruptions made by the speaker upon detecting a problem, allowing time to plan the repair (Levelt, 1983, as cited in Shriberg, 1994, p. 23). According to Levelt's "Main Interruption Rule," speakers may cut themselves off at any point, even within a word, and such within-word cut-offs are not constrained by morphological or syllable boundaries. In this context, unfilled pauses play an important role in both speech production and perception. Howell and Young (1991, as cited in Shriberg, 1994, p. 23) found that, by marking an increase of information, unfilled pauses may help listeners identify the onset of the repair region. In adult speakers without any known speech or language disorders, a planning pause is generally considered to last around 200 milliseconds. However, a lower rate of DFs has been observed in cases where subjects engage in preplanning or rehearsal (Goldman-Eisler, 1968; as cited in Shriberg, 1994).

2.2.3 Repetition

According to Dickerson (1971; as cited in Shriberg, 1994) repetitions (of words, syllables, or sounds) have two potential roles in the speech of nonnative English speakers: 1) repetitions may serve as a "pause device" used to gain time for lexical search; 2) but they can also help "bridge the gap" following an extended pause.

2.2.4 Interruption

Interruptions (referred to as "fragmets" in Shriberg, 1994) are instances of mid-word or mid-syllable cut-off in speech. It has been observed that interruptions always mark the right edge of the RM, and this is particularly helpful for automatic correction of DF, as interruptions specifies the right boundary of material to be deleted, as in this example:

Now, the way they are designed, of course, we have to: what,	what	we have to design?
RM	IM	RR

2.2.5 False starts

False starts involve sentence-initial deletions or aborted words that are challenging to detect because they often lack clear cues of DF, such as repeated or substituted words or similar syntactic units in succession (Shriberg, 1994, p. 121). Most false starts are simply followed by restarts, typically in the form of repetitions occurring in close proximity to the aborted syntactic unit, as in this example: *assuming that you have a particle of sand that is entering into the sand trap (.)* ***with (.) by*** *(.) staying at the free surface*. Conversely, recognizing false starts involving changed words – as in *I think that () okay, if I: go to the 30 of May* – is more difficult, although previous studies have shown that these can be acoustically distinguished by a pause – in the example above transcribed with rounded brackets () – usually less than 400 ms (O'shaughnessy, 1994).

2.2.6 Lengthening

Although Shriberg (1994) does not explicitly classify sound lengthening (also known as prolongation or elongation) in her typology of DFs, Eklund's (2004) work offers an overview of prolongations in Swedish and other languages, highlighting consistent patterns in the segmental characteristics of extended lexical material, thereby supporting the recognition of lengthening as a distinct category. Nevertheless, Shriberg (1994) does differentiate between "hesitation lengthenings" (i.e., prolongation DFs) and "lengthenings associated with accentuation or prosodic phrasing" (p. 186). This distinction is important because different types of lengthened syllables convey different information; for instance, accented syllables usually carry high semantic value, while lengthenings tend to occur on less significant words. In this respect, according to Munoz et al. (2014), prolongations frequently serve a speech management function and are often associated with specific lexical items, such as functional words with extended vowels in contexts where vowel reduction or elision would typically be expected. Furthermore, this distinction is particularly relevant to the scope of this thesis, as pronunciation issues typical of Italian L1 speakers of English often include final lengthening – i.e., an increase in segmental duration at the right edge of prevocalic consonant-ending words – accompanied by the addition of a *schwa* sound between and at the end of words (Wheelock, 2016) – e.g., "masses" pronounced as /ˈmæsɪzə/ and "rates" as /reɪtsə/. This phenomenon occurs because Italian is a language in which words typically end in a vowel. In the methodology section, we will elaborate further on this point; however, as a preview, to differentiate DF lengthening from the lengthening phonetic effect typically found in monolingual Italian speakers, we only considered durations greater than 28 milliseconds (see Beller-Marino, 2014).

In addition, as noted by Shriberg (1994, p. 7), studies in linguistics, conversation analysis, psycholinguistics, and computational linguistics have independently observed that most same-turn DFs typically share a similar surface structure. Although there is consensus regarding this structure, a wide range of terms has been used to describe the corresponding regions. Therefore, we will adopt the standardized terms introduced by Shriberg (1994), as illustrated in Figure 2.

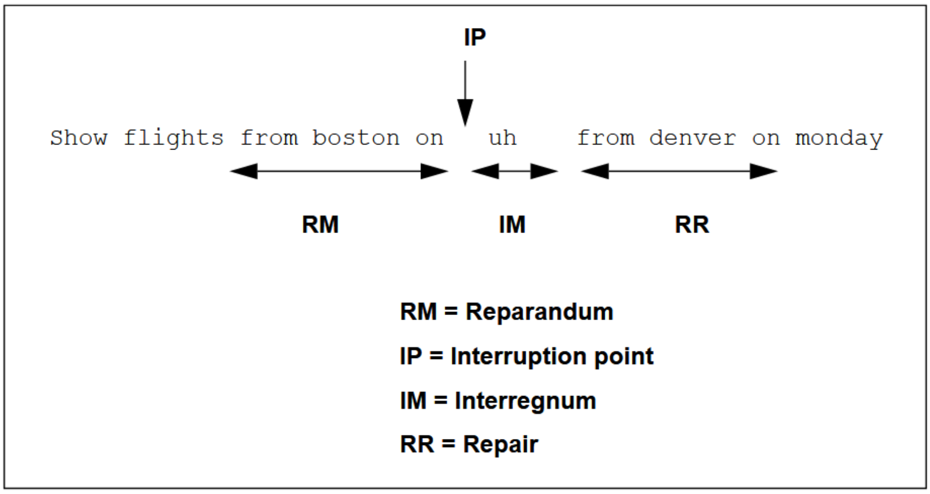

Figure 2. Terminology of DF regions in Shriberg (1994, p. 8).

The reparandum (RM) is the segment that needs correction. The interruption point (IP) is where the speaker halts their speech to fix the previously spoken content, marking the boundary between disfluent and fluent speech. The interregnum (IM) is an optional element that may include silent pauses, filled pauses, or overt editing phrases such as “I mean”, “I’m sorry” or “no.” Finally, the repair involves the corrected version of the linguistic content.

In addition, it is well-established that each of these regions exhibits distinct acoustic characteristics that differentiate them (see Moniz et al., 2014 for a comprehensive review). Specifically, an “edit signal” mechanism exists, where speakers signal an impending correction to their listeners. This signal is conveyed through features like repetition and speech fragments, as well as glottalization, co-articulatory gestures, and voice quality attributes such as jitter (variations in pitch frequency) within the reparandum, along with notably different pause durations compared to fluent boundaries. Additionally, the repair phase may involve shifts such as increased pitch and energy.

The example provided below, extracted from the corpus analyzed in this thesis, illustrates a “compound” DF (DF) structure. It features an unfilled pause followed by an interruption and a filled pause, which are then followed by a repair.

The flood flow: is: a flow that	**(.)hap" erm**	occurs (.) once every 30, 50, 100 years.
RM	IM	RR

These segments are contiguous, and removing a continuous segment containing the *error* (in this case the filled pause "erm") results in the intended utterance.

In this context, Kosmala (2021) highlights that the term "DF" is problematic because it suggests a pathological condition, i.e., a speech disorder, notably stuttering. However, several authors have emphasized that DFs are not always related to uncertainty or production challenges. In L1 speech, DFs can fulfill different roles in speech, such as initiating a conversational turn, holding the floor, gaining time during planning, or signaling discourse structure (see Kosmala 2021, p. 71-72, for a comprehensive review). Likewise, in L2 speech, the term "fluency" can be misleading or unsuitable, as Lennon (1990, p. 292; as cited in Kosmala, p. 72) argues:

> It is often assumed that the fluency target of the language learner is "native-like levels." However, a moment's reflection shows that the idea of monolithic and unitary fluency for native speakers is mythical. Native speakers clearly differ among themselves in fluency, and, more particularly, any individual native speaker may be more or less fluent according to topic, interlocutor, situation, "noise," stress, and other factors.

Native speakers naturally exhibit DFs (DFs). Studies exploring the cognitive mechanisms behind these DFs suggest they may signal cognitive effort due to processing challenges, retrieval difficulties, or speaker uncertainty (see Kosmala, 2021). Furthermore, previous research has shown that L1 and L2 DFs are closely related, as they arise from similar cognitive and social processes across languages (Derwing et al., 2009; Lopez-Ozieblo, 2019).

However, the presence of DFs when speaking a second language does not necessarily indicate poor proficiency. In fact, DFs, along with discourse markers and backchanneling, are also associated with strategic language use, with moderate use of these features considered a strong indicator of conversational fluency – a key element of spoken language (Gürbüz, 2017, as cited in Kosmala, 2021, p. 72). The same forms can actually function as markers of both fluency and DF, depending on their duration, contributing factors, and distribution within the micro and macro contexts. Consequently, researchers have adopted a functionally ambivalent perspective on *(dis)fluency* – with the "*dis*" in brackets – recognizing its dual role.

However, differences between native and non-native speech are well established, with L2 speech typically characterized by slower speech rates and more frequent pauses (Tavakoli, 2011). Research indicates that L2 speakers produce more mid-clause pauses and fewer end-clause pauses compared to native speakers (Tavakoli, 2011), and they more frequently employ non-juncture pauses – i.e., pauses occurring within syntactic units – reflecting planning challenges (Cenoz, 1998).

2.3 (Dis)fluencies in EMI

Research on DFs in the EMI context remains limited. Recently, Martin-Rubió and Diert-Boté (2023) investigated differences in fluency measures and pronunciation accuracy among EMI lecturers at varying CEFR levels. Interestingly, their study diverges from previous research on lecturers' speech rate in EMI settings, which typically focuses on calculating the total number of words or syllables uttered by the lecturer divided by the total lecture time in minutes, resulting in words per minute (wpm) or syllables per second (sps) (Tauroza & Allison, 1990; Hincks, 2010; Thøgersen & Airey, 2011; Johnson & Picciuolo, 2022). Instead, they employed three distinct fluency rate measures in their analysis:

> (1) Mean Syllables per Run (or MSR), which represents the average numbers of syllables produced between pauses (both filled and silent) and calculated by dividing the total number of syllables by the number of bp-units [i.e., between-pauses unites] (or runs);
>
> (2) Rate of Speech Time (ROST), which measured the speed at which syllables are delivered, and results from dividing the syllables by the Speech Time; and
>
> (3) the three Ratios – Speech Time Ratio (STR), Silent Pause Time Ratio (SPTR), and Filled Pause Time Ratio (FPTR) – which indicate the proportion of time (in percentages) the lecturer spends delivering meaningful syllables, silent-pausing, or filled-pausing, respectively. (Martin-Rubió & Diert-Boté, 2023, p. 43)

Their precise methodological approach enabled them to identify differences between two lecturers with B2 and C1/C1- proficiency levels on the CEFR scale, respectively. Despite both producing approximately 500 syllables on average, the C1 lecturer spoke slightly faster (4.2 syllables per second) by using filled pauses less than 1% of the time, whereas the B2 lecturer spoke at 3.7 syllables per second and used filled pauses around 13% of the time on average.

Interestingly, the authors also examined common pronunciation deviations from the standard. They found that STEM lecturers with lower levels of English proficiency were more likely to mispronounce specialized terms compared to their colleagues in the Social Sciences. The study concluded that technical terms are more prevalent in STEM classroom discourse, while Social Sciences lecturers tend to rely more on everyday language. Consequently, the authors suggested that STEM lecturers teaching in EMI settings should focus particularly on the accurate pronunciation of specialized terms to prevent isolated pronunciation errors that might cause confusion and increase cognitive effort. However, despite these issues, the study noted that overall message comprehensibility was not significantly affected.

2.3.1 Pragmatic Functions of DFs and How They Are Interpreted by Students

As discussed in the previous section, the literature suggests that the primary function of DFs is largely related to planning, monitoring, and managing speech processes (see Kosmala, 2021). However, DFs also serve pragmatic functions when speakers use them for intersubjective or discursive purposes. In this regard, DFs can fulfill various pragmatic roles, including planning/structuring speech, expressing uncertainty, managing the speaker-listener relationship, facilitating turn-taking, and handling choice- or change-related mechanisms in speech (see Kosmala, 2021, for a comprehensive review). The following table (Table 6), adapted from Kosmala (2021, p. 78), outlines the key pragmatic functions performed by DFs.

Category	Description	Criteria
Speech management (SM)	Basic feature of DFs. Related to speech processes (planning, repair, lexical search).	- occur in initial position, - occur before content words.
Discursive structuring (DS)	Discursive contexts in which DFs are used to structure, mark, punctuate, emphasize parts of speech, similar to discourse markers.	- occur in initial position, - co-occur with discourse markers which serve ideational and sequential functions, - are accompanied by parsing gestures.
Interactive/communicative	Interactional/communicative contexts in which (Dis)S contribute to the flow and sequencing of the interaction by indexing a stance, reacting to what the other one is saying, turning to the interlocutor, holding/yielding a turn etc.	- occur during dialogic sequences (i.e., question/answer), - co-occur with interpersonal discourse markers (*you know*, *you know what I mean*), interactional gestures, and/or a gaze towards the interlocutor.
Uncertainty	Contexts in which the speakers overtly display/signal their uncertainty.	- occur with epistemic stance markers (*I'm not sure*, *I don't know*), - and/or with gestures (frown, thinking face) in order to save face.

Table 6. Pragmatic functions of DFs (source: Kosmala, 2021, p. 78)

In addition, psycholinguistic studies (e.g., Brennan & Schober, 2001) have shown that sentences containing DFs tend to be either more complex or associated with more demanding tasks. In this context, Munoz et al. (2014) compared the frequency of DFs in the interactions of native Portuguese speakers across two different settings: academic lectures and dialogues. They observed that, firstly, sentence-like units (SUs; whether fluent or containing DFs) in lectures generally include more words than those in dialogues, indicating that lectures are less dynamic, with information delivered at a slower pace. Secondly, after analyzing three different measures – articulation rate, speech rate, and phonation ratio – they found that, despite speakers producing more DFs per SU in dialogues, there was no significant difference in the overall speech rate between the two corpora. They suggested that this could be attributed to the inherent differences between dialogic and monologic communication, which are typical of expository lectures. As they rightly argue, in dialogues, there is a direct cooperative exchange between two interlocutors, whereas in lectures, the lecturer predominantly holds the floor, needing to elaborate on concepts in detail, often using paraphrases, explanatory sentences, and examples to illustrate theoretical ideas. This suggests a higher cognitive load for lecturers, as they alone are responsible for repairs and cannot rely on negotiation with students, especially given the high hierarchical asymmetry characteristic of lectures (Anderson & Ciliberti, 2002).

Thirdly, they found that lectures contain proportionally more silent pauses than dialogues, aligning with the multifunctional role of silent pauses used by teachers, such as to give the floor, emphasize key points, or encourage students to reflect on the presented topic before introducing a new one. Notably, they found that silent pauses are the most frequent type in both corpora, consistent with previous research (e.g., Shriberg, 1994; Eklund, 2004), followed by complex sequences – mostly repetitions and substitutions used for lexical search – and repetitions. In addition, lecturers used interruptions (sometimes referred to as "fragments") less frequently than speakers in dialogues, highlighting the need to preserve word integrity (Levelt, 1989). Both settings also shown infrequent use of false starts (sometimes referred to as "deletions"), as these are more cognitively demanding to process (Fox-Tree, 1995; as cited in Munoz et al., 2014). Interestingly, Munoz et al. (2014) also found that lecturers tend to provide students with prosodic cues as signals for most types of DFs and their surrounding contexts. If lecturers aim to maintain the congruity and continuity of the message during instances of DF, then how does this impact comprehension?

DFs can also be examined from the listener's perspective. Traditionally, linguists and psycholinguists have focused on DFs in terms of production, studying how and why speakers produce them. In contrast, computational linguists have approached DFs from a recognition angle, often aiming to improve machine processing of spontaneous speech. Studies in linguistics and psycholinguistics (Fox Tree, 1995; Brennan & Williams, 1993; as cited in Brennan & Schober, 2001) explored DFs from the listener's viewpoint, suggesting that the information within DFs can help listeners manage disruptions in spontaneous speech and potentially infer speakers' intentions by recognizing patterns in speech errors. Cues are not limited to specific types of DFs but also include linguistic features of the structured segments within a disfluent event, particularly the transition back to fluency, which is crucial for understanding the message.

Brennan and Schober (2001) further contributed to this inquiry by examining how human listeners cope with disfluent speech. They asked listeners to identify a unique referent on a display in response to both fluent utterances (e.g., "Move to the yellow square") and disfluent versions (e.g., "Move to the pur- <uh> yellow square," "Move to the purple <pause> yellow square," and "Move to the <pause> yellow square", among others). Their findings revealed that listeners responded more quickly to target words (i.e., repairs) following DFs with longer edit intervals compared to cases where DFs were absent. Although longer intervals typically involve fillers, the study shown that when the interval before the repair was controlled, the presence of a filler made little difference.

The authors caution that the finding of a "DF advantage" does not suggest that DFs are preferable to fluent speech. Fluent utterances had lower error rates overall, underscoring that fluency remains desirable from a listener's perspective. They also clarify that their results should not be seen as guidelines for speakers, as there is no evidence that speakers deliberately choose to be disfluent (Brennan & Schober, 2001, p. 295). In the following section, we will instead focus on instances where speakers employ deliberate communicative strategies to address linguistic challenges and ensure clear communication.

2.3.2 Communicative Strategies to Overcome Linguistic Challenges

Thus far, we have observed that speakers make spontaneous repairs when encountering DFs in their speech. This prompts the question: what specific types of repairs do they perform, and to what extent are these repairs deliberate? In the following section, we will examine the different types of repairs, while considering whether these actions are carried out with conscious intent. By analysing these strategies, we can gain deeper insights into how speakers address communication challenges and maintain clarity in their interactions.

Research in Conversational Analysis (hereafter referred to as CA) has focused on uncovering the systematic principles that structure the exchanges in naturally occurring conversations. A key aspect of CA research is its emic perspective, which involves examining and uncovering the systematic properties that emerge as speakers and listeners make sense of each other's actions during naturally occurring conversations, turn by turn. This includes examining how participants rely on the full range of linguistic and non-linguistic resources available to them to demonstrate their understanding of one another's talk (Huth, 2011).

One key feature of talk-in-interaction recognized by conversation analysts is the concept of repair. As Schegloff et al. (1977, p. 381) put it:

> If language is composed of systems of rules which are integrated, then it will have sources of troubles related to the modes of their integration (at the least). And if it has intrinsic sources of trouble, then it will have mechanism for dealing with them intrinsically. An adequate theory of the organization of natural language will need to depict how a natural language handles its intrinsic trouble. Such a theory will, then, need an account of the organization of repair.

Repair is a regular phenomenon in conversation where speakers can identify and potentially resolve issues related to speaking, hearing, or understanding.

Repair can be analyzed based on its position relative to the trouble source in turn constructional units (henceforth referred to as TCUs), who initiates it, who resolves it, and whether the attempt is successful. Accordingly, repair is classified (Schegloff, et al., 1977; Schegloff, 2000) into four categories depending on who initiates and who resolves the repair, namely: self-initiated self-repair (SISR), Other-initiated Self-repair (OISR), Self-initiated Other-repair (SIOR), and Other-initiated Other-repair (OIOR). First, in SISR the repair is initiated by the same speaker of the trouble source. For example:

> She was givin me a:ll the people that were go:ne **this yea:r** I mean **this quarter** y" //know
> (Schegloff et al., 1977, p. 364)

Second, in OISR the recipient indicates a problem in the talk and the speaker resolves the problem, as in:

> Ken: Is Al here today?
> **Dan**: Yeah.
> **Roger**: He <u>is</u>? Hh eh heh
> **Dan**: Well he **was**. (Schegloff et al., 1977, p. 364)

Third, in SIOR the repair is initiated by the recipient, as in:

> B: He had dis **uh Mistuh W– whatever k– I can"t think of <u>his first name</u>**, Watts on, the one thet wrote // that piece,
> A: **Dan Watts**. (Schegloff et al., 1977, p. 364)

Finally, in OIOR the repair is initiated and resolved by the recipient, as in:

> B: Where didju play ba:sk//etbaw.
> A:(The) gy:m.
> B: In the gy:m?
> A: Yea:h. Like grou(h)p therapy. Yuh know=
> B: Oh:::.
> A: Half the group thet we had la:s' term wz there en we jus'<u>playing</u> arou:nd.
> B: Uh– **fooling** around.
> A: Eh– yeah ... (Schegloff et al., 1977, p. 365)

It is important to note that repair does not always involve correcting an error. It can also include instances like word searches, where speakers struggle to find the right word without necessarily replacing an incorrect term. In the following excerpt taken from the corpus investigated in this study, a word search occurs where the lecturer initiates a repair operation but does not immediately succeed:

> <Lecturer> And the engine: has **<u>maneuvers</u>** at the gate **by rotating (.) erm (.) by rotating this,**
> **erm this:**
> **I don't I don't get this, I' I' it's a (.) like a,**
> **You know?**

The repair is successfully completed around 20 seconds later through a student's intervention:

> **<Student> A screw.**
> <Lecturer> Yeah, yeah, yeah, a screw, a screw, yeah.

Similarly, Schegloff et al. (1977) also found that both self- and other-initiated repairs may "fail", as in the following OI:

> Roger: It's kinduva– // kinduv weird
> Dan: Heh
> Roger: Whadda you think
> **Ken: Hm?**
> **Roger: <u>Ferget it</u>.** (Schegloff et al., 1977, p. 366)

Furthermore, Schegloff et al. (1977) found that in everyday conversation, L1 speakers generally prefer self-repair over other-repair. This preference is attributed to a structural bias inherent in conversational interactions, largely shaped by the turn-taking mechanism. These interactions follow implicit rules governing turn-taking and topic shifts, creating a framework that favors specific conversational patterns and behaviors (Seedhouse, 2004).

In this context, research has shown that conversational patterns are "strongly constrained by the institutional setting, participant roles and the overall and local goals that need to be achieved for the task at hand" (Williams et al., 1997; as cited in Björkman, 2014, p. 129). Notably, studies in ELF settings have demonstrated that ELF speakers employ various types of repairs as part of their broader repertoire of communicative strategies (CSs) used to achieve communication goals. Unlike early SLA research, which primarily viewed CSs as evidence of L2 speakers' inadequacies, ELF research has shown that ELF speakers utilize a wider range of strategies to both pre-empt and resolve communicative challenges (Björkman, 2014; Kaur, 2011; Mauranen, 2007). Furthermore, "self-repair constitutes a powerful self-regulating mechanism that allows the speaker to not only make corrections when linguistic and factual errors occur but to also make talk more specific, explicit and clear" (Kaur, 2011, p. 2712). Self-repair is considered an "explicitness strategy" (Mauranen, 2007) that helps speakers enhance the clarity of their statements and improve the comprehensibility of speech, thereby supporting effective communication.

In this respect, Mauranen (2007, 2010a) emphasizes that making discourse more explicit is an important strategy in ELF interactions. She particularly identifies three main strategies ELF speakers use to improve clarity and comprehensibility, i.e., rephrasing, topic negotiation, and discourse reflexivity.

Despite being long deprecated by linguists, rephrasing has been found to play a significant role in academic writing and in making lectures more comprehensible (Suviniitty, 2012). Interestingly, it does not seem to be influenced by greater language proficiency or dependent on the speaker's L1, as similar types of repetitions and rephrasings are observed in both L2 and L1 speech – though in proportionally greater quantities in L2 (Mauranen, 2012). Rephrasing or restructuring a previous statement serves important functions such as (1) organizing discourse and maintaining coherence, especially when (2) dialogue shifts into other-repetition (allo-repetition). The examples below are taken from Mauranen (2010b). (1)

<S6> yeah but these salaries are not public so *you cannot compare*
<S1> yeah </S1>
you you cannot compare in the same place you have to to make a meeting and to compare each er salary sheet to to see wheth- if the people don"t know what can they do it's not public.

(2)
<S9> [they] said there are no occupational disease
<S1> mhm </S1>
this occupational [health]
<S2> [no] </S2>
manager </S9>
<S2> but of course they have occupational related [disorders and so on]
<S10> [dis- mhm] yeah </S10>
I mean a occupational diseases in Finland er you have to prove it very [carefully]
<S5> [mhm] </S5>
so it's it's not so easy to get (xx) as one
<S10> yeah </S10>
[but that's very good that they don"t have] </S2>
<BS7> [well I asked the managers]
<S10> mhm </S10>

what they have is mainly **injuries** </BS7>
<S2> **injuries** [yeah] </S2>
<S10> [which] is accident (xx)
<S2> yeah </S2>
accident but from the manager
<S2> yeah </S2>
he said no occupational health disease
<S2> no </S2>
but for occupational health accidents
<S2> yeah </S2>
they have [some] </S10>

In addition, in multi-party conversations, rephrasing also fosters a cooperative atmosphere by signaling that participants are focused on the same topic and making sense of it together.

<S2> I think that it's yeah sometimes it's really like this you said
<S1> mhm-hm </S1>
but not always </S2>
<S1> no no of course not </S1>
<S2> [but in general] </S2>
<S1> [this is] </S1>
<S2> [in general I agree] </S2>
<S1> [yeah on average] </S1>
<S2> yes I agree for example I worked in one company and i was in...

Topic negotiation involves signaling a topic change by first introducing it with a noun phrase and then using a co-referential pronoun, a technique known as left-topic dislocation, which helps align both speaker and listener before continuing. As in the following example from Mauranen (2010b, p.13):

one of my friends she tried to enter to the university

As Mauranen (2010b, p. 14) points out topic negotiation "is likely to facilitate comprehension by helping ensure that speaker and hearer have the same topic in mind, and it serves an interactional function in indicating willingness to cooperate." Lastly, through discourse reflexivity or metadiscourse – referred to as "discourse about the ongoing discourse" (Mauranen, 2010b, p. 14) – speakers explicitly guide listeners, whether in organizing the text or managing the interaction (Mauranen, 2010a). As Mauranen (2010b, p. 14) notes, metadiscourse is a crucial explicitness strategy and is "ubiquitous in all genres in ELFA." Unlike in written language, in dialogues, metadiscourse takes on a much wider range of functions, particularly in discourse management. It primarily serves four roles: (1) introducing topics, (2) clarifying the speaker's stance toward the content, (3) indicating a shift in the discussion, thereby guiding the flow of the interaction. Below are examples from Mauranen (2010b) that illustrate these roles.

(1) <S1> what about we've we've talked about groups before and for example in Spanish... </S1>

(2) <S12> cause there's a difference betwe- in [the rules if it's]
<SU> [yeah @@] </SU>
if it's like, you know I think I'm not too familiar with the differences but I think it refers to that I don't know what they're allowed to do but some things are not allowed...

(3) <S1> er okay before we go to the next topic, I I think that. in a way the question <NAME> made what made you study or be- become interested on this issue it is a relevant question cause this your topic leads us a bit further to more general
<S2> yeah </S2>
discussion about human rights or in general whether we can... </S1>

These practices collectively enhance communicative clarity and illustrate how speakers manage the complexities inherent in ELF interactions.

Building on Björkman's (2014) review, Table 7 presents "CSs following traditional conceptualizations," derived from Dörnyei (1995, p. 58, as cited in Björkman, 2014, p. 125).

Avoidance or Reduction Strategies	1	**Message abandonment**	leaving a message unfinished because of language difficulties.
	2	**Topic avoidance**	avoiding topic areas or concepts which pose language difficulties.
Achievement or Compensatory Strategies	3	**Circumlocution**	describing or exemplifying the target object or action (e.g., *the thing you open bottles with* for *corkscrew*).
	4	**Approximation**	using an alternative term which expresses the meaning of the target lexical item as closely as possible (e.g., *ship* for *sailboat*)
	5	**Use of all-purpose words**	extending a general, empty lexical item to contexts where specific words are lacking (e.g., the overuse of *thing*, *stuff*, *make*, *do*, as well as using words like *thingie*, *what-do-you-call-it*).
	6	**Word-coinage**	creating a non-existent L2 word based on a supposed rule (e.g., *vegetarianist* for *vegetarian*).

	7	**Use of non-linguistic means**	mime, gesture, facial expression, or sound imitation.
	8	**Literal translation**	translating literally a lexical item, an idiom, a compound word or structure from L1 to L2.
	9	**Foreignizing**	using a L1 word by adjusting it to L2 phonologically (e.g., with a L1 pronunciation) and/or morphologically (e.g., adding to it a L2 suffix).
	10	**Code-switching**	using a L1 word with L1 pronunciation or a L3 word with L3 pronunciation in L2.
	11	**Appeal for help**	turning to the conversation partner for help either directly (e.g., *What do you call …?*) or indirectly (e.g., rising intonation, pause, eye contact, puzzled expression).
Stalling or Time-gaining Strategies	12	**Use of fillers/ hesitation devices**	using filling words or gambits to fill pauses and to gain time to think (e.g., *well*, *now let me see*, *as a matter of fact*).

Table 7. "CSs following traditional conceptualizations," derived from Dörnyei (1995, p. 58, as cited in Björkman, 2014, p. 125).

Additionally, building on Björkman's (2014) study, Table 8 provides a review of the CSs identified within the field of ELF pragmatics as for CSs used in academic ELF settings.

Mauranen (**2006b**) (**see also Mauranen, 2007**)	**On preventing misunderstandings:**	
	Confirmation checks	
	Interactive repair	
	Self-repair	
	Clarifications	
	Repetitions	
	Co-construction	
Kirkpatrick (**2007**)	**Speaker strategies:**	**Listener strategies:**
	Spell out the word	*Lexical anticipation*
	Repeat the phrase	*Lexical suggestion*
	Be explicit	*Lexical correction*
	Paraphrase	Don"t give up
	Avoid local/idiomatic referents	Request repetition
		Request clarification
		Let it pass
		Listen to the message
		Participant paraphrase
		Participant prompt

Cogo (2009)	Repetition and code-switching
Bjørge (2010)	Backchanneling
Kaur (2010, 2011)	**On achieving mutual understanding:**
	Repetition
	Paraphrase
	Requests for confirmation of understanding
	Requests for clarification
Björkman (2011, 2013)	*Comment on terms and concepts*
	Comment on details of task
	Comment on discourse structure and content
	Comment on intent
	Comment on common ground
	Comment on signalling importance
	Backchanneling repair (self and other) (the first five strategies were adopted from Penz, 2008)

Table 8. Reported strategies and other pragmatic phenomena in ELF studies in instructional settings (from Björkman, 2014, p. 126-127).

As observed in the table above, different ELF studies have proposed various taxonomies, often because these studies focus on "a selection of strategies" (Björkman, 2014, p. 127). Moreover, as Björkman notes, these proposed taxonomies tend to overlap in terms of the functions that CSs may perform in discourse. For example, repetition is widely recognized as a versatile strategy that supports communication by enhancing production, comprehension, and interaction. It is crucial in both preventing and resolving misunderstandings (Kaur, 2014). According to Norrick (1987; as cited in Kaur, 2014), when a speaker repeats themselves, they can hold onto their turn in conversation, gain time for formulating their thoughts, manage interruptions, improve the coherence of their message, and aid in understanding, among other purposes. Conversely, when another speaker repeats what was said, it shows active listening, signals acknowledgment and interest, expresses agreement or disagreement, initiates repair, and serves other communicative functions. Additionally, in interactions involving non-native speakers, repetition takes on added significance, such as enhancing clarity or accuracy of statements (Murata, 1995) and verifying correctness (Sawir, 2004) after a native speaker has prompted a repair.

To address the multifunctional nature of CSs, and in an effort to summarize the strategies identified in earlier studies, Björkman (2014) introduced a taxonomy of strategies identified in naturally occurring ELF interactions, which are primarily divided into two main categories: self-initiated and other-initiated communication strategies. This taxonomy is summarized in the two tables below (Table 9 and Table 10), with examples taken from Björkman's (2014) study. The references are retained as cited in the original study.

Self-initiated CSs (in Björkman, 2014)			
1. Explicitness strategies	(a) Repetition	i.e., restating or echoing the previous speaker's utterance	<S1> he said er **higher surface area per volume** er er er lets you increase the temperature it he said, er er er **higher surface area per volume** will er mean that you can increase the temperature </S1>
	(b) Simplification	i.e., similar to paraphrasing, but for lexical items	<S1> it will be **double**, *I mean* **two times** </S1>
	(c) Signaling importance	i.e., when the speaker emphasize the significance of an item by explicitly highlighting it within the discourse	<S1> you of course you saw this information from internet or some from some handbook 4 you should write down the reference **it's very important** for us </S1>
	(d) Paraphrasing	i.e., providing the same content by modifying the previous utterance or ongoing utterance	<S1> but I think we should do because before we hand in our paper **we should correct** some 7 mistake **or improve it** I think and **we should er organize the contents** yeah I think so </S1>
2. Comprehension check		i.e., generally questions posed by the speaker to check whether the listener is able to follow along with what is being said	<S4> [...] so these forces are 5 always somehow a little too late **you know what I mean** [...] <S4>

3. Word replacement	i.e., a type of repair which can occur even when there is nothing to “correct”. Instances of word replacement may pertain either to language usage or to content-related issues.	<S4> [...] but I would guess so because he always **argumented** like **argued** like if we increase the air speed the stiffness of the wing decreases so if the stiffness decreases I would say the frequency **increases** er **decreases** as well </S4>

Table 9. Self-initiated CSs observed in Björkman (2014)

Other-initiated CSs (in Björkman, 2014)			
1. Confirmation checks	(a) Paraphrasing	i.e., speakers rephrased the previous statement to ensure they had correctly understood it. This approach has also been termed the “interpretive summary strategy” in a previous SLA study (Jamshidnejad, 2011; as cited in Björkman, 2014, p. 132).	**<S1> So how about your feeling about the last trip trip to the waste water plant** <NAME OF PLACE> **</S1>** **<S4>** Sorry **</S4>** <S1> *I mean* **we have already visited waster water plant of** <NAME OF PLACE> **and how do you think it** *I mean* **do you have any feeling?** </S1>
	(b) Repetition	i.e., repetition of other’s statements in order to fostering cooperation in ELF interactions.	**<S1>** good ok, and then, at the end, at the end then he said something about transparency what is that, in the binder (xx) special he said transparency you said yes what is that **</S1>** **<S2>** ah yes **slides slides** because he used the Swedish word **slides slides </S2>**
	(c) Overt question	i.e., speakers ask questions about the previous statement when it’s unclear, using confirmation checks to do so.	**<S2>** [...] so erm I I put very big mass er which is half the mass of the total wing but I think in a real real case you can’t do that because [...] **</S2>** **<S1>** with **with putting a mass do you mean** that you made the wing more stiff </S1> **<S2>** no you just put a mass without any stiffness **</S2>** **<S1> without any stiffness? </S1>** <S2> yeah you just you know just just like a something you you glue on it or you you stick on it or something </S2>

2. Clarification requests	i.e., Requesting explanations or additional information on something the speakers haven't fully understood, thereby necessitating further interaction. Clarification requests usually consists of questions or question repeats (Dörnyei and Scott, 1997:16; as cited in Björkman, 2014, p. 129).	<S1> I can ask them if they have have a lease a lease program </S1> **<S2>** lease? **</S2>** <S3> lease like you </S3> **<S1> rent </S1>** <S3> rent </S3> <S2> rent </S2>
3. Co-creation of the message/anticipation (in Kirkpatrick, 2007)	"where speakers fill in the blanks in each other's utterances in an effort to produce a complete utterance, which in turn means a complete message. Kirkpatrick terms this type of usage "lexical anticipation" (Kirkpatrick, 2007)" (Björkman, 2014, p. 133).	**<S1>** yeah yeah but I am gonna ask him what what does it what does it **</S1>** **<S2> consume </S2>** <S1> yeah consume and </S1>
4. Word replacement		<S1> yeah check transparencies </S1> <S2> **slides** </S2> <S1> slides slides </S1>

Table *10*. Other-initiated CSs observed in Björkman (2014)

Findings from Björkman's (2014) study reveal that instances of other-initiated CSs occurred more frequently than self-initiated CSs, particularly in the form of confirmation checks and clarification requests. Among self-initiated CSs, explicitness strategies – especially paraphrasing and repetition – were more common, followed by comprehension checks. The author suggests that the higher frequency of other-initiated CSs may be attributed to the highly interactive nature of the dialogic exchanges examined, specifically within students' group-work discussions.

Two key reflections from Björkman's (2014) study are noteworthy. First, the author focuses on the function of identified CSs within interaction, emphasizing the "cooperative nature" of these English as a Lingua Franca (ELF) interactions, where speakers employ proactive strategies to enhance explicitness and prevent potential misunderstandings (p. 135). Importantly, Björkman (2014) examined polyadic interactions during group-work sessions where multilingual and multicultural students with varying levels of English language competence collaborated on tasks without a lecturer's presence. Despite the asymmetries in English proficiency and subject knowledge among participants, the higher occurrence of other-initiated CSs could be due to the absence of a socially recognized authority figure (i.e., the lecturer) to whom students typically direct questions. Additionally, the interactions among students in this study do not seem to follow the traditional Initiation-Response-Follow-up (IRF) sequence (Sinclair & Coulthard, 1975) typical of educational settings.

Second, Björkman (2014, p. 128) highlights that identifying CSs is challenging, largely because it depends on how one defines a CS and whether it is viewed as an intraindividual or interindividual phenomenon. While there is little consensus on the best way to identify CSs, it is widely agreed that the clearest evidence of CSs emerges in discourse that is explicitly marked by the speaker (Björkman, 2014, p. 129). Moreover, identifying CSs requires careful analysis of the surrounding discourse, considering both preceding and subsequent conversational turns. In this context, a close examination of the provided excerpt reveals that what the author identified as other-initiated CSs may not fully align with this classification. Moreover, the author does not explicitly consider who repairs the trouble source – or, in other words, who accomplishes the CS – but focuses solely on who initiates it. Before providing some examples, it is useful to recall the definitions provided by Schegloff et al. (1977, p. 364): self-initiated repairs are those "initiated by the speaker of the trouble source", while other-initiated repairs are initiated "by any party other than the speaker of the trouble source". Furthermore, as previously mentioned, self-initiated repairs can be resolved either by the speaker themselves or by other speakers, and the same applies to other-initiated repairs.

For example, the following instance from Björkman (2014, p. 133) illustrates an other-initiated word-replacement CS. In this example, S1 identifies an issue in his/her own speech (self-initiated, as the speaker who produced the trouble source also flagged it). The issue is then resolved by another speaker, S2, who replaces the word "transparency" with "slide".

<S1> good ok, and then, at the end, at the end then he said something about transparency what is that, in the binder (xx) special he said transparency you said yes what is that **</S1>** **<S2>** ah yes **slides slides** because he used the Swedish word **slides slides </S2>**	SIOR	where is the trouble source?	S1
		who raises concerns (i.e., questions, non-verbal cues) about that trouble?	S1
		who repairs?	S2

The same applies for the following example (Björkman, 2014, p. 134), which is an instance of what Björkman refers to as co-creation of the message.

<S1> yeah yeah but I am gonna ask him what what does it what does it **</S1>** **<S2> consume </S2>** <S1> yeah consume and </S1>	SIOR	where is the trouble source?	S1
		who raises concerns (i.e., questions, non-verbal cues) about that trouble?	S1
		who repairs?	S2

These comments are made solely to demonstrate the challenges in identifying CSs in discourse, as well as to illustrate how CSs have been analyzed in this thesis, as discussed in Chapters 3 and 4.

As noted in Björkman (2014), one of the most frequent self-initiated CSs is repetition. Kaur (2012) analyzed instances of repetition found in naturally occurring interactions among graduate students at a university in Malaysia, who were ELF speakers with varying levels of English proficiency. She identified four main types of self-repetition – i.e., same-speaker repetition within an ongoing turn – that were not used to repair intelligibility issues (Watterson, 2008, p. 393; as cited in Kaur, 2012, p. 604), but rather "to enhance the clarity of expression for improved understanding" (p. 610): parallel phrasing, key word(s) repetition, combined repetition, and repaired repetition.

Parallel phrasing, first identified by Norrick (1987; as cited in Kaur, 2012, p. 599), is a repetition that "is not verbatim but displays slight variation and results in a numbered list of items." (Kaur, 2012, p. 600). The example provided below is taken from Kaur (2012, p. 600-601):

1 V: wherever their target groups are: located it is no longer
2 S: °uhhuh°
3 V: er **distance is no longer a barrier**,
4 S: yeah
5 V: **distance is no longer an impedi**[**ment**,
6 S: [yes
7 V: it **is no longer an obstruction**,
8 S: °uhhuh°=
9 V: =you understand? because no matter where you are you"ll get the
10 information.

In the excerpt above, the comprehension check "you understand?" in line 9 indicates that this instance of self-repetition is indeed focused on ensuring comprehension.

A second type of self-repetition identified by Kaur is key-word(s) repetition, i.e., "the recycling of a lexical item(s) oriented to by the speaker as crucial for purposes of understanding the message or idea being put across." (Kaur, 2012, p. 602). An example of key-word repetition is provided below:

1 L: I think- I think- . . .(0.9) other than this starting point there should be
2 also be one- the other starting point which is . . .(0.8) the level of- . . .(1.3)
3 the- I don"t know whether you like- you know this word er density
4 V: yeah population density=
5 L: =okay we don"t talk about population density but we talk-talk about
6 . . .(1.0) the density of . . .(0.6) **SME network**, **SME network**.
7 V: °yeah°

As demonstrated in the examples above, repeating key words enables the speaker to emphasize and highlight elements that are central to conveying their message. In situations where speakers do not share the same linguistic variety, such as in ELF contexts, reusing key words in the way described can help focus attention on the most important points needed to achieve effective communication.

Thirdly, combined repetition refers to "[t]he practice of combining exact repetition with repetition with slight variation or a reformulation" (Kaur, 2012, p. 604), for example, by "substituting a key word in the preceding repeated segment with a synonym" (Kaur, 2012, p. 604). For example, in the following excerpt from Kaur (2012, p. 604), the speaker first repeats the word *borderless* (line 2), then reinforces the concept by providing a definition (i.e., doesn"t have borders), which is immediately followed by a synonym, *boundary*.

1 V: okay okay let-let me comment. you know this world is: . . .(0.5) is a
2 **borderless** world, [**borderless** it doesn"t have borders, no boundary
3 S: [mm mm
4 uhhuh no [border
5 V: [you understand?=
6 S: =yeah.

Finally, repaired repetition refers to instances where the speaker replaces words or segments of speech to ensure clarity, particularly in response to speech disturbances, such as overlapping speech. The example provided below is from Kaur (2012, p. 607).

1 D: why you: not come °tomorrow° ah **yester**[**day**
2 S: [yesterday
3 D: why you not come **yesterday?**=
4 S: =yesterday I was trying to do the: the: this thing.

Interestingly, as Kaur points out, repaired repetition differs from the "reformulation repetition" identified by Murata (1995, p. 353; as cited by Kaur, 2012, p. 606), where non-native speakers use self-repetition in an attempt "to find appropriate words or phrases in the process of interaction" to enhance the accuracy and clarity of their utterances. This is illustrated in the following example, where the substitution of the preposition *about* with *in* suggests that the self-repetition is intended to repair the original utterance:

S: So have you taught erm have you taught **about**- have you taught **in** Japan- Japanese universities? (Murata 1995: 352; as cited in Kaur, 2012, p. 606).

As evidenced by the research discussed above, effective communication between ELF speakers – defined as "speakers' competence in making meaning and achieving communicative goals" (Kaur, 2011, p. 2713) – increasingly relies on their awareness of linguistic and cultural diversity in ELF contexts. Given that comprehension is often difficult to achieve purely through linguistic means (Mauranen, 2007), ELF speakers employ various strategies to mitigate potential communication challenges. Specifically, they use repair practices and CSs as methods to enhance the clarity of their utterances and improve the chances of being understood. However, research has also shown that "lecturers in ELF settings make less frequent use of pragmatic strategies than students, who deploy these strategies frequently in group-work sessions" (Björkman, 2011, p. 950). As Ollinger (2012, p. 72) points out, this is unsurprising "considering the strikingly different interactive dynamics of these two speech events". Consequently, "[i]t is reasonable to assume that, in the absence of appropriate pragmatic strategies used often in lectures, there is an increased risk for covert disturbance" (Ollinger, 2012, p. 950). Nonetheless, Ollinger (2012, p. 78) emphasizes that "[w]hile the majority of studies researched multi-party and thus dialogic talk, elements of monologue talk occurred occasionally, especially in the form of lecturer talk".

In EMI settings, research into the pragmatic strategies used by EMI lecturers has focussed for the most part on discourse reflexivity or metadiscourse. This is because metadiscourse markers hold particular significance in lecture comprehension, especially in EMI contexts, which are often characterized by monologic and highly informative lectures (Broggini & Murphy, 2017; Molino, 2018). Studies suggest that while the lecture genre, cultural factors, and disciplinary culture may influence lecturers' use of interactive metadiscourse in EMI, certain metadiscourse features are more closely associated with this context. Notably, personal metatext forms like reformulations, metalinguistic comments (e.g. *let me clarify what I mean by*), and self-mentions (e.g. *personally, I think…*) are more prevalent (Molino, 2018; Broggini & Murphy, 2017). Conversely, impersonal metadiscourse expressions, such as connectives (e.g. *moreover*, *in conclusion*), appear less frequently in the spoken language of EMI lecturers, who also demonstrate a limited variety of connectors (Broggini & Murphy, 2017). Similarly, Zhang and Lo's (2012) analysis on

EMI lecturers' use of interactive metadiscourse (i.e., transition markers, such as *in addition*, *but*, *thus*, *and*, *because*, *so*; frame markers such as *first*, *then*, *next*, *now*, *well*; reminders, e.g. *as noted above*; code glosses, e.g. *this is called*, *it can be defined as*) revealed that, despite differences in the nature of the courses, the use of interactive metadiscourse shared common features across the four courses analyzed. Transition markers – which are crucial for maintaining the cohesion and coherence of the ongoing discourse – were the most frequently used category, while reminders were the least commonly used. In contrast, frame markers and code glosses were only marginally employed. The study also found that the linguistic realizations of transition and frame markers were limited to a narrow range. However, the authors emphasize that the use of transition markers is closely tied to the genre of subject teaching, particularly in science education, where "providing a scientific accounting of theories and phenomena constitutes an essential part of classroom instruction" (p. 70).

In this regard, Mauranen (2012), in a comparison of word frequencies in the ELFA (English as a Lingua Franca in Academic Settings) and MICASE corpora, revealed that non-native speaker lecturers, despite using a smaller vocabulary range, rarely impede comprehension with their more limited lexical and syntactic repertoire. Similarly, a corpus-based contrastive study by Deroey and Johnson (2021) on lecturers' use of importance markers found little difference between how L1 and EMI lecturers employ these markers. However, intra-corpus differences were identified, suggesting that lecturerss' teaching experience and educational culture, rather than language proficiency alone, influence their use of metadiscourse markers.

2.4 Gestures and Non-verbal Communication in EMI

According to Canale (1983; as cited in Lazaraton, 2004, p. 80) strategic (communicative) competence "is composed of mastery of verbal and non-verbal communication strategies . . . a) to compensate for breakdowns in communication due to . . . insufficient competence . . . and b) to enhance the effectiveness of communication (pp. 10–11). Nonetheless, as Allen (2000; as cited in Lazaraton, 2004, p. 88) pointed out, SLA research on comprehensible input has almost always been concerned with verbal input, with no attention to the nonverbal aspects of L2 teacher talk.

2.4.1 The Role of Gestures in Supporting Verbal Communication

Gestures, primarily consisting of hand movements that speakers perform unconsciously while talking (cf. Kendon, 1986, 2004; McNeill, 1992), are systematically and closely linked to language and speech. In his seminal paper "So You Think Gestures Are Non-Verbal?", David McNeill (1985) challenged the prevailing view that gestures are insignificant additions to communication, irrelevant to our understanding of language and linguistic processing, by arguing that gestures are indeed linguistic, as evidenced by the simultaneous development of speech and gesture in childhood, their concurrent deterioration in cases of language impairment, and their joint processing during "crossmodal information integration" (Graziano & Gullberg, 2018, p. 1). Substantial research now supports this view, (for more detailed references and further reading on this topic, see Graziano & Gullberg, 2018, p. 1) demonstrating that gestures are both temporally and semantically aligned with speech, mirroring both lexical and discursive structures within language (Gullberg, 2006a; Graziano & Gullberg, 2018).

2.4.2 Types of Gestures and Their Functions

"Gestures are semantically coexpressive with speech" (Gullberg, 2006a, p. 158) such that they often express meanings already conveyed in speech, either iconically, metaphorically, or indexically. Several taxonomies of gesture types and functions have been provided in the literature (see eg. McNeil, 1992; Kendon, 2004; Streeck, 2009).

Following McNeill (1992), gestures are of two main types: representational, and emblems. Representational gestures, which are co-speech gesture, are classified into four main types: iconic, metaphoric, deictic, and beats. Iconic gestures visually represent referential meaning – i.e., "concrete physical features of the world" (Straube et al., 2010, p. 521) – directly illustrating what is being described – e.g., lowering the right hand to depict an arch while saying, "There is a bridge over the river" (Straube et al., 2010, p. 521). Metaphoric gestures, on the other hand, visually represent abstract ideas or categories – e.g., lowering the right hand to depict an arch while saying, "The politician builds a bridge to the next topic" (Straube et al., 2010, p. 521). Deictic gestures involve pointing, either to refer to specific objects in the physical space or to draw attention to abstract referents that are "spatially anchored in gesture space" (Gullberg, 2006a, p. 160). Finally, beats gestures are rhythmic hand movements that align with the flow of speech but do not carry specific representational meaning (e.g., quick hand flicks that match the emphasis or rhythm of spoken words). Emblems, a term initially introduced by researcher David Efron (1972), refer to gestures with specific meanings that are widely recognized within a particular ethnic, cultural, or subcultural group (e.g., the "thumbs up" gesture). These gestures are used intentionally and consciously, much like spoken words, and are distinctive in that they can be employed either alongside or instead of words.

In contrast to McNeill (1992), who argues that gestures and words originate from the same mental process, Kendon (1980) views gestures and speech as distinct but interconnected elements. While both play crucial roles in the "process of utterance" – how we communicate – Kendon notes that gestures can evolve over time to take on more structured meanings, much like words do. According to Kendon:

> Whereas, in speech, linguistically significant units can be organized in temporal sequence only, in the kinesic medium forms may be constructed which contrast in their spatial organization, as well. Furthermore, the instrument of sign language expression, the body, has a number of spatially separated parts which can be used as articulators, simultaneously. (Kendon 1988, p. 6; as cited in Streeck, 2009, p. 22)

Kendon (2004) first identifies some formal properties of gestures, distinguishing them based on the articulators involved (e.g., hands, arms, eyebrows), the space used while gesturing (e.g., extreme periphery, periphery, center, center-center), and the movements performed (e.g., back/forth, up/down, left/right). Additionally, Kendon (2004) describes the stages involved in gesturing, often referred to as gesture units. The act of gesticulation typically involves the hand moving from a rest position (preparatory phase) to a hold position (pre-stroke), which anticipates the execution of the gesture (stroke). This may sometimes be followed by another hold position (post-stroke) before the hand returns or retracts. Gesture units can consist of one or multiple gesture "phrases." For example, this could include a series of repetitive strokes, potentially with a change in hand configuration, followed by a return to the original stroke pattern, continuing until the hand finally returns to its rest position. Such extended gesture units are often observed when participants take longer turns during conversation.

As Streeck (2009, p. 23) points out, speech-gesture coordination is so complex that, in order to understand how gestures contribute to meaning,

> it is never sufficient during research to identify tokens of a certain gesture, count their frequency, and correlate these with some other variables: what a token of a gesture *does* and how it is taken is contingent on its precise temporal relationship to the utterance or utterance sequence within which it occurs, the unit that it precedes, the ones that it follows, and on how vocal and kinesic units are delivered moment by moment in relation to one another.

In this respect, since the 1980s, Kendon has taken a functional approach to studying gestures, focusing on how hand movements contribute meaning to utterances, which he describes as "semantic interaction" (Kendon, 2004, p. 158). He identifies three primary functions of gestures: referential, pragmatic, and interactive (or interpersonal). Referential gestures "contribute to the propositional content of utterances" (Streeck, 2009, p. 24) – e.g., drawing a square shape in the air while discussing a box. Pragmatic gestures are non-referential; they do not directly convey specific content related to the topic of speech or refer to concrete objects, or actions. Pragmatic gestures include modal, parse, and performative gestures. Modal gestures express attitudes or modalities, such as certainty or doubt – e.g., repeatedly tapping fingers to emphasize a point rather than conveying any specific meaning. Parse gestures help structure speech and make it easier to follow – e.g., pausing and holding hands outward when introducing a new idea. Performative gestures accompany speech acts like commands, requests, or promises – e.g., pointing a finger when giving an order. Finally, interactive or interpersonal gestures help manage interaction, such as signaling turn-taking, eliciting feedback, or marking agreement, as well as pointing to draw attention.

Nonetheless, single gestures can perform different functions. As Kendon also points out this is "a typology of functions, not of gestures. Any given gestural form may, according to context, function now in one way, now in another" (Kendon, 2004, p. 225).

As Schegloff (2007) pointed out, human societies rely heavily on "an organization of interaction informed by the use of language" (as cited in Streeck, 2009, p. 25). Gestures are integral to this interactional organization. More specifically, as Streeck (2009) argues, gestures are crucial in addressing the *turn-taking problem* – the need for participants to manage speaking turns sequentially, given the linear nature of speech and the human limitation of being unable to speak and comprehend simultaneously. Gestures also play a vital role in resolving the *action-formation problem*, which involves determining how language, bodily movements, environmental cues, and interactional positioning coalesce into actions that are recognized by others as specific communicative acts. Additionally, gestures help address the *sequence-organization problem*, which requires structuring turns at talk to be responsive to preceding actions and to signal the kind of response being sought. Lastly, gestures assist in managing the *trouble problem*, handling contingencies that arise from potential difficulties in hearing or understanding during interaction. For instance, gestures are often involved in word-searches (Goodwin, 1986; Streeck, 1993; as cited in Streeck, 2009, p. 26) and self-repair (Seyfeddinipur, 2006). Nonetheless,

> there has actually been relatively little research on gesture within a conversation analytic framework (but see Lerner 2002; Schegloff 1984) – most certainly due to the fact that many of its practitioners regard *talk* in interaction – and not interaction per se or action in interaction – as the crucible of human sociality; gesture, accordingly, is relegated to its customary role as a hand-maiden to speech. (Streeck, 2009, p. 27)

2.4.3 Interaction Between Gestures and Co-occurring Speech

Following Streeck's (2004) review of gesture studies, Charles and Marjorie Goodwin are among the few conversation analysts who emphasize the significance of gestures alongside spoken language. Their initial studies focused particularly on word-search activities, revealing that these instances are marked by distinct speech patterns such as pauses, hesitations, and disruptions, along with non-verbal cues like gaze aversion or a "thinking face" (Goodwin & Goodwin, 1986, p. 57; as cited in Streeck, 2004, p. 27). While these signals alert listeners that a word-search is taking place, the preceding conversation often hints at the type of word being sought and a possible solution. Within this structured context, listeners can interpret the gestures that the speaker uses in place of the missing word. When they find a suitable word or phrase that conveys the intended meaning and fits grammatically, they effectively resolve the word-search and complete the speaker's thought.

In two experiments analyzing speakers' conversational interactions during living space description tasks, Seyfeddinipur (2006) examined instances of speech DFs and associated gestural behaviors, providing evidence that gestures are sensitive to these DFs. The study found that speakers typically do not interrupt their speech immediately upon detecting an error, as suggested by the Main-Interruption-Rule hypothesis. Instead, they tend to interrupt only when they are ready to repair the error, aligning more with the Delayed-Interruption-For-Planning hypothesis. This finding implies that some degree of cognitive replanning occurs before speech is suspended. Consequently, DFs appear to be linked to cognitive load during speech planning: as planning becomes more complex, speakers are more likely to produce DFs, characterized by pauses typically lasting longer than 200 milliseconds, which in turn affects the timing and execution of gestures.

Seyfeddinipur (2006) also compared gestural behavior during disfluent and fluent utterances. While the overall rate of gestural activity remained consistent across both conditions, differences emerged in the timing of gesture suspension and its position within the gesture phrase. In disfluent utterances, gestures were often suspended before speech in after-word suspensions, but not within-word suspensions. The occurrence of gesture suspension suggests that speakers monitor not only their speech but also their gestures, along with the coordination between these modalities. This monitoring likely considers factors like clarity and visibility to ensure effective communication with the listener. However, the study noted that in this specific, less interactive task, pointing and iconic gestures were more common, while other gesture types, such as beat gestures and emblems, were rare. This finding indicates that further research is needed across different discourse genres to understand how various gestures are suspended.

In the study "When Speech Stops, Gesture Stops," Graziano and Gullberg (2018) tested the proposed compensatory role of referential gestures, particularly during word searches. While several theories implicitly suggest a compensatory function for gestures during speech DFs, they do not explicitly address this. For example, the Information Packaging Hypothesis (Alibali et al., 2000; Kita, 2000, as cited in Graziano & Gullberg, 2018) proposes that referential gestures aid in the conceptual planning of spoken messages, especially for spatio-motoric concepts. An expanded version, the Gesture-for-Conceptualization Hypothesis (Kita et al., 2017, as cited in Graziano & Gullberg, 2018), suggests that speakers can activate, manipulate, package, and explore spatio-motoric information for both speaking and thinking through referential gestures. While these theories imply a compensatory role for gestures, they do not explicitly state this. In contrast, the Lexical Retrieval Hypothesis (Krauss & Hadar, 1999; Krauss et al., 2000; Morsella & Krauss, 2005, as cited in Graziano & Gullberg, 2018) explicitly posits that “the main role of referential gestures is to facilitate lexical retrieval from the mental lexicon by means of cross-modal priming” (Graziano & Gullberg, 2018, p. 2).

However, previous studies (e.g., Gullberg, 1998, 2011, as cited in Graziano & Gullberg, 2018) have empirically challenged this view, showing that when speakers gesture during DFs, they often produce gestures that comment on the breakdown itself rather than represent the referential content of the sought words. These gestures often involve continuous wrist-turning to expose palms (referred to as metapragmatic or “thinking gestures” by Gullberg) or palm-up gestures directed toward the interlocutor. Kendon (2004) categorized many of these non-referential gestures as pragmatic gestures. Nonetheless, evidence on the specific functions of gestures during DFs is limited, and direct cross-linguistic comparisons are scarce.

Some earlier studies provide insights into cross-linguistic differences in gesture use. For example, Efron (1972) observed that Italian immigrants in the United States produced more referential gestures than Yiddish-speaking immigrants, who favored more pragmatic gestures. Similarly, Kendon (2004) found that Italian speakers used a wider range of pragmatic gestures compared to British and American English speakers. Gullberg (1998) also noted that native Swedish speakers produced more referential gestures than native French speakers, who favored non-referential gestures, specifically beats. Despite these observations, the relationship between gestures and DF has not been extensively examined across languages.

To address this gap, Graziano and Gullberg (2018) compared gestural behavior during fluent and disfluent stretches of narratives by competent speakers in two languages

(Dutch and Italian). Their study revealed three key findings: (1) across all groups, speakers overwhelmingly produced gestures during fluent speech and rarely during DFs. However, L2 learners were significantly more likely to gesture during DFs than other groups; (2) in all groups, gestures during DFs tended to be holds, consistent with Seyfeddinipur's (2006) findings; (3) the few gestures completed during DFs across all groups had both referential and pragmatic functions. These results strongly support the notion that speech and gestures form an integrated system.

In a recent analysis of tandem interactions between undergraduate French and English students, Kosmala (2021) compared DFs in the speech of both L1 and L2 speakers from a multimodal perspective. Kosmala found that DFs are more frequent and complex – such as disfluent sequences involving multiple DFs like a prolongation followed by a filled pause – in L2 speakers, especially during moments of lexical search. As Kosmala (2021, p. 95) notes,

> A non-native speaker may resort to a combination of more stalling mechanisms (more complex (Dis)S) in order to deal with their lexical difficulties and keep their partner's attention, while native speakers produce less complex (Dis)S because they do not experience as many production difficulties, but also because they may not be as self-conscious about their productions since it is their native language.

Kosmala (2021) also examined the visual-gestural features of DFs, their interactional context, and the individual strategies of both L1 and L2 speakers. Interestingly, there were no significant differences in the distribution of DFs based on their functions (e.g., speech management, discursive structuring, interactive/communicative purposes, or displaying uncertainty), indicating that language proficiency might have a limited effect on DF patterns. The only slight difference observed was in contexts of uncertainty, where DFs occurred slightly more in L2 than in L1. Regardless of whether speakers were using their L1 or L2, they employed DFs to achieve various interactional goals, suggesting that overall frequency alone isn"t a sufficient measure of DFs in L1 and L2; individual differences also need to be considered, as the use of DFs varies significantly across speakers.

Kosmala (2021, pp. 94-95) found that both L1 and L2 speakers use gestures in conjunction with DF markers as communication strategies, deploying them differently to resolve language difficulties or maintain the speaker-hearer relationship. Notably, L2 speakers tend to use more gestures, particularly holds (consistent with previous studies), during DF sequences compared to L1 speakers. This suggests that L2 speakers have a greater need to buy time in their discourse, often requiring a complete suspension of speech in different modalities.

Additionally, Kosmala observed that both L1 and L2 speakers predominantly produced pragmatic gestures during DFs, aligning with the findings of Graziano & Gullberg (2018). However, L2 speakers used more pragmatic and thinking gestures – also known as adaptors (Ekman & Friesen, 1969), such as rubbing one's forehead or fidgeting, which help manage emotions and stress. In contrast, L1 speakers produced more parsing gestures. This supports the idea that the pragmatic gestures used by L2 speakers during verbal DFs are not necessarily tied to speech content but rather serve as a metalinguistic comment on communication breakdowns, signaling issues in speech production or that the speaker is engaged in a word search (Graziano & Gullberg, 2018, p. 13).

Holler et al. (2013) provided further evidence for the communicative function of gestures during lexical retrieval. Their study analyzed participants' – all L1 speakers – gestures during tip of the tongue (TOT) experiences – when conceptual information is available but insufficient to access the word form – in different contexts: face-to-face (FTF) communication, separated by a screen, and speaking into a voice recorder alone. The results showed that participants in the FTF context produced significantly more representational gestures than those in the solitary condition. This suggests that, even during lexical retrieval difficulties, representational gestures primarily serve a communicative role.

As Feyereisen (2006) suggests, it would be interesting to analyze the interactions between gesture and language production systems by examining the types of gestures that co-occur with different categories of words, such as manufactured objects and action verbs. Similarly, as Gullberg (2006b, p. 112) highlights, several issues in the presumed relationship between speech and gesture require further clarification, such as whether compensation and facilitation are intended mainly for native speakers, learners, or both, and at which linguistic level these compensatory processes occur – e.g., at the word formulation stage, the conceptual level, or during interaction.

2.4.4 Gestures and Non-verbal Communication in teaching

Gestures are "interactional phenomena with rich semiotic affordances to all interlocutors involved", and therefore "constitute input" (Gullberg, 2006b, p. 115). In his review of existing literature across anthropology, linguistics, psychology, and education, Roth (2001) highlights the limited research on the role of gestures in teaching, particularly in subjects like science and mathematics, which are often characterized by abstract concepts. Roth (2001) suggests that such research could be crucial in enhancing our understanding of the role of gestures in learning and knowing science (p. 365). Among the few studies available, Roth (2001) observed that gestures, especially deictic ones, help scientists orient each other to specific entities and coordinate discussions about abstract concepts and visual displays. Additionally, iconic and metaphorical gestures, due to their spatio-motoric nature – meaning they represent concepts through coordinated physical movements that map out spatial relationships or shapes – allow scientists to depict phenomena and concepts that are challenging to describe using typological words (Lemke, 1999b; as cited in Roth, 2001, p. 372).

Building on Gullberg's (2006b) work, several studies have demonstrated that gestures significantly contribute to learners' comprehension and learning. These studies have focused on "teachers' gestures as conveyors of speech-related meaning" (Gullberg, 2006b, p. 116), revealing that teachers' gestures enhance listening comprehension in L2 learners, particularly benefiting those with lower proficiency levels.

In language learning specifically, gestures may also aid in the recall of lexical items (e.g., Lazaraton, 2004). Lazaraton's (2004) microanalytic study of a single teacher's use of gestures during unplanned vocabulary explanations in an L2 context found that gestures play a significant role in supporting and enhancing verbal explanations. By serving as visual aids, gestures can make abstract or difficult concepts more concrete and understandable, potentially improving student comprehension, especially for those who struggle with purely verbal explanations. These findings suggest that gestures can be an effective tool for ESL teachers, helping to bridge language gaps and facilitate learning.

2.4.5 Gestures and Non-verbal Communication in EMI

As illustrated above, in instructional settings, research has shown that the discourse of a discipline encompasses a full range of semiotic resources integral to conveying meaning and knowledge within that field. These resources include images, spoken and written language, mathematical symbols and formulae, tools like software, gestures, and working practices (Airey & Linder, 2009; as cited in Malmström & Pecorari, 2021, p. 218). In some interactions, gestures, gaze, and head movements may take precedence over language, which could be subordinated or even absent (Norris, 2004).

The concept of language as encompassing not just written and spoken words, but also visual elements, typographical features, gestures, facial expressions, eye contact, and posture, has increasingly gained significance in EMI contexts (Morell, 2018; 2020).

Following Dearden (2015, p. 23; as cited in Morell, 2018, p. 78), EMI lecturers "should not only have a sufficiently high level of proficiency, but also need to have similar skills as those required for EFL teachers," namely, "to know how to modify their input, assure comprehension via student-initiated interactional modifications, and create an atmosphere where students operating in an L2 are not afraid to speak."

Research has indeed shown that EMI lecturers often need to effectively use multimodal resources to meet students' needs and enhance comprehension (Fortanet-Gómez & Ruiz-Madrid, 2014; Crawford Camiciottoli & Fortanet-Gómez, 2015; Morell, 2018, 2020; Ruiz Madrid & Valeiras-Jurado, 2020). This is true not only in face-to-face EMI settings (see e.g., Querol-Julián, 2021). In this respect, research has particularly focused on EMI lecturers' interactional competence – i.e., their use of multimodal resources to promote student interaction and engagement – with earlier research substantiating the view that interaction within lecture discourse fosters a learning-friendly environment by enhancing understanding, capturing attention, and stimulating critical thinking (Deroey & Taverniers, 2011).

For example, Morell (2018) explored how an EMI lecturer's use of multimodal ensembles contributes to effective pedagogy by promoting student engagement and eliciting conceptual meaning. Her analysis shown that a coordinated use of complementary mode ensembles – i.e., gaze, gestures, speech, written material, and movement in the classroom space – enabled the lecturer to "textually organize and interpersonally involve students to elicit meaning" (p. 77). She identified recurring patterns of multimodal ensembles that co-occurred in specific pedagogical functions, showing that gestures are omnipresent in EMI lecturer discourse, except during the reviewing function, where Morell noted that the lecturer was engaged in writing on the board.

More recently, Morell et al. (2022) identified the moves EMI lecturers instantiate when carrying out pair-work activities – as instances of engagement episodes (EEs) –

exploring the ways in which lecturers implemented these pedagogical functions from a multimodal perspective. They found five basic moves – namely, contextualizing, setting up, monitoring, eliciting, and summarizing. They also found that in these EEs, lecturers made use of a combination of multimodal resources, namely spoken, written, non-verbal materials (NVMs; e.g., graphs, tables, images, realia), space, and posture. Findings show that in the five moves, EMI lecturers were constantly moving and making use of all modes except for NVMs, which were implemented only occasionally in the contextualizing, setting up, and summarizing moves to fulfill expository and illustrative functions. However, this study did not focus on other semiotic resources such as gaze and gestures, and it only examined lecturers' multimodal discourse during pair-work activities. As the authors point out, "[N]ot only could we explore more semiotic resources, but we could also examine how lecturers set up EEs in other types of activities such as group work, debates, and oral presentations" (Morell et al., 2022, p. 17).

Taking an SLA perspective, Costa (2023) recently referred to the linguistic and extra-linguistic cues used by EMI lecturers to facilitate student comprehension as "input presentation strategies" (p. 4). These strategies include discourse strategies – such as repetitions, examples, summaries, definitions, paraphrasing, reformulating, asking questions, code-switching, and humor – as well as speech strategies – e.g., slowing speech pace, emphasizing through intonation, and clearly articulating words. Beyond these, strictly multimodal strategies like the use of visual aids, gestures, and videos are also crucial. However, as argued by Costa (2023), comprehensible input goes beyond merely showing pictures during a lesson; it requires "a conscious effort to make the lesson understandable through various means" (Echevarría et al., 2008, as cited in Costa, 2023, p. 5).

Chapter 3. Research Methodology

3.1 Research Design

This study employs a single-case study research design as outlined by Yin (2009), focusing on an in-depth analysis of a single Italian L1 lecturer teaching in an international engineering master's program delivered through EMI at an Italian university. According to Yin (2009), a case study is an empirical investigation that examines a phenomenon within its real-life context, typically utilizing multiple data collection methods to explore the phenomenon comprehensively. Priya (2021) further emphasizes that case studies are not merely data collection techniques but research strategies that involve detailed examinations of cases within their natural settings.

This single-case study is explanatory in nature, aiming to explore the causal factors that explain how and why particular communicative strategies and behaviors emerge in the EMI classroom. The study specifically investigates how elements such as disfluencies, speech rate, self-repairs, pragmatic functions, and non-verbal resources (particularly hand gestures) interact and influence one another. By analyzing these variables in a real-time classroom context, the research seeks to understand the mechanisms that lead to communication breakdowns or successful explanations and how these factors collectively shape the teaching and learning experience in an EMI environment.

The study's focus on one complete lecture allows for a detailed examination of how communicative challenges manifest and how the lecturer navigates them. Unlike many EMI studies that rely on comparative analyses across multiple disciplines or sample episodes, this single-case approach offers a deeper exploration of both verbal and non-verbal communicative strategies. This method helped us avoid the risk of our analysis being biased by episodes with apparent theoretical appeal; instead, we allowed our sample to guide us in unexpected directions. The research not only aims to identify communicative challenges but also to explain the dynamic interplay between verbal and non-verbal strategies that either facilitate or hinder the lecturer's ability to convey intended meanings.

To enhance the depth and reliability of the findings, this study integrates both qualitative and quantitative methods. This dual approach involves supplementing qualitative case study methods with statistical analysis, examining the frequency of co-occurrence of the selected variables. This comprehensive framework ensures a robust analysis that captures both the nuanced, context-specific details of the lecturer's behavior and the broader patterns in EMI discourse, offering a well-rounded understanding of how verbal and non-verbal strategies interact in an engineering education context.

3.2 Data Collection Methods

The study triangulates data from three key sources to provide a comprehensive exploration of the lecturer's communicative challenges.

(1) *Survey responses and semi-structured interviews*. As part of a previous research project titled "Teaching in English at the University of Bologna" conducted between 2018-2020 (Picciuolo & Johnson, 2020), lecturers from the Department of Civil, Mechanical, and Environmental Engineering (DICAM) at the University of Bologna participated in a survey aimed at analyzing their perceived difficulties and needs when teaching their academic subjects through English. The head of the DICAM department at the time, who is the focus of this case study, also participated in this research. He provided insights through a semi-structured interview, the results of which served as a foundational starting point for this analysis.

(2) Student feedback via surveys. In the same research project, students enrolled in EMI courses at the DICAM department were asked to complete an anonymous survey administered online by their lecturers (Picciuolo & Johnson, 2020; Johnson & Picciuolo, 2023). This survey aimed to compare lecturers' perceptions with those of their students, assessing the alignment between the two groups' views on challenges and needs in EMI settings. Although the survey included responses from students attending various EMI engineering courses, the insights gained from this data will be utilized as a secondary source of information for this study.

(3) Videorecording of a lecture. A lecture video recorded by the lecturer under investigation was downloaded from his YouTube Channel (https://www.youtube.com/user/albertomontanari/), with permission granted to share it for research purposes. The lecture, part of the "Sustainable Management of Water Resources Systems" course within the master's degree programs in Civil Engineering, was delivered on May 2, 2019, for a duration of 83.77 minutes, corresponding to 1 hour and 40 minutes. The total word count for the lecture was 9,970 words, with 11,621 tokens in total. Notably, the lecturer's speech alone accounted for 9,824 words and 11,509 tokens, reflecting a significant portion of the verbal content. The interview with the lecturer, conducted in November 2019, provides contextual information about his teaching and communicative practices, which will be analyzed in conjunction with the recorded lecture.

3.3 Data Analysis

The data collected was analyzed using both qualitative and quantitative techniques. First, the lecture was transcribed using OpenAIWhisper[5], a transcription software

[5] https://whisperui.com/

operating in a Python environment. The transcription was then manually reviewed to ensure accuracy, particularly addressing any potential misinterpretations of the lecturer's speech caused by his non-standard English accent, pronunciation, or the use of highly technical jargon. The transcription was then uploaded into ELAN[6], a software that enables extensive annotation of multimodal data. 9 tiers were created based on the variables investigated, as shown in Table 11 below.

	Annotation tiers	description
1	Clause transcript lect	The lecture transcript was initially segmented manually into boundary clauses, which were delimited by silent pauses lasting more than 200ms. Silent pauses occurring at clause boundaries or following discourse markers – not signaling disfluencies (DFs) – were annotated with '()' in the same tier as 'Clause transcript lect,' but as separate segments.
2	Clause transcript lect TOKENIZED	ELAN software enables automatic alignment of the orthographic transcription with the audio using an integrated recognizer. The transcript was automatically segmented into individual tokens and then matched with corresponding segments of the audio.
3	Clause transcript student 1-5 Clause transcript studentS	Each instance of students' interventions was annotated separately according to the speaker. Since 5 students intervened during the lectures, 5 tiers were created. In instances where multiple students spoke simultaneously – such as saying 'yes' – an additional tier labeled 'clause transcript students' was added.
4	Pragmatic categories	Pragmatic categories of pedagogical functions were annotated according to a framework (illustrated below in this section) that integrates Alsop (2016) and Kunioshi et al. (2016).
5	(Dis)fluencies (DFs)	DFs were annotated both within the text using orthographic symbols (e.g., '(.)', 'erm') and separately on a specific tier (i.e., DFs).
6	Communicative Strategies (CSs) type	CSs were annotated by drawing from Schegloff's (2008) framework and integrating it with Bjorkman's (2014), as illustrated below in this section.
7	Gesture type	Gesture types were annotated based on Kendon (2004), categorizing them into representational gestures – i.e., iconic, metaphoric, deictic, beats – and emblems.

[6]ELAN (Version 6.8) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan (31.05.2024).

8	Gesture function	Gesture functions were annotated based on Kendon (2004), categorizing them into referential, pragmatic (including modal, performative, and parse), and interpersonal functions.
9	Gesture description	The gesture descriptions include details of the formal properties of the gestures, particularly the articulator and movement (e.g., finger pointing at the projector timetable). This was done to determine how many instances of different gestures the lecturer performed throughout the lecture. ELAN allows for the creation of a 'controlled vocabulary,' enabling the user to import a list of descriptors. Consequently, a drop-down field appears for each segment, allowing the user to annotate by selecting from multiple options. This feature also ensures greater consistency in the annotation process. A list of gestures identified in this analysis is provided in Appendix X.

Table *11*. List of tiers used in ELAN to annotate the videorecording of the selected lecture

From classroom observation, three additional tiers were created, although they were not systematically annotated, to keep track of instances that might support the analysis: 'misunderstandings' – this occurred only once during the lecture, but it was an interesting episode during a dialogic interaction with a student; 'lecturer mistakes' – these were annotated according to the type of deviation, such as pronunciation or vocabulary; and 'Actions lecturer' – while we focused on hand gestures, we also tracked the lecturer's physical actions to identify any links between disfluencies and hesitations that might occur due to multitasking or simultaneous physical activities. Provided below is a comprehensive account of the methodology used in the analysis.

3.3.1 Speech Rate

The recorded lecture was transcribed and analyzed for speech rate using the methodology outlined by Martin-Rubió and Diert-Boté (2023). Time intervals were first defined, as summarized in Table 12 below.

Time intervals	Description
Speech Time (**ST**)	The total duration during which the lecturer is actively speaking and delivering meaningful syllables.
Silent Pause Time (**SPT**)	The total duration of silent pauses during the lecture.

Filled Pause Time (FPT)	The total duration of filled pauses (such as "erm") during the lecture.
Total Time (TT)	The sum of all three intervals, i.e., *TT=ST+SPT+FPT*

Table *12*. Description of the time intervals.

Next, speech rate was measured using three distinct fluency rate metrics, as summarized in Table 13 below.

<table>
<tr><th colspan="2">Fluency rate measures</th><th>Description</th><th>Formula</th></tr>
<tr><td colspan="2">Mean Syllables per Run (MSR)</td><td>This metric represents the average number of syllables produced between pauses (including both filled and silent pauses) and is calculated by dividing the total number of syllables by the number of between-pause (both filled and silent pauses) units (bp-units or runs).</td><td>$\frac{Total\ number\ of\ syllables}{total\ number\ of\ runs\ (bp-units)}$</td></tr>
<tr><td colspan="2">(2) Rate of Speech Time (ROST)</td><td>This measures the pace at which syllables are articulated, determined by dividing the total syllable count by the ST.</td><td>$\frac{Total\ number\ of\ syllables}{ST}$</td></tr>
<tr><td rowspan="3">(3) Three Ratios</td><td>Speech Time Ratio (STR)</td><td>The STR ratio quantifies the proportion of time, expressed as percentage, that the lecturer spends on syllable articulation.</td><td>$\left(\frac{ST}{TT}\right) \times 100$</td></tr>
<tr><td>Silent Pause Time Ratio (SPTR),</td><td>The SPTR ratio quantifies the proportion of time, expressed as percentage, that the lecturer spends on silent pauses.</td><td>$\left(\frac{SPT}{TT}\right) \times 100$</td></tr>
<tr><td>Filled Pause Time Ratio (FPTR):</td><td>The FPTR ratio quantifies the proportion of time, expressed as percentage, that the lecturer spends on filled pauses.</td><td>$\left(\frac{FPT}{TT}\right) \times 100$</td></tr>
</table>

Table *13*. Description of fluency metrics *used to measure* speech rate.

The 'Clause transcript lect TOKENIZED' annotation was first exported to an XLS file, then words were automatically segmented into syllables while keeping the timestamps aligned using a Python script, an example of which is provided below.

```
import pandas as pd
import pyphen

# Path to your Excel file
```

```
file_path = r'C:\Users\Mariangela\Desktop\mona 3 syllables.xlsx'

# Load the Excel file
df = pd.read_excel(file_path)

# Initialize the pyphen syllable splitter for English
dic = pyphen.Pyphen(lang='en')

# Function to split words into syllables
def split_into_syllables(text):
    if pd.isna(text):
        return text  # If the cell is empty, return as is
    if isinstance(text, (int, float)):
        return str(text)  # Convert numbers to strings
    text = str(text).strip()  # Ensure it's a string and strip any surrounding
whitespace
    words = text.split()  # Split the text into words
    syllables = [dic.inserted(word) for word in words]  # Split each word into
syllables
    return ' '.join(syllables)  # Join the syllables back together

# Apply the syllable splitting function to the tokenized column
correct_column_name = 'Clause transcript_lecturer TOKENIZED'
df['Syllables'] = df[correct_column_name].apply(split_into_syllables)

# Save the modified DataFrame back to Excel
output_path = r'C:\Users\Mariangela\Desktop\mona_3_syllables_split.xlsx'
df.to_excel(output_path, index=False)

print(f"Syllable splitting complete. File saved to: {output_path}")
```

Importantly, numbers were transcribed into words to facilitate syllabification, and formulas (e.g., LB4) as well as years (e.g., 1980) were spelled out to ensure consistency in the syllable count. The data were prepared for statistical analysis and quantitative measurement of speech rates through the following steps. First, in a separate Excel file, timestamps were recorded along with the transcript of the lecturer's speech, which was split into syllables and counted. Silent pauses at clause boundaries or after discourse markers, as well as disfluencies – such as filled pauses (e.g., "erm") and unfilled pauses (with a minimum duration of 200 ms, occurring only within clauses) – were identified and analyzed. Additionally, sequences of different types of disfluencies (e.g., "combo (lengthening + unfilled pause)") were included in the analysis. Next, a Python script was run, as illustrated in the example below, to automatically calculate fluency rates.

```
import pandas as pd
```

```
# Load the spreadsheet
file_path = ‘/mnt/data/MONA 1_SPEECH RATES.xlsx’
df = pd.read_excel(file_path)

# Convert the ‘Duration - msec’ column to numeric values, ignoring errors for non-
numeric entries
df[‘Duration - msec’] = pd.to_numeric(df[‘Duration - msec’], errors=‘coerce’)

# Calculate Speech Time (ST)
# Assuming ‘number of syllables’ > 0 indicates active speech
ST = df[df[‘number of syllables’].notna()][‘Duration - msec’].sum()

# Calculate Silent Pause Time (SPT)
# Assuming ‘Silent pauses (<200 ms)’ indicates silent pauses
SPT = df[df[‘Silent pauses (<200 ms)’].notna()][‘Duration - msec’].sum()

# Calculate Filled Pause Time (FPT)
# Assuming ‘DFs (fp, unp)’ indicating filled pauses
FPT = df[df[‘DFs (fp, unp)’].notna()][‘Duration - msec’].sum()

# Calculate Total Time (TT)
TT = ST + SPT + FPT

# Calculate Mean Syllables per Run (MSR)
total_syllables = df[‘number of syllables’].sum()
total_runs = df[‘number of syllables’].count()  # Counting the number of non-null
syllable entries
MSR = total_syllables / total_runs if total_runs > 0 else 0

# Calculate Rate of Speech Time (ROST)
ROST = total_syllables / ST if ST > 0 else 0

# Calculate the Ratios
STR = (ST / TT) * 100 if TT > 0 else 0
SPTR = (SPT / TT) * 100 if TT > 0 else 0
FPTR = (FPT / TT) * 100 if TT > 0 else 0

# Create a summary dataframe to display the results
summary_df = pd.DataFrame({
    ‘Metric’: [‘Speech Time (ST)’, ‘Silent Pause Time (SPT)’, ‘Filled Pause Time (FPT)’,
‘Total Time (TT)’,
               ‘Mean Syllables per Run (MSR)’, ‘Rate of Speech Time (ROST)’,
               ‘Speech Time Ratio (STR)’, ‘Silent Pause Time Ratio (SPTR)’, ‘Filled
Pause Time Ratio (FPTR)’],
```

```
    'Value': [ST, SPT, FPT, TT, MSR, ROST, STR, SPTR, FPTR]
})

# Display the summary dataframe
summary_dfcorrected_results = {
    'Corrected Total Time (TT)': actual_total_time_ms,
    'Speech Time (ST)': speech_time,
    'Silent Pause Time (SPT)': silent_pause_time,
    'Filled Pause Time (FPT)': filled_pause_time,
    'Unfilled Pause Time': unfilled_pause_time,
    'Mean Syllables per Run (MSR)': msr,
    'Rate of Speech Time (ROST)': rost,
    'Corrected Speech Time Ratio (STR)': str_ratio_corrected,
    'Corrected Silent Pause Time Ratio (SPTR)': sptr_ratio_corrected,
    'Corrected Filled Pause Time Ratio (FPTR)': fptr_ratio_corrected
}

# Output the corrected results
corrected_results
```

3.3.2 (Dis)fluencies

(Dis)fluencies (DFs) in the lecturer's speech were annotated in ELAN using the taxonomy illustrated in Table 14, which provides examples from the corpus under investigation.

	Types	Abbrev.	Label	E.g.
DFs	filled pauses	fp	"erm"	And then we have in: erm erm next altitude,
	unfilled pauses (min. duration 200 ms)	unp	" (.) "	It's much better to have the flow (.) landing over water,
	interruptions	int	" ' "	And sometimes we see boulders tha' that are large,
	repetitions	rep	Not transcribed ortographically.	the pool is is, can be made by concrete,
	lengthening	leng	" : "	Because once that (.) the river starts: excavating downstream,
	false starts	fs	Not transcribed ortographically.	We got to () Okay, this is the figure.
	combination	combo (e.g. rep + leng + unp)	combinations of these categories involving at least two different kinds of disfluencies appearing in immediate sequence.	And then it depends also on the viscosity of: of the: (.) of the fluid,

Table *14*. Taxonomy of DFs with *e*xamples from the corpus.

Importantly, only intra-clausal occurrences of DFs were considered. Phenomena occurring at clause boundaries or following discourse markers – such as 'so,' 'and,' or 'well' – were not considered instances of DFs and were therefore annotated with the orthographic symbol '()' in the same tier as 'Clause transcript lect,' but in separate segments to facilitate retrieval when analyzing fluency rates in Excel. Conversely, DF markers were annotated both in the orthographic transcription of the file, using the symbols illustrated in Table 14 above, and on a separate tier – i.e., 'DFs' – to ease retrieval during quantitative analysis and to facilitate the analysis of the distribution of DF markers throughout the lecture. There are only two exceptions: *repetitions* were transcribed orthographically, resulting in occurrences of words that appear consecutively more than once in the transcript. However, in the specific 'DFs' tier, repetitions were annotated only when they were not used for emphasis. Similarly, *false starts* were not transcribed orthographically in the lecture transcript but were annotated in the 'DFs' tier as 'fs'.

The decision to annotate disfluencies using ELAN, rather than Praat[7] or other software specifically designed for phonological analysis, was intentional. The focus was not on disfluencies in isolation, but rather on examining how disfluencies co-occur with the lecturer's gestures, actions (e.g. spatial movement and writing on the blackboard), speech (i.e., co-occurring words), and *phases* of the lecture (i.e., pragmatic or pedagogical functions). This approach aimed to explore whether these verbal and non-verbal cues provided insights into the challenges faced by the lecturer. ELAN was chosen because it not only supports the integration of multiple modalities – including gestures and actions – but also offers reliable tools for analyzing speech, such as spectrograms, which enabled accurate identification and counting of pauses.

[7] https://www.fon.hum.uva.nl/praat/

After identifying the DFs, the analysis proceeded with a systematic approach using both Excel and the corpus tool Sketch Engine[8]. Initially, total occurrences of DFs were counted in Excel, both in absolute terms and relative to the ST of the lecture. The distribution of these occurrences over the duration of the lecture was then analyzed to identify periods with a higher density of DFs. This analysis included examining the CSs the lecturer employed to repair these instances.

In a separate analysis, the CSs employed by the lecturer to repair these disfluencies were examined. This involved identifying the types of repairs made for different kinds of DFs and quantifying the frequency of these repair strategies throughout the lecture. The goal was to understand the lecturer's approach to managing and correcting disfluencies, and to identify patterns in how repairs were executed in response to specific types of DFs.

Subsequently, time periods within the lecture where DFs were more concentrated were scrutinized further to determine the specific phases of the lecture – defined by the pragmatic functions being performed – where DFs occurred most frequently. Attention was also given to the words co-occurring with these DFs, with a focus on whether these were technical terms or general words, and on identifying the particular types of DFs associated with each word.

For lexical analysis, Sketch Engine was employed to compare the sample lecture text against the British Academic Spoken English Corpus (BASE). The Keyword Analysis feature was used to distinguish words that were significantly more common in the sample lecture than in the reference corpus, indicating they might be technical terms. Words identified as statistically more frequent in the sample were likely domain-specific, while those more common in the BASE corpus were categorized as general

[8] https://www.sketchengine.eu/

academic English. A Word List was then generated for the sample text to assess word frequency, which was compared with the word frequency list from the BASE corpus to classify words as general or technical.

Following this, the Concordance feature of Sketch Engine was used to examine the context in which specific words appeared within both the sample text and the BASE corpus. This step confirmed whether certain words were used in a general sense or within a specialized technical context. Collocations – words that frequently co-occur – were analyzed to determine whether the terms in the sample were associated with technical jargon or general language.

Finally, the co-occurrence of DFs with gestures was analyzed, focusing on how many DF markers co-occurred with gestures and identifying the types of DFs associated with specific gestures and their corresponding functions. This comprehensive analysis provided insights into the intersection of verbal and non-verbal communication in the lecture, shedding light on the challenges faced by the lecturer.

3.3.3 Communicative Strategies

The communicative strategies (CSs) employed by the lecturer were identified and categorized following Schegloff (1997, 2000), integrating his taxonomy of repair strategies with the explicitness strategies identified by Mauranen (2007) and Björkman (2014), as shown in Table 15 below.

1	**Self-initiated self-repair** (**SISR**)	Replacing		
		Inserting	Paraphrasing	
			Topic negotiation	
			Code-switching/ translanguaging	
			Self-Repetition	Parallel Phrasing
				Key Word(s) Repetition
				Combined Repetition
				Repaired Repetition
				Parenthetical Remarks (i.e., definition; description; illustration; comparison)
		Deleting		
		Aborting		

2	**Other-initiated Self-repair (OISR)**	Clarification Requests
		Confirmation Checks
		Correction Prompts
		Elaboration Prompts
		Repetition Prompts
3	**Self-initiated Other-repair (SIOR)**	Request for Assistance
		Open Request for Correction
		Partial Word/Incomplete Utterance
		Tag Question
		Explicit Statement of Uncertainty
4	**Other-initiated Other-repair (OIOR)**	

Table *15*. Overview of the annotations used for identifying the CSs deployed by the lecturer investigated.

We felt the need to integrate these taxonomies because our corpus revealed that the lecturer employed CSs typical of ELF spoken interactions, such as code-switching, topic negotiation, and self-repetitions. While Schegloff's framework provides a clear understanding of the organization of repair, which helps researchers identify the source of *trouble*, we are aware that in the context of English as a Lingua Franca in Academic Settings (ELFA) – to which EMI programs belong – terms like *repair* and *trouble* might raise some eyebrows. However, the purpose of this investigation is to find empirical evidence from the lecturer's communicative conduct to determine the sources of trouble when the lecturer encounters difficulty in *conveying* the intended meaning, both in monologic sequences and in dyadic interactions with students. Additionally, it aims to provide further evidence of the communicative strategies (CSs) that the lecturer, in cooperation with the students, employs to *deliver* the information. In this respect, we emphasize that teaching is not merely a matter of delivering knowledge, as Wellington and Osborne (2001, p. 3) so aptly pointed out:

> We recently read a strategic plan for a university which talks of 'delivering learning' as if it was some sort of package or commodity which is passed on to the student, stored in a kind of pigeon hole and later redelivered to a higher authority when assessment or examinations come around. We hope to show in this book that learning and teaching in science classrooms is (and always has been) a bit more complex than the delivery model,

> mainly because human beings rather than post office sorting machines are involved. The message of the past 20 years of research in science education has been that learners are much more than post boxes.

3.3.4 Gesture Analysis

Hand gestures were coded and analyzed alongside verbal content to identify when the lecturer used gestures to compensate for expressive difficulties and linguistic problems. However, the gesture analysis also aimed to identify instances where the lecturer used hand gestures not only as compensatory devices—a common but overly simplistic assumption—but also to facilitate positive interaction with students. Specifically, we examined how the lecturer's gestures aligned with the students during interactions to find joint solutions, such as eliciting lexical help from the students.

To achieve this, we aimed to identify at what linguistic level facilitation or compensation was occurring—whether at the level of formulating a word, concrete or abstract words, at the conceptual level, or at the interactional level.

Roth (2001, p. 384) pointed out, "gesture studies should be driven by the questions we ask rather than by one or the other research methodology". In line with this perspective, we drew on Kendon (2004) for the classification of gesture types and their description in terms of formal and structural properties, particularly articulators and movement. However, we decided to exclude the space occupied by the hand during the movement from our analysis. Similarly, we did not account for the phases accomplished during the gesture *phrase* – such as preparatory, stroke, and hold – despite recognizing their importance in understanding the precise timing and impact of gestures. Our primary interest was to explore how gestures contribute to the overall meaning of utterances, with a focus on their semantic interaction. To this end, we concentrated on analyzing the hand gestures as a whole, annotating their occurrence from the start (preparation) to the end (return) of each gesture, including some instances of post-stroke movements. This approach allowed us to capture the broader communicative function of gestures within the context of interaction, rather than becoming overly focused on the detailed mechanics of each phase. The table below (Table 16) provides a clearer overview of the types of gestures analyzed and their corresponding functions.

Gesture types	**representational**	iconic metaphoric deictic beats		
	emblems			
	self-adaptors			
Gesture function	interactive			
	pragmatic	modal parse performative		
	referential	concept	facilitation	action descriptive spatial temporal
		word	compensation	

Table *16*. Overview of the annotations used for identifying the types and functions of gestures performed by the lecturer investigated.

While a definition of gesture types and functions has already been provided in Chapter 1 of this study, we will now offer a more detailed explanation, along with examples of the referential function of gestures.

Referential gestures are used to provide specific information or refer to something related to the content being discussed. For example, pointing at a graph on a projector (a deictic gesture) to refer to data or specific points on the graph has a referential function. In this case, the gesture enhances or supports spoken language by reinforcing or clarifying the verbal message, making it easier to understand. Simply put, the gesture facilitates comprehension without replacing spoken words.

For instance, assume the lecturer is explaining the structure of a bridge while pointing at a diagram of the bridge on the screen and says, "The *beam here* supports most of the load". The deictic gesture in this scenario aids in understanding the word "beam" by directly identifying it on the diagram. It clarifies which part of the bridge structure the *word* "beam" refers to, helping students link the spoken term to the correct element of the bridge. As such, similar gestures in our corpus were annotated as *deictic* gestures with a referential *facilitation word* function.

Now, suppose the lecturer is teaching the concept of torque in mechanical systems. While explaining torque, the lecturer points to various positions along a lever and says, "The farther you apply force from the pivot point, the greater the torque." Here, the deictic gesture helps convey the *concept* of torque by visually indicating the distance from the pivot point. By pointing to different areas on the lever, the lecturer reinforces the verbal explanation, aiding students in grasping how the distance affects the force applied. As such, similar gestures in our corpus were annotated as *deictic* gestures with a referential *facilitation concept* function.

Conversely, imagine that during a lecture on structural engineering, the lecturer forgets the technical term for a part of a bridge. The lecturer points at the suspension cables on a diagram and says, "The... *erm...* these things hold the weight of the deck." In this case, the pointing gesture compensates for the forgotten technical term (*suspension cables*) by indicating the part of the bridge being discussed. The gesture fills the gap left by the forgotten word, manifested, in this case, through speech DFs such as filled pauses and unfilled pauses. As such, similar gestures in our corpus were annotated as *deictic* gestures with a referential *compensation word* function.

Finally, consider a lecturer explaining a complex fluid dynamics concept but struggling to describe it verbally. The lecturer moves their hands in a circular motion to illustrate the swirl of fluids in a pipe, saying, "*It's hard to explain*, but *it's like* this flow that goes in a circular motion around the pipe." Here, the gesture compensates for the difficulty in verbally explaining the *concept* of vortex flow. By mimicking the swirling motion with their hands, the lecturer visually represents the concept of fluid circulation, compensating for the lack of a precise verbal explanation, manifested, in this case, through CSs such as SIOR, namely Explicit Statement of Uncertainty. As such, similar gestures in our corpus were annotated as *deictic* gestures with a referential *compensation concept* function. Overall, facilitation gestures enhance or reinforce the verbal message, helping clarify or emphasize the concept being discussed, whereas compensation gestures occur when the spoken word is lacking, unclear, or not fully expressed, and the gesture *compensates* by providing the missing information. Furthermore, *conceptual*-related gestures are often broader, focusing on the entire idea or content being conveyed, whereas *word*-related gestures are more specific, addressing gaps or ambiguities in the actual words used.

In the taxonomy used to annotate gesture types, we also included *self-adaptor* gestures – gestures where individuals unconsciously touch or manipulate their own body, objects, or clothing. These gestures often serve as a way to manage emotions or stress and are not necessarily aimed at communication. Examples include scratching, fidgeting, rubbing one's face, or glancing at one's watch. In the context of a lecture, glancing at one's watch, for example, might indicate a sense of time pressure. A lecturer feeling pressed for time may experience additional stress to deliver content clearly within the allotted period. This could influence their speech delivery, causing them to rush through explanations or skip important points in an effort to stay on schedule.

Observing self-adaptor gestures in studies on lecturers – especially in a multilingual context like EMI – can provide valuable insights into their emotional and cognitive states. Although these gestures do not have a semantic relationship with the content being taught, they may reveal cognitive overload that could indirectly affect the quality of the teaching. For instance, a lecturer feeling rushed might speed up their speech or cut down on explanation time for complex topics, potentially diminishing the effectiveness of the lecture. Additionally, while these gestures are not inherently communicative, they are still interactional, as they can convey meaning or affect the dynamics of engagement with students.

Furthermore, when referential gestures were used to facilitate or compensate for either a word or a concept, annotation was performed to determine the type of word or concept – i.e., descriptive, action, temporal, or spatial – as explained in the table provided below (Table 17).

Level of linguistic facilitation/compensation	Definition	Example
Spatial	Gestures that represent spatial relationships, directions, or locations.	The gesture of moving the hand straight forward likely represents the spatial location of the downstream wall.
Descriptive	Gestures that describe the characteristics or features of an object or concept.	Using hand shapes to describe the size or shape of an object.
Action	Gestures that depict actions or movements.	Mimicking the flow of water, indicating a path or movement direction.
Temporal	Gestures that represent the passage of time, sequences, or durations	Moving a hand from left to right to represent the progression of events from past to future.

Table *17*. Gesture functions categorized by the level of linguistic facilitation or compensation they provide.

An example is provided in Figure 3, where the lecturer moves his hands, palms down, moving downward while saying, “in order to try to *accompany* water downstream”. The gesture here performs a *referential* function, aiming to *facilitate* the explanation of the *downstream flow process*. While it co-occurs with the verb *accompany*, it refers to the concept of downstream water flow – i.e., the movement of water in a river, stream, or other body of water in the direction the current naturally flows, typically from a higher elevation to a lower one. Indeed, he moves his hands downward to simulate that action.

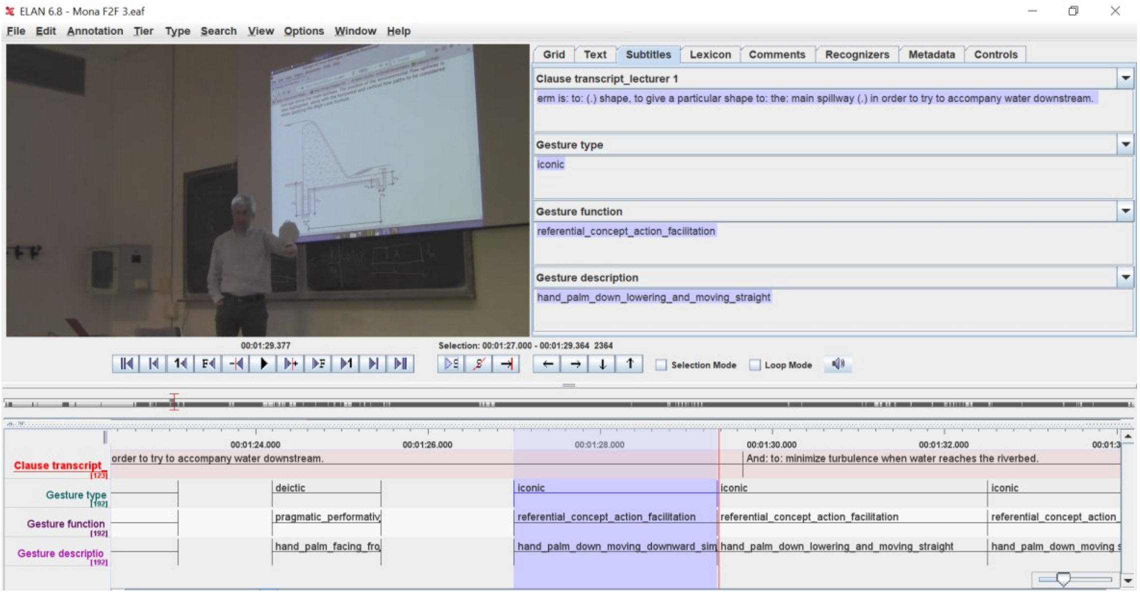

Figure 3. Example of a gesture function annotation categorized by the level of linguistic facilitation or compensation. In this case, the lecturer performs an action facilitation gesture.

Regarding gesture description, gestures were also annotated by examining their structural properties. The focus was primarily on the articulator involved (e.g., hand/hands, finger/fingers, fist, palm/palms), the movement performed (e.g., pointing, turning, circling, moving), and the direction (e.g., up/down, back/forth, inward/outward). For example, the figure below (Figure 4) shows the lecturer performing a beat gesture that serves a pragmatic parsing function—i.e., helping to structure the discourse by breaking speech into manageable chunks, marking transitions, emphasizing points, or aiding in the rhythm of speech. Beats, in particular, serve this purpose. In this instance, the lecturer moves both hands laterally, with the backs of the palms facing forward, moving left and right. The resulting annotation is: hands_palms_side_facing_front_moving_left_right.

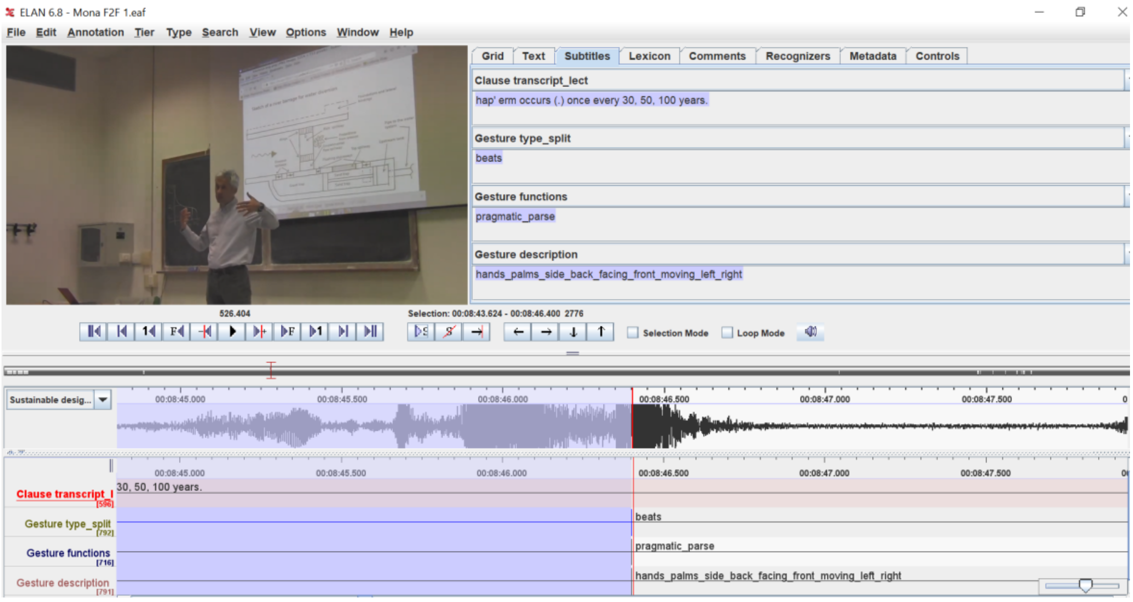

Figure 4. Example of gesture annotation on ELAN.

Annotating gestures on ELAN involved several iterations of re-elaboration and re-labeling. We began by analyzing a small, 9-minute sample of the lecture. After annotating this sample, we exported the annotations into a CSV file and created a controlled vocabulary list, which we then re-imported into ELAN to annotate the remainder of the lecture. This approach aimed to ensure higher consistency across the data. The structure of the gesture (i.e., the "gesture description" tier) was annotated with the intent of identifying how many instances of the same gesture the lecturer performed, analogous to identifying word tokens and types in text analysis.

As the sole annotator, ensuring both intra-rater reliability (i.e., the consistency of one's own annotations over time) and the quality of the final annotations was critical. To achieve this, we followed a systematic process. The file was annotated in multiple phases: initially, the dataset was annotated, and after a delay of two weeks, a subset of the data was re-annotated without referencing the original annotations – a process known as Bootstrap Validation. Consistency in our annotations was statistically measured across these bootstrapped samples using Cohen's Kappa, a statistical measure used to evaluate the degree of agreement between two sets of annotations (or in this case, two rounds of annotations by the same individual) beyond what would be expected by chance. Bootstrapping was repeated twice on samples of 300 annotations each, with the κ value ranging between 0.61 and 0.80, indicating substantial agreement. However, for "gesture description", the κ value was lower than 0.61. We nevertheless decided not to pursue further validation of the gesture description annotations, as consistent annotation of gesture types and functions was deemed more crucial for the scope of our analysis. The same procedure was followed for the validation of the annotation of the CSs and DFs.

Additionally, while our analysis predominantly focused on the examination of hand gestures, we extended our annotation to include other relevant actions performed by the lecturer, such as walking, writing on the board, and other forms of embodied activity. Since the lecturer was recording himself, he consistently remained in the public space, occasionally moving into the social-consultative space (Lim et al., 2012) – near the blackboard and the classroom projector – to ensure he was always on camera. As a result, analyzing spatial movement might have been less meaningful. Nevertheless, we annotated, albeit not systematically, the embodied actions he performed beyond speaking and gesturing, such as walking, sitting down, standing up, and writing on the board, to contextualize DFs occurrences, which might have been due to multitasking.

Furthermore, although we did not analyze lecturer's gaze, and spatial and orientational arrangements systematically, we considered them when the function or type of gesture was difficult to discern. For example, when the lecturer's gestures were in reference to an inscription (e.g. photos, graphs, diagrams, maps, etc.) drawn on the blackboard or projected onto a screen, distinguishing between a deictic, an iconic or a metaphoric gesture was easier by observing lecturer's gaze, as well as his body orientation, and spatial arrangement with respect to the inscription and to the students. In this respect, Roth and Lawless (2002) observed that when performing gestures around an inscription, the same gesture may be deictic if performed in the "graphic space" (p. 19), where the lecturer "normally limits themselves to the two dimensions spanned by the inscription" (p. 19). However, the same gesture may become iconic if the speaker shifts their entire body "into a position in narrative space" (p. 18), thereby "exploiting three dimensions" (p. 18) to "articulate the typology" (p. 18) of the term or concept being discussed. In such a case, the lecturer's gesture has "referents both in the lived world" and the diagram (p. 27). This shift is immediately noticeable as the speaker visibly rotates their body to face the audience, as shown in Figures 5 (*deictic* gesture) and 6 (*iconic* gesture) below. In Figure 5, the lecturer deictically points to a graphical point in the diagram displayed on the projector. In contrast, in Figure 6, the lecturer moves his hand, palm down, backward along an imaginary horizontal line, simulating the length of a pool. In this case, his gesture makes use of the three-dimensionality of the local space. Furthermore, the utterance "both in width (.) and length" and the gesture are continuous. As Streeck (2009, p. 24) further pointed out:

> Gesture itself symbolizes by constructing structures in space. The shape, orientation, trajectory, and motion pattern of a gesture can therefore not be explained solely by reference to the content that is conveyed through talk and gesture; rather, the features of gestures and the effects that they achieve are also contingent upon the locale and spatial organization of the encounter.

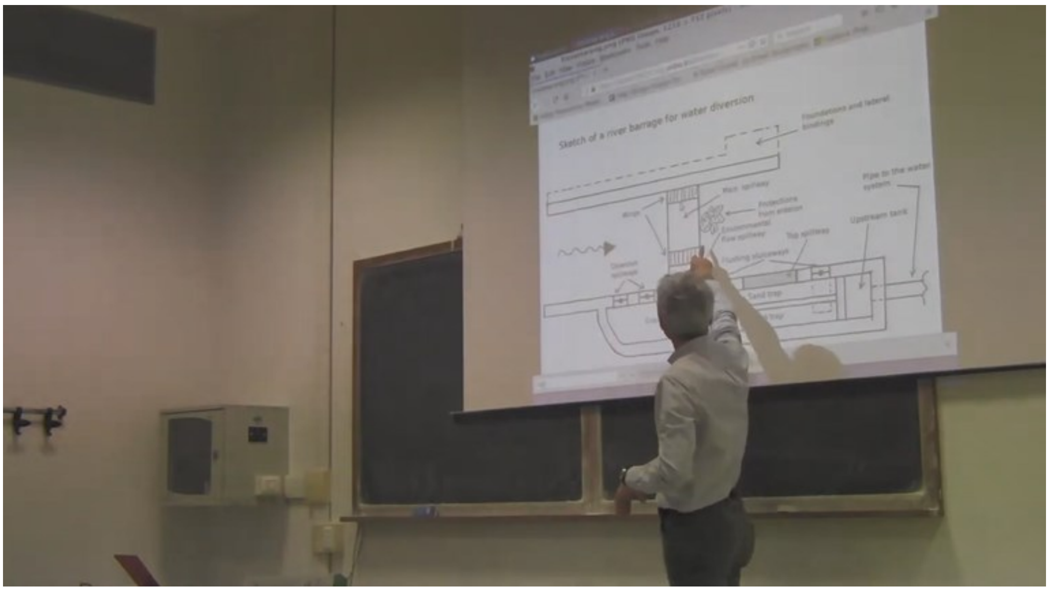

Figure 5. Example of a deictic gesture performed in the graphical space of the inscription.

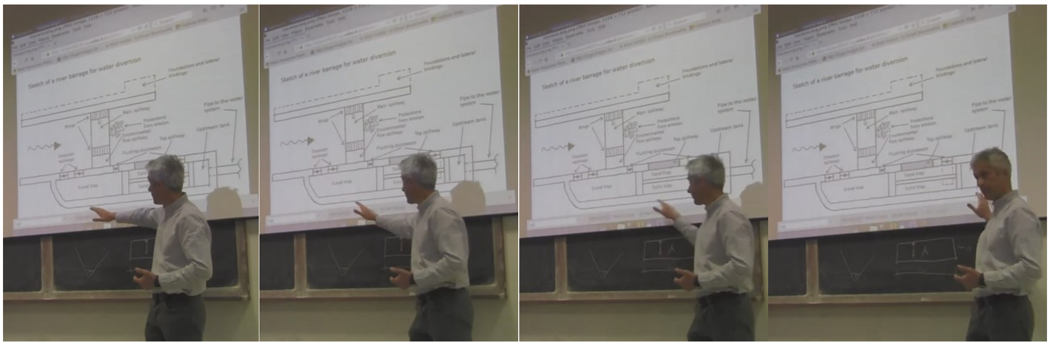

Figure 6. Example of an *iconic* gesture (mimicking the length of a pool) performed in the narrative space of the inscription. The gesture co-occurs with the utterance "both in width (.) and length". The lecturer's gaze and body orientation serve as key indicators for identifying this iconic gesture.

Gesture analysis is a complex and demanding task, one that current NLP and AI technologies are not yet fully capable of executing. The interpretation of meaning conveyed through articulators, which symbolize concepts by constructing spatial structures, remains a uniquely human ability.

3.4 Ethical Considerations

Although the lecturer uploads his lectures on YouTube, making them freely accessible and not subject to copyright restrictions, we nonetheless obtained his explicit consent for their use in this study. Additionally, the lecturer was informed of his right to withdraw from the research at any time, and all students involved were anonymized to protect their identities. The study strictly adheres to ethical research guidelines, ensuring confidentiality throughout the research process.

3.5 Limitations

While this study offers valuable insights into the experiences of one lecturer and a cohort of students within an EMI setting, certain limitations should be acknowledged. Firstly, while the single-case study design allows for a deep, context-specific exploration, it may limit the generalizability of the findings to other disciplines, departments, or universities. Different lecturers and student cohorts in other fields or institutions may face unique challenges that are not reflected in this specific case. Secondly, the lecturer's responses and reflections may be influenced by self-perception and personal biases. Since the study relies on self-reported data from the lecturer, there may be an inclination to downplay certain challenges or emphasize strengths, which could limit the objectivity of the findings. Thirdly, the study is conducted within a specific educational and cultural context at an Italian university, focusing on engineering courses. EMI environments vary significantly across regions, disciplines, and educational systems. The experiences and strategies described here may differ in settings with a higher or lower prevalence of EMI, such as universities in countries with greater English proficiency or distinct pedagogical approaches. Fourthly, by focusing on a single lecturer's approach, the study limits its scope to that individual's teaching methods and experiences. Other lecturers may use different techniques, face distinct challenges, or have varying levels of EMI training and experience, which could provide alternative perspectives on EMI strategies. Fifthly, the study provides a snapshot of the lecturer's and students' experiences at a specific point in time. However, the evolving nature of language proficiency and teaching practices could be better understood through a longitudinal approach, which might reveal shifts in lecturer strategies, student engagement, and the effectiveness of different pedagogical methods. Sixth, while the study investigates the lecturer's views on EMI, it does not deeply explore the broader institutional or policy context influencing EMI practices. Factors such as faculty development opportunities, administrative support, and institutional policies on EMI could significantly impact the implementation and success of EMI but remain underexplored in the analysis. Seventh, although student feedback is collected on various aspects of the course, it would be valuable to gather more detailed data on how students perceive specific EMI strategies used by the lecturer (e.g., interactive methods, the use of open resources). Understanding the direct impact of these strategies on student learning could offer a more comprehensive evaluation of their effectiveness. In this respect, the survey responses do not account for potential differences in learning outcomes between students of varying proficiency levels. Without a detailed analysis of how these language barriers impact academic performance, the findings may overlook important nuances.

These limitations point to areas where further research or data collection may be needed to develop a more complete understanding of the challenges and opportunities in EMI teaching.

Chapter 4. Findings and Discussion

Although this study focuses on analyzing the communicative practices of one EMI lecturer in the specific field of engineering, the analysis is informed and driven by students' perceptions of their major challenges and needs. In this section, the key findings of the research will be presented, addressing the main RQs which focus on both the lecturer's and students' perceptions and the cross-influence of speech rate, disfluency markers (DFs), communication strategies (CSs), gestures, and speech during the EMI lecture investigated. Specifically, the following RQs will guide the analysis:

RQ1. What are the lecturer's perceptions regarding their own teaching performance in EMI, and how do they perceive students' needs and major challenges?

This question explores the lecturer's self-reflection on teaching practices and challenges in the EMI context, offering insight into how he adjusts to students' learning needs.

RQ2. What are students' perceptions of their lecturer's teaching performance in EMI engineering classes, and what challenges do they report?

This focuses on gathering students' views on the effectiveness of the lecturer's teaching strategies and identifying their own challenges in comprehending EMI instruction.

RQ3. What are the key misalignments between the lecturer's perceptions of his teaching performance and the students' reported challenges in the EMI context?

This question investigates the discrepancies between the lecturer's self-assessment of their teaching practices and the challenges students face in comprehending EMI instruction. It aims to identify areas where the lecturer's strategies may not fully align with students' linguistic and cognitive needs, with the goal of translating these perceptions into linguistic terms and providing evidence to support these interpretations.

RQ4. During which pedagogical phases does the lecturer tend to alter speech rate (e.g., speed up or slow down)? What pragmatic functions correspond to these shifts in speech rate?

This question links two key dimensions: speech rate and the lecturer's pragmatic functions, aiming to identify patterns in how speech rate shifts according to the lecture's structure.

RQ5. How do slower speech rates correspond to pauses, and are these pauses markers of (dis)fluencies (DFs)? How are these markers distributed throughout the lecture, and with which pragmatic functions do they tend to occur most frequently?

This question investigates the relationship between slower speech, disfluency markers and pragmatic functions. It aims to map out where disfluencies tend to occur, and which pragmatic functions are associated with these markers.

RQ6. How are DF markers and CSs connected, and what repair strategies do the lecturer and students use to address DFs?

This question focuses on examining the lecturer's and students' strategies for managing DFs, connecting them to broader communication strategies. It aims to explore the relationship between DFs and CSs, particularly how these strategies are employed during the lecture to maintain communication flow.

RQ7. Can the lecturer's gestures reveal more about linguistic challenges, such as difficulties in word formulation or conceptualization? How do gestures interact with the lecturer's verbal explanations?

This question examines the connection between the lecturer's gestures and linguistic challenges, exploring how body movements (i.e., gestures) complement or reveal difficulties in speech production and concept delivery.

The analysis will identify the cross-influences among five major categories – speech rate, DF markers, CSs, gestures, and speech – which will be presented and discussed in detail. Each research question aims to deepen the understanding of how verbal and non-verbal strategies affect communication during EMI instruction.

4.1 The Lecturer's Views and Misalignments

The lecturer investigated in this study is a full professor of Water Engineering and Hydrology at the DICAM department of the University of Bologna since 2012. He is an Italian L1 speaker aged 55-60 years old, and he is currently teaching five courses all for master's degree programmes, among which four are taught in English. Furthermore, as evidenced by his personal website (https://www.albertomontanari.it/opinions) not only is the lecturer used to writing in English both for professional and academic purposes, but also for general communication purposes.

The interview responses from the lecturer provide rich insights into his approach to teaching in an EMI context, revealing both the strategies he employs and the challenges he faces. Below is a comprehensive analysis, integrating observations on the lecturer's experience, teaching practices, and reflections on student engagement.

4.1.1 Lecturer's Experience and EMI Training

The lecturer has more than 10 years of experience teaching in English but has not attended any formal EMI training courses. Despite not participating in EMI-specific training, he feels his extensive experience compensates for this, believing that such training might only offer marginal benefits. The lecturer's confidence in handling technical terminology is evident, as he prioritizes using open and regularly updated digital resources over traditional textbooks, which he views as quickly becoming outdated. He reports holding a C1 level of English proficiency, which supports his perceived ability to manage lectures effectively in a non-native language.

4.1.2 Perceptions of Interactive Engagement and Its Importance

The lecturer emphasizes that interactive engagement is more critical in international programs than in Italian ones, due to greater linguistic diversity – both between the lecturer and the students, as well as among the diverse cohort of students, who present not only different L1s but also varying levels of English language proficiency. In EMI settings, interaction becomes a key tool for overcoming language barriers and enhancing comprehension. The lecturer sees interaction as a facilitator of learning, allowing students to clarify doubts and learn from each other, which is especially important when students may not fully grasp the language being used in instruction. This view aligns with broader EMI research, which highlights the effectiveness of interactive and student-centered teaching in contexts where linguistic challenges are prevalent. In the lecturer's view, his proactive use of strategies such as asking questions, encouraging peer discussions, and inviting students to present reflects his awareness of how interaction helps students overcome their fear of participation and improves their understanding.

4.1.3 Class Size and Student Participation

The lecturer's classes usually consist of 25 to 50 students, with 50 to 75% of them being international students. In the context of EMI and university-level education, smaller class sizes have been shown to enhance personalized engagement and active learning strategies. For example, research suggests that instructors in smaller university classes more frequently use learner-centered activities such as group work, simulations, and case studies, resulting in higher student engagement and retention (see e.g., Wright, et al. 2019). However, the lecturer still notes challenges in fully engaging students, attributing this to factors like cultural norms, language anxiety, and students' hesitation to ask questions due to fear of making mistakes or appearing overly eager.

This observation is critical because it highlights how even in smaller groups, factors beyond class size, such as linguistic and cultural dynamics, significantly impact student participation in EMI settings. The lecturer's continuous efforts to encourage interaction suggest a deep understanding of these dynamics, as he tries to create a classroom environment where students feel comfortable participating.

4.1.4 Adapting Teaching Materials for EMI

The lecturer has recognized that international students often find the slides too lengthy and overwhelming. In response, he has begun creating more concise and targeted content. This is particularly important for students who may struggle with English proficiency. The lecturer reported that approximately 80% of the slides are created specifically in English, while the remaining 20% are translations from Italian. This transition highlights the complexities involved in adapting teaching materials for a diverse student body. Additionally, the lecturer's increased reliance on slides and digital tools in EMI reflects an effort to provide a clear structure that supports both his delivery and students' comprehension. He recognizes that "the screen" (e.g. a projector connected to the classroom pc, slides) provides him with a helpful guide, particularly in an EMI context where both lecturers and students may not have a high level of proficiency in the language.

4.1.5 Interactive Methods and Time Management

The lecturer incorporates interactive methods as a core part of their teaching approach, recognizing the value of student engagement. However, these activities require additional time, prompting the lecturer to allocate two extra hours outside the regular schedule for interactive sessions. This reflects a strong commitment to maintaining student engagement while acknowledging the challenge of balancing content coverage with active learning. The decision to extend class time highlights the lecturer's understanding of the pedagogical value of interaction, even if it demands additional effort. This also underscores a common challenge in EMI: finding a balance between delivering comprehensive content and providing opportunities for interactive, student-centered learning.

4.1.6 Use of Technology and Open Resources

The lecturer is a strong advocate for open access resources and uses platforms like Wikimedia Commons[9] to overcome copyright issues. He believes that textbooks are outdated and prefers to use digital materials that can be continuously updated. This approach aligns with his commitment to providing accessible, flexible learning resources that can be easily adapted to changing needs. The lecturer's emphasis on open information also extends to his preference for digital platforms over print materials, enabling students to access updated resources even after the course ends.

4.1.7. Challenges and Institutional Support

The lecturer highlights the challenges associated with engaging in innovative teaching practices, such as a lack of immediate recognition and institutional support. Despite these obstacles, he continues to push for interaction and open information sharing as he believes it benefits both his teaching and the department's visibility. The lecturer points out that while some colleagues resist innovative practices, others have found value in these approaches and have begun to embrace them. The lecturer also observes that increasing opportunities to engage with colleagues in discussions about teaching practices, rather than focusing solely on administrative tasks, could lead to a more positive and enriching experience. He emphasizes the importance of

[9] https://zh.wikipedia.org/wiki/c:

recognition and support from the institution, suggesting that acknowledgment of these efforts can be a powerful motivator for sustained engagement in innovative practices in EMI.

The lecturer's responses provide a detailed and reflective account of the complexities of teaching in an EMI context. He demonstrates a clear understanding of the importance of interaction, adaptability, and the use of open resources to address the linguistic and cultural challenges faced by international students. His proactive approach to encouraging engagement, along with his commitment to maintaining open and flexible learning materials, offers valuable insights into effective EMI teaching practices. Furthermore, his willingness to adapt teaching methods and materials to meet the needs of diverse learners highlights his continuous efforts to create inclusive and effective learning experiences. However, the lecturer also acknowledges ongoing challenges related to student participation, the diverse linguistic, cultural, and academic backgrounds of students, as well as time management and institutional support.

4.2 Students' perceptions and misalignments

As part of two previous studies (Picciuolo & Johnson, 2020; Johnson & Picciuolo, 2023) investigating engineering students' self-perceived challenges and needs in EMI at the targeted university, student feedback was gathered through online surveys distributed by lecturers who agreed to assist the researchers between 2019 and 2023. Both surveys focused on specific areas such as the lecturer's speed of delivery, clarity of speech, understanding of technical terminology, and the perceived effectiveness of other semiotic resources beyond the lecturer's speech (e.g., slides, inscriptions displayed on the projector).

Students were asked to reflect on their experiences during EMI lectures, not limited to the lecturer investigated in this study. By collecting perspectives from students across other courses within the same department, the study gains a broader understanding of the challenges faced in this EMI setting and how these challenges are perceived differently by lecturers and students. The students' responses revealed valuable insights into their experiences and perceptions of EMI courses in engineering programs. Below is an analysis of key observations and trends that emerged from the student survey data.

4.2.1 Student Background and Language Proficiency

The respondents come from diverse linguistic backgrounds, including Italian, Persian, Urdu, Chinese, and Norwegian, among others. Most students have a B2 or C1 level of English proficiency, indicating that they have a solid command of the language, with a few students at a B1 or C2/native speaker level.

4.2.2 Comfort with EMI Courses

Many students report feeling either "very comfortable" or "fairly comfortable" with following courses in English. This comfort level is generally linked to their previous experience with English or familiarity with international academic environments.

Some international students mention challenges related to specific accents (e.g., Italian) or pronunciation differences that make comprehension more difficult, especially for those accustomed to American or British English accents.

4.2.3 Perception of Lecturers' Pronunciation and Clarity

The feedback on lecturers' pronunciation varies, but many students rate it as "fairly clear" or "very clear." However, there are consistent mentions of challenges with understanding due to accent or mispronunciations, especially among students with lower levels of English proficiency (B1/B2). In this respect, at the DICAM department, students are required to pass a B2-level English language test administered by the university's Language Center or provide proof of having an English language certification of at least B2.

In cases where the lecturer's pronunciation is perceived as unclear, students typically resort to coping strategies such as asking their peers for clarification, trying to guess the meaning from context, or taking notes to check later.

4.2.4 Usefulness of Visual Materials

A significant number of students find visual aids like PowerPoint slides, black/whiteboard writing, and handouts useful for following the lecture. Visual materials are frequently cited as essential tools for bridging gaps in comprehension, especially for students who might struggle with oral language processing.

Additionally, while students appreciate concise and well-organized slides, some, particularly international students, note that overly detailed slides can be overwhelming.

4.2.5 Suggestions for Improvement

Common suggestions for improving lectures include: incorporating more real-life examples and practical exercises; slowing down the pace of delivery, especially for complex topics; offering recordings of lectures to allow students to revisit content at their own pace; providing breaks during long lectures to maintain attention and improve retention.

4.2.6 Interaction and Engagement in Class

While interaction was not explicitly assessed in the responses, it's clear from the comments that students value opportunities for clarification, either through peer discussions or directly with the lecturer. A few students express frustration with the fast pace or unclear explanations during lectures, indicating that there may be a need for more interactive elements like quizzes or small group discussions to enhance understanding.

4.2.7 Cultural and Linguistic Challenges

Several students point out that while they are generally comfortable with English, the Italian accent can present difficulties. This is particularly relevant for non-European students who may have limited exposure to European varieties of English.

Some international students mention that despite being comfortable with English, they occasionally struggle with idiomatic expressions or informal language used by lecturers.

4.2.8 Student Strategies for Managing Difficulties

When students encounter challenges, they often rely on peer support, either by asking questions or discussing unclear points after class. Others prefer independent strategies such as taking detailed notes to review later or guessing the meaning based on context. The survey data highlights that most students are generally comfortable with EMI courses, but consistent challenges arise related to accent, pace, and the clarity of explanations. Students from diverse linguistic backgrounds, in particular, reported facing additional barriers due to differences in pronunciation or unfamiliar academic vocabulary. In this regard, the survey responses reveal noticeable differences across students' language proficiency levels and first languages (L1). Below are the key differences based on these variables.

4.2.9 Differences Across Language Proficiency Levels

Students with higher proficiency (C1/C2) generally report being more comfortable and confident in following EMI courses. They tend to have fewer issues with pronunciation, pace, or content comprehension. They often describe their experience as “very comfortable” or “fairly comfortable.” Nonetheless, even at high proficiency levels, students occasionally mention challenges related to specific accents (e.g., Italian-accented English), though these are more minor and do not significantly hinder their understanding. Furthermore, since high-proficiency students' feedback is more likely to emphasize improvements in content delivery (e.g., more examples, clearer slides) rather than linguistic aspects, this suggests that they are more focused on the lecture content rather than language issues.

Students with lower English proficiency (B1/B2) more frequently report difficulties with understanding lecturers due to pronunciation, accent, and speed. Italian-accented English and fast-paced speech are recurring themes that contribute to comprehension challenges for this group. Lower-proficiency students also rely more heavily on visual materials, handouts, and slides. They often suggest adding more structured resources (like glossaries) and express a preference for simplified slides with less content per slide. As regards coping strategies, B1/B2 students tend to rely on strategies like peer support, guessing from context, and taking detailed notes to revisit later. They are more likely to express frustration if these strategies do not fully address their comprehension issues.

Nevertheless, differences across L1s also emerged from the students' replies. L1 speakers of English and students from geographical contexts known for high-quality English education and widespread use of English in media and daily life (e.g., Norway) typically express little to no difficulty in following the lectures. For these students, any challenges they face are more likely related to content delivery or teaching style rather than language comprehension. In this respect, these students sometimes note that lecturers' English could be improved, particularly in terms of pronunciation or choice of words. They express that these issues occasionally make it challenging to focus on content, especially when the lecturer's English level is perceived as lower than theirs.

Conversely, students from non-European backgrounds (e.g., Persian, Chinese, Urdu speakers) more frequently mention difficulties with the Italian accent, pronunciation, and specific language issues. These students are more likely to cite problems with understanding due to lecturers' mispronunciations or intonation that doesn't align with what they're used to (e.g., American or British English). For these students, the language barrier is often intertwined with content comprehension. They express a need for clearer, more deliberate explanations and prefer a slower pace to allow them time to process both the language and the subject matter. Additionally, these students are more likely to request additional support resources, like lecture recordings, glossaries, or simpler slides, to help bridge the gap between their linguistic challenges and the technical content of the course.

From the student survey, we may observe combined effects of L1 and English language proficiency. Particularly, the challenges faced by students tend to be more pronounced among those with a combination of lower English proficiency and non-European L1 backgrounds. For instance, a B2-level student whose L1 is Chinese or Persian may face multiple layers of difficulty, including pronunciation differences, unfamiliar vocabulary, and issues with pace. These students tend to request more visual aids, structured outlines, and lecture recordings. They rely more on independent strategies like revisiting notes or asking peers after class to compensate for what they miss during the lecture.

4.3 Misalignments Between Lecturer's Perceptions and Students' Perceptions

The analysis reveals several misalignments between the lecturer's perceptions and those of the students regarding key aspects of EMI instruction.

One notable area of divergence lies in the *speech rate* during lectures. The lecturer, while focusing primarily on content delivery and student interaction, does not emphasize speech rate as a significant issue. In fact, he perceives his ability to deliver technical content effectively in English as sufficient, especially given his self-reported proficiency and experience in the EMI context. However, students, particularly those with lower English proficiency (B1/B2) or from non-European backgrounds, frequently highlight the pace of the lecturer's speech as a challenge. These students often express difficulty in keeping up with the rapid delivery of complex material, with several suggesting a slower pace to allow for better processing of both language and content. In contrast, higher-proficiency students, especially from geographical regions with strong English education systems (e.g., Norway), report fewer issues with speech rate and instead focus on the quality of content and teaching style.

Another significant misalignment is observed in the *clarity of pronunciation and explanations*. The lecturer seems confident in his ability to communicate technical content in English, likely supported by his perceived C1 level of proficiency and years of teaching experience. He acknowledges the importance of interaction but does not explicitly address pronunciation as a barrier in his teaching practice. However, students from diverse linguistic backgrounds, particularly those from non-European contexts (e.g., Persian, Chinese, or Urdu-speaking students), often report difficulties in understanding the lecturer due to his Italian-accented English. These students mention that mispronunciations, unfamiliar intonation patterns, and the overall clarity of speech sometimes impede their comprehension, particularly when compounded by the fast pace of the lecture. In some cases, these students rely on external coping strategies, such as asking peers for clarification or reviewing notes after class, to mitigate these challenges.

In terms of *student engagement and interaction*, the lecturer views interaction as a critical component of his teaching, especially in international programs where linguistic diversity is higher. He actively encourages questions, peer discussions, and student presentations, believing that such methods help students overcome language barriers and enhance comprehension. Despite these efforts, some students, particularly those with lower proficiency or from different linguistic backgrounds, may still feel hesitant to participate due to language anxiety, fear of making mistakes, or unfamiliarity with interactive pedagogical practices. These students might appreciate the lecturer's efforts but continue to struggle with participation, which they attribute more to linguistic challenges than to the interactive teaching methods themselves.

Overall, while the lecturer focuses on delivering content and fostering interaction, students more frequently highlight speech rate and clarity during the explanation of complex topics as key areas for improvement.

For students, especially those from non-European backgrounds and with lower English proficiency, these aspects of the lecturer's performance significantly impact their ability to comprehend the material and engage effectively in the classroom. On the other hand, students with stronger English backgrounds or higher proficiency levels tend to focus less on language issues and more on content delivery and the effectiveness of teaching methods. This misalignment between the lecturer's perceptions and the students' challenges underscores the need for deeper investigation in these areas. The following sections present the findings from our analysis, focusing particularly on these aspects.

4.4 Speech rate

In order to address *RQ4. During which pedagogical phases does the lecturer tend to alter speech rate (e.g., speed up or slow down)? What pragmatic functions correspond to these shifts in speech rate?*, we first analyzed the lecturer's speech rate. As outlined in the research methodology section in Chapter 2, we used Python to code and measure the lecturer's speech rate using three distinct fluency metrics: Mean Syllables per Run (MSR), Rate of Speech Time (ROST), and three ratios – Speech Time Ratio (STR), Silent Pause Time Ratio (SPTR), and Filled Pause Time Ratio (FPTR).

First, annotations regarding the lecturer's speech, which had been tokenized and marked with DF markers, were exported from ELAN. Then, the data were processed from an XLS file containing the lecturer's transcribed speech, which was automatically syllabified using Python (though online tools are also available for this task). Silent pauses were annotated orthographically using "()" and "(.)" to facilitate the count of speech runs.

Additionally, filled pauses, marked as "fp", and unfilled pauses, marked as "unp", were considered. Disfluency sequences (e.g., lengthening, stuttering) that included filled or unfilled pauses – such as lengthening or repetition combined with "fp" or "unp" – were classified according to the type of pause present. If a filled pause occurred in the sequence, it was considered a filled pause ("fp"). Similarly, if the sequence included an unfilled pause, it was marked as an unfilled pause ("unp").

In cases where both filled and unfilled pauses appeared within a disfluency sequence, the pause was classified as an unfilled pause ("unp") since these typically last longer.

Table 18 below reports the findings for each of the resulting measures. Time is expressed in both milliseconds (msec) and minutes (min) for all relevant measures.

Speech Rate Measure	**Description**	**Value (msec)**	**Value (min)**
Speech Time (ST)	Total duration of meaningful speech	3,867,597	**64.46**
Silent Pause Time (SPT)	Total duration of silent pauses	1,291,612	**21,53**
Filled Pause Time (FPT)	Total duration of filled pauses	65,837	**1.10**
Total Time (TT)	Total time (ST + SPT + FPT)	5,225,046	**87.08**
Mean Syllables per Run (MSR)	Average number of syllables between pauses	13,05	
Rate of Speech Time (ROST)	Syllables articulated per millisecond of speech time	0,003614648	
Speech Time Ratio (STR)	Percentage of time spent speaking	74%	
Silent Pause Time Ratio (SPTR)	Percentage of time spent in silent pauses	25%	
Filled Pause Time Ratio (FPTR)	Percentage of time spent in filled pauses	1%	

Table *18*. Speech rate measures resulting from the analysis of the lecturer's speech.

The data provided in Table 18 presents a comprehensive analysis of the temporal characteristics of the lecturer's speech, offering valuable insights into both the fluency and rhythm of the speaker. The total speech time (ST), which measures the duration of meaningful speech, amounts to 3,867,597 milliseconds, or approximately 64.46 minutes. This substantial figure indicates that the speaker maintained an extended period of active speech, comprising nearly three-quarters of the total observed time. Indeed, the speech time ratio (STR) confirms that 74% of the overall time was dedicated to the articulation of meaningful content, reflecting a predominant engagement in verbal communication throughout the session.

Conversely, the silent pause time (SPT) stands at 1,291,612 milliseconds, equivalent to 21.53 minutes, which constitutes approximately 25% of the total duration. The silent pause time ratio (SPTR) mirrors this figure, showing that a quarter of the time was spent in silence. While pauses of this nature are integral to natural speech, allowing the speaker moments for cognitive processing and breath control, the balance between speech and silence suggests a measured and thoughtful delivery, rather than a rapid or pressured manner of speaking.

The filled pause time (FPT), measuring the duration of non-verbal fillers such as "erm", is recorded at 65,837 milliseconds, or 1.10 minutes. This minimal time spent on filled pauses, accounting for only 1% of the total duration, indicates a high degree of fluency. The speaker's low reliance on fillers may reflect confidence, preparation, and cognitive clarity, all of which contribute to a perception of polished and uninterrupted communication.

The total time (TT), which aggregates the speech time, silent pause time, and filled pause time, reaches 5,225,046 milliseconds, or 87.08 minutes. This figure provides an overall context for the session, encapsulating both active verbalization and the intermittent pauses. In this respect, the 3-minute discrepancy between the total lecture time and the TT resulting from the analysis is due to the inclusion of combinations of DFs, which encompass unfilled pauses, filled pauses, and other DF markers not typically counted in such speech measures. Nonetheless, this discrepancy does not affect the relevance of the findings. The extended duration of speech within this total time reinforces the speaker's capacity for sustained articulation, while the controlled use of pauses punctuates the speech without hindering its flow.

The mean syllables per run (MSR) measure stands at 13.05, indicating the average number of syllables spoken between pauses. This figure suggests a moderate pace of speech with regular interruptions for pauses, allowing the speaker to articulate several words before requiring a moment of silence. Such a pattern is indicative of a well-paced delivery, balancing fluidity with natural pauses for emphasis or reflection.

The rate of speech time (ROST), which calculates the number of syllables articulated per millisecond of speech time, is recorded as 0.003614648 syllables per millisecond. This rate, when converted, equates to approximately 3.6 syllables per second, which falls within the range of a typical conversational speech rate. This suggests that the speaker's pace is neither rushed nor excessively slow, allowing for clear articulation while maintaining a rhythm appropriate for understanding.

In summary, the data presents a speaker who engages in extended periods of meaningful speech, with a well-balanced use of silent pauses and minimal reliance on filled pauses. The overall temporal structure of the speech reflects a controlled and fluid delivery, characterized by moderate pacing and clear articulation. The speaker's ability to sustain speech over a lengthy period, while incorporating pauses in a measured manner, suggests a high degree of fluency and effective communication skills. The speech rate, combined with the relatively high number of syllables per run, further reinforces the perception of a coherent and well-structured speech performance.

We then ran a Python code to identify variations in speech rate over time. Below, I provide the Python code used for this analysis.

```
# Load the updated files to process the full analysis
new_file1_path = '/mnt/data/MONA 1_SPEECH RATES.xlsx'
new_file2_path = '/mnt/data/MONA 2_SPEECH RATES.xlsx'
new_file3_path = '/mnt/data/MONA 3_SPEECH RATES.xlsx'

# Reading the relevant sheet with speech run data from the new files (1, 2,
and 3)
new_file1_speech_runs   =   pd.read_excel(new_file1_path,   sheet_name='speech
runs')
new_file2_speech_runs   =   pd.read_excel(new_file2_path,   sheet_name='speech
runs')
new_file3_speech_runs   =   pd.read_excel(new_file3_path,   sheet_name='speech
runs')

# Preparing a similar analysis by extracting duration and speech rate for each
file
new_file1_speech_runs['Speech Rate'] = 1 / (new_file1_speech_runs['Duration -
msec'] / 1000)
new_file2_speech_runs['Speech Rate'] = 1 / (new_file2_speech_runs['Duration -
msec'] / 1000)
new_file3_speech_runs['Speech Rate'] = 1 / (new_file3_speech_runs['Duration -
msec'] / 1000)

# Add cumulative time to all three files for comparison
new_file1_speech_runs['Cumulative  Time  (ms)']  =  new_file1_speech_runs['End
Time - msec'].cumsum()
new_file2_speech_runs['Cumulative  Time  (ms)']  =  new_file2_speech_runs['End
Time - msec'].cumsum()
new_file3_speech_runs['Cumulative  Time  (ms)']  =  new_file3_speech_runs['End
Time - msec'].cumsum()

# Now let's plot all three parts together for comparison
plt.figure(figsize=(12, 8))

plt.plot(new_file1_speech_runs['Cumulative            Time            (ms)'],
new_file1_speech_runs['Speech Rate'], label='Part 1')
plt.plot(new_file2_speech_runs['Cumulative            Time            (ms)'],
new_file2_speech_runs['Speech Rate'], label='Part 2')
plt.plot(new_file3_speech_runs['Cumulative            Time            (ms)'],
new_file3_speech_runs['Speech Rate'], label='Part 3')

plt.title('Speech Rate Over Time - Full Lecture (All Parts)')
plt.xlabel('Cumulative Time (milliseconds)')
plt.ylabel('Speech Rate (syllables per second)')
plt.legend()
plt.grid(True)
plt.show()
```

The resulting variations in speech rate are visualized in Figure 7, which presents the speech rate across the duration of the lecture.

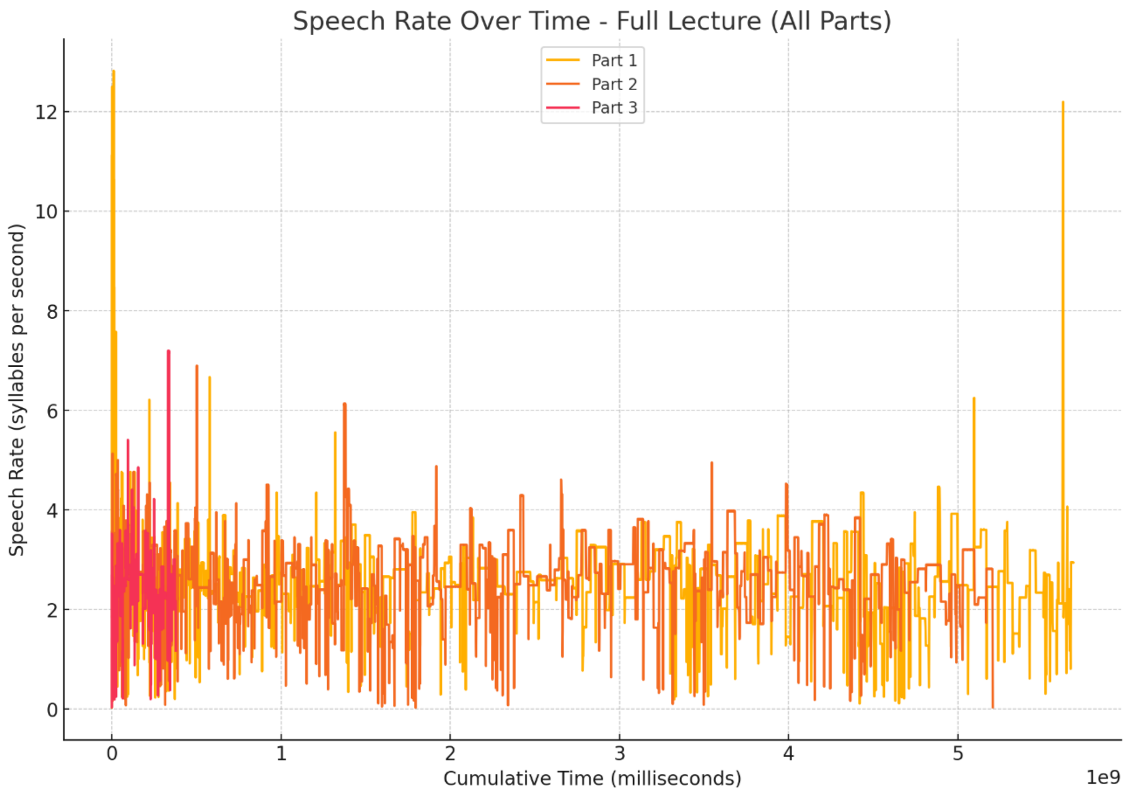

Figure 7. Speech rate over lecture time.

We analyzed an audio file obtained from the YouTube channel of the lecturer, which was originally divided into three parts. The smallest part, consisting of 9 minutes, was used for initial trial analysis. Subsequently, we continued our analysis while keeping the three parts separate. The graph shown in Figure 7 illustrates the variations in speech rate across all three parts of the lecture. Each part of the lecture shows its own trend, allowing us to observe when the lecturer speeds up or slows down. This comparison offers insight into how the speech rate evolves throughout the full session.

We now provide a more detailed breakdown by focusing on key aspects: (1) peak and trough speech rates, identifying the highest and lowest speech rates in each part of the lecture to determine when the lecturer speaks the fastest or slowest; (2) speech rate comparison, performing a statistical comparison between the parts to identify which section had the overall fastest and slowest speech rates; and (3) trends over time, observing general trends in each part to see whether the lecturer tends to speed up or slow down over time. Table 19 below presents a summary of the speech rate analysis for each part of the lecture.

Detailed Speech Rate Analysis Summary	Part 1	Part 2	Part 3
Peak Speech Rate	12.82 sps	6.90 sps	7.19 sps
Through Speech Rate	0.02 sps	0.02 sps	0.03 sps
Mean Speech Rate	2.65 sps	2.63 sps	2.47 sps
Std Dev of Speech Rate	0.78	0.69	0.77

Table *19*. Detailed speech rate (syllables per second – sps) analysis summary for the three parts of the lecture.

To facilitate the interpretation of the lecturer's variation in speech rate over time, Figure 8 visually illustrates the general trends throughout the lecture.

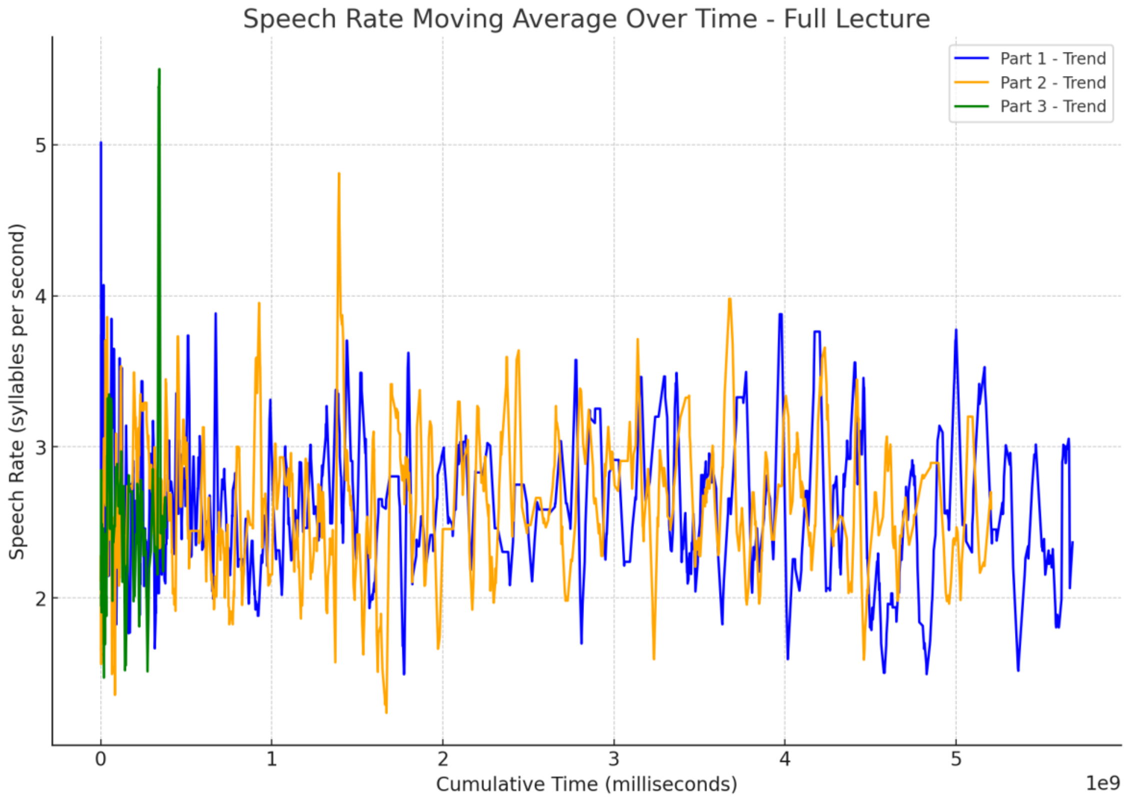

Figure *8*. Speech rate moving average over lecture time.

The plot of the moving averages across the three parts of the lecture shows general trends in speech rate throughout the lecture. The analysis of the speech rates reveals notable variations in the lecturer's delivery pace. In the first part of the lecture, we observe a higher peak speech rate of 12.82 syllables per second, indicating moments where the lecturer speaks significantly faster. However, the trough in this section, where the speech rate drops to a minimal 0.02 syllables per second, suggests periods of near silence or very slow articulation, likely due to long pauses or transitions between topics. The mean speech rate for this part, at 2.65 syllables per second, along with a standard deviation of 0.78, indicates moderate variability in the speaker's pace, with notable instances of acceleration and deceleration.

In the second part of the lecture, the peak speech rate decreases to 6.90 syllables per second, which, while still relatively fast, is considerably lower than the peak in the first part. This suggests that the speaker tends to maintain a more consistent, but slower, pace in this section. The trough speech rate, again close to 0.02 syllables per second, suggests similar pauses or breaks in speech. The mean speech rate in this part, 2.63 syllables per second, is comparable to the first part, but the lower standard deviation (0.69) implies a slightly more stable pace, with fewer fluctuations in the speed of delivery.

The third part of the lecture shows a peak speech rate of 7.19 syllables per second, slightly higher than in the second part but still notably lower than the peak in the first part. This might indicate a slight increase in urgency or focus towards the end of the lecture, where the lecturer briefly speeds up again. The trough rate is slightly higher than in the previous parts, at 0.03 syllables per second, which suggests fewer or shorter pauses. The mean speech rate in the third part is slightly lower at 2.47 syllables per second, with a standard deviation of 0.77, showing a moderate range of variation in speech pace, similar to the first part.

The moving average trends across all three parts indicate that the lecturer's speech rate fluctuates over time, with periods of increased speed followed by slower phases, possibly aligned with changes in the complexity of content or the natural rhythm of lecturing. The first part of the lecture seems to have the most pronounced variation, while the second part maintains a more consistent and steady rhythm. The third part appears to strike a balance between these two extremes, with a slight acceleration towards the end, perhaps as the lecturer approaches a conclusion or wraps up the content.

Overall, the analysis suggests that while the lecturer maintains a generally consistent pace across the entire lecture, there are distinct phases of increased speed and more deliberate, slower moments. The first part appears to have the greatest variability, with the most extreme speeds both fast and slow, whereas the second part is steadier, and the third part shows a slight increase in pace towards its conclusion.

Table 20 shows when the peaks and troughs in speech rate occur during the full lecture (Table 20).

	Peak Speech Rate (min)	**Trough Speech Rate (min)**
Part 1	1.72	0.79
Part 2	48.09	*59.95*
Part 3	83.66	*75.52*

Table *20*. Times when the peak (highest) and trough (lowest) speech rates occur during the full lecture.

These timestamps (in minutes) correspond to moments in the lecture where the lecturer speaks the fastest (peaks) and the slowest (troughs) in each part. These variations likely reflect shifts in content complexity, emphasis on key points, or natural pauses for thought and transitions. To explore these hypotheses further, the following section will analyze the pragmatic functions performed by the lecturer during different pedagogical phases of the lecture.

4.5 Pragmatic Functions

An analysis of the pragmatic functions employed by the EMI engineering lecturer reveals distinct patterns in the use of explanatory strategies, rhetorical questions, and summary statements, highlighting their prevalence and potential impact on speech rate and pedagogical effectiveness. Table 21 below shows the number of occurrences and total time spent (in seconds) on each pragmatic function in order to assess how much time the lecturer devotes to each type of function.

Pragmatic functions	occur-rences	total_duration _msec	total_duration _min
<explaining "defining">	59	2088925	34,82
<explaining "reasoning">	37	814477	13,57
<explaining "picture">	4	533546	8,89
<explaining "demonstrating">	8	344824	5,75
<housekeeping>	3	274901	4,58
<explaining "managing the channel">	5	133037	2,22
<explaining "rhetorical question">	27	114012	1,90
<summary "review previous lecture content">	5	98069	1,63
<summary "preview current lecture content">	4	90888	1,51
<explaining "story">	4	78107	1,30
<explaining "reasoning" "managing the message">	2	53440	0,89
<explaining "defining" "managing terminology">	4	43407	0,72
<explaining "reasoning" "managing audience behaviour">	1	43260	0,72
<summary "preview future lecture content">	1	41060	0,68
<explaining "defining" "code-switching">	3	31864	0,53
<explaining "Q&A">	8	25361	0,42
<summary "review current lecture content">	4	21558	0,36
<explaining "managing message">	1	18204	0,30
<explaining "reasoning" "managing terminology">	1	10660	0,18
<explaining "defining" "managing the channel">	2	8864	0,15
<explaining "code-switching">	1	7259	0,12
<explaining "defining" "managing the message">	1	5370	0,09

<explaining "translanguaging">	1	5360	0,09
<explaining "reasoning" "translanguaging">	1	4920	0,08
<explaining "defining" "managing terminology" "onomatopoeia">	1	3760	0,06
<explaining "defining" "translanguaging">	1	1792	0,03
<explaining "defining" "managing comprehension">	1	1360	0,02

Table *21*. Quantitative breakdown of the different types of pragmatic functions identified in the dataset.

The analysis of the pragmatic functions used throughout the lecture reveals several key insights into how the lecturer structures their communication. The most frequently occurring function is *explaining defining*, especially when paired with "managing terminology" or "managing comprehension". These functions, used across various parts of the lecture, show a clear emphasis on ensuring students understand key concepts and terms. The significant amount of time dedicated to these functions, particularly "explaining defining" paired with "managing terminology", highlights the lecturer's focus on providing clarity and reinforcing technical language for the students.

There are also moments when the lecturer shifts to *code-switching*, although this occurs far less frequently. The shorter time allocated to these segments suggests that while code-switching is part of the lecture strategy, it is used selectively, perhaps to bridge gaps in comprehension or provide contrast between languages.

We identified instances where the lecturer performed pragmatic functions that did not fit within the taxonomy established by Alsop (2016). These were cases in which the lecturer engaged in moves such as managing comprehension, as described by Ädel (2023), which ensures that both the lecturer and students are "on the same page" (p. 14). To account for these significant moves throughout the lecture, we applied Ädel's (2023) taxonomy to track them.

The presence of pragmatic functions like *managing comprehension* reflects the lecturer's attempt to ensure that students not only receive information but are following along and grasping the material. This reflects a teaching approach that balances delivering content with checking for understanding, using explanations as a primary tool to facilitate learning.

The distribution of these functions throughout the lecture suggests a structured pedagogical approach where more foundational explanations and definitions occur early and periodically, while reasoning and complex problem-solving may emerge later. This deliberate pacing ensures that students are equipped with the necessary vocabulary and conceptual grounding before moving on to more advanced discussions.

We then analyzed the distribution of pragmatic functions across the lecture timeline in order to observe how different functions are used at various points during the lecture.

The scatter plot in Figure 9 provides a visual representation of how the lecturer structures the session.

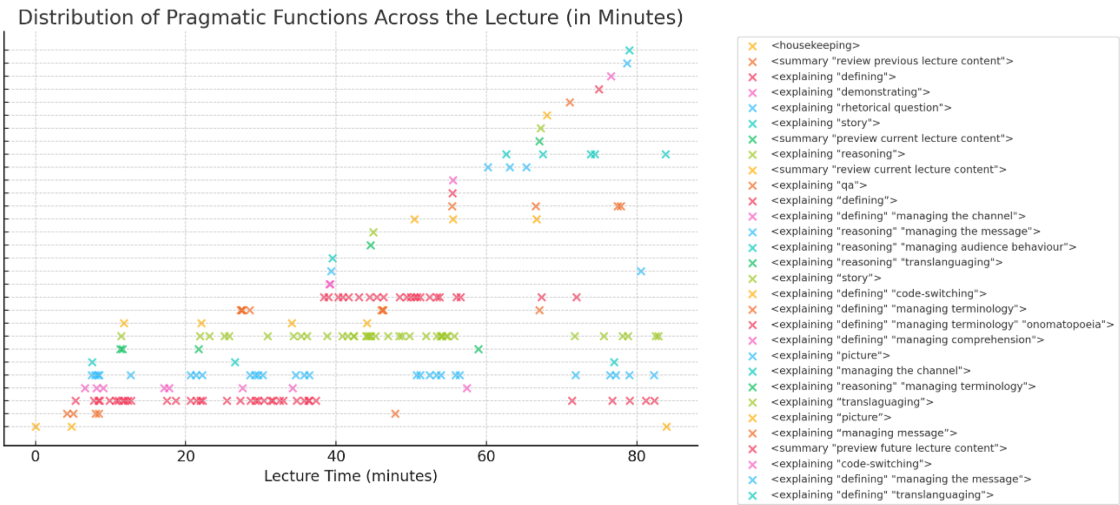

Figure 9. Distribution of pragmatic functions across the lecture timeline

The scatter plot depicting the distribution of pragmatic functions across the 84-minute lecture provides a detailed overview of how the lecturer structures his communication over time. Notably, key pragmatic functions such as "explaining defining" and "explaining reasoning" are distributed throughout the lecture, but their frequency appears more concentrated in specific segments, suggesting a deliberate pedagogical strategy.

In the initial stages of the lecture, there is a noticeable emphasis on "explaining defining", which likely reflects the lecturer's focus on introducing foundational concepts and ensuring that students have a clear understanding of technical terminology. As the lecture progresses, the frequency of this function decreases, signalling a shift from introductory definitions to more complex discussions.

Around the middle of the lecture, there is a rise in "explaining reasoning" and "rhetorical questioning", which suggests a transition to deeper engagement with the content. At this stage, the lecturer likely begins to challenge students to think critically, engage with the material more analytically, and consider the application of earlier definitions in problem-solving contexts. The appearance of rhetorical questions at these key moments may also indicate the lecturer's effort to stimulate student reflection and interaction, even in a potentially monologic setting.

Toward the latter part of the lecture, the functions become more evenly distributed, with a blend of "explaining reasoning" and "demonstrating functions". This likely reflects the lecturer's attempt to consolidate learning, offering practical applications or demonstrations of the concepts discussed earlier. The even spread of pragmatic functions in the final segment may suggest a more balanced, reflective approach as the lecturer guides students through the application and reinforcement of key ideas.

Overall, the distribution of pragmatic functions reveals a carefully structured lecture that begins with foundational definitions, transitions to critical engagement and reasoning, and concludes with demonstrations and practical applications. This structure aligns with common pedagogical approaches that build student understanding gradually before encouraging higher-order thinking and applying learned concepts in a practical context.

Finally, the peak and trough speech rates previously mentioned refer to specific pragmatic functions. Therefore, we cross-referenced the given peak and trough times with the timeline of pragmatic functions in the lecture.

We identified a first peak at 1.72 minutes and a first trough at 0.79 minutes. Both the peak and trough correspond to the "housekeeping" function, which likely involves organizational aspects of the lecture.

A second peak was found at 48.09 minutes. This peak corresponds to the "summary "review previous lecture content"" function, where the lecturer is summarizing content from a previous session. Additionally, we found a trough at 59.95 minutes, which corresponds to the "summary "preview current lecture content"" function, likely indicating a slower speech rate as the lecturer transitions to new content.

Interestingly, we identified a third peak at 83.66 minutes, which corresponds to "explaining "reasoning"", suggesting a faster speech rate when the lecturer is explaining complex reasoning or problem-solving, whereas a third trough was found at 75.52 minutes corresponding to "summary "preview future lecture content"", where the lecturer is likely summarizing or previewing future topics at a slower pace. These correspondences provide insight into how the lecturer adjusts speech rate depending on the function being performed.

4.6 Linguistic Analysis of (Dis)fluencies

This section presents the results of the linguistic analysis of DFs in the lecturer's speech. The occurrences of each DF type and their durations were counted, as shown in Table 22 below.

Types of DFs	Count	Total Duration (seconds)
unp	435	275.32
leng	140	97.106
rep	47	37.696
int	35	26.96
fp	32	20.408
combo (unp + rep)	21	25.967
combo (unp + fp)	10	14.509
combo (leng + unp)	8	10.473
combo (int + rep)	7	7.212
combo (leng + rep)	7	7.053
fs	5	5.736
combo (rep + unp)	4	7.982
combo (fp + rep)	4	6.037
combo (leng + fp)	4	5.061
combo (unp + rep + unp)	3	7.683
combo (rep + fp)	2	3.41
combo (unp +rep)	2	2.76
combo (int + fp)	2	2.71
combo (leng + unp + fp)	2	5.771
combo (int + unp)	2	2.802
combo (fp + rep + unp)	1	2.02
combo (unp + rep +unp)	1	1.87
combo (unp + fp + int)	1	1.84
combo (leng + rep + leng + unp)	1	1.76
combo (unp + fp + rep)	1	1.3
combo (int + fp + rep)	1	1.07
combo (leng +unp)	1	0.83
leng	2	1.756
combo (unp + fp)	1	0.711
combo (fp +int)	1	0.687
combo (unp + leng)	1	0.686
unp	1	0.386
combo (unp + fp + rep + unp)	1	4.229
combo (fs + int + fp)	1	3.283
combo (leng + fp + unp)	1	2.981
combo (unp + fp + unp)	1	2.205
combo (int + leng + rep)	1	1.769
combo (leng + leng)	1	1.732
combo (unp + int + rep)	1	1.694
combo (leng + unp + rep)	1	1.602
int	1	1.447
combo (int+ leng + rep)	1	1.313
combo (fp + leng)	1	1.159
combo leng + rep)	1	1.147
combo (int + leng)	1	1.132
combo (leng + int)	1	1.114
combo (fp + int)	1	1.048

Table 22. Frequency and *d*uration of DFs in *l*ecturer's *s*peech.

The data on DF types and their durations provides valuable insight into the lecturer's speech patterns and the cognitive processes involved during the lecture. The most common type of disfluency is unplanned pauses (unp), occurring 435 times with a total duration of 275.32 seconds. This suggests that the lecturer frequently pauses to gather thoughts or clarify content, which could indicate moments of reflection or hesitation, particularly when discussing complex topics or ensuring accuracy in technical explanations.

Lengthening (leng) and repetitions (rep) also appear frequently, with 140 and 47 occurrences, respectively. The fact that lengthening has a total duration of 97.106 seconds suggests that the lecturer often stretches sounds or syllables, possibly to maintain control over the discourse or signal to the audience that more information is forthcoming. Repetitions, with a duration of 37.696 seconds, indicate self-correction, where the lecturer repeats words or phrases to clarify what was just said.

Interruptions (int) and fillers (fp) occur less frequently but still play a significant role in the lecturer's speech. Interruptions, with a total duration of 26.96 seconds, may indicate external or internal disruptions to the flow of the lecture, while fillers, totaling 20.408 seconds, suggest that the lecturer uses these as temporary placeholders during moments of hesitation or thought processing.

The combination disfluencies (e.g., combo (unp + rep) or combo (unp + fp)) show that multiple types of disfluencies often occur together. This indicates moments of higher cognitive load or complexity, where the lecturer might pause, repeat, or use fillers simultaneously. These combos have varied durations, with some lasting over 25 seconds, such as combo (unp + rep), and others being shorter but still reflective of moments where the lecturer is struggling to maintain fluent speech.

These data suggests that unplanned pauses and lengthening are the most common disfluencies, reflecting the lecturer's efforts to manage cognitive processing during the lecture. Repetitions, interruptions, and fillers also play a significant role, with combination disfluencies indicating moments of heightened cognitive effort where multiple disfluency types occur together. The duration of these disfluencies provides further insight into how long the lecturer spends managing these moments, which could correlate with more challenging sections of the lecture content.

To verify this hypothesis, we examined the relationship between disfluencies (DFs) and the pragmatic functions employed during the lecture. Table 23 below illustrates the frequency of specific DF types within different pragmatic functions.

	unp	**leng**	**combo**	**rep**	**int**	**fp**	**fs**	**int**	**leng**	**unp**	**Total**
<explaining "reasoning">	114	36	21	12	11	8	1	0	0	1	**204**
<explaining "defining">	170	51	43	17	15	13	2	0	0	0	**311**
<explaining "demonstrating">	37	5	9	2	1	0	1	0	0	0	**55**
<explaining "rhetorical question">	16	10	7	2	0	1	0	0	0	0	**36**
<explaining "picture">	12	4	5	2	4	1	0	0	1	0	**29**
<explaining "story">	5	5	4	3	3	2	0	0	1	0	**23**
<explaining "managing the channel">	12	4	1	0	0	0	0	0	0	0	**17**
<explaining "Q&A">	10	3	2	0	0	2	0	0	0	0	**17**
<explaining "reasoning" "managing the message">	8	6	1	1	0	0	1	0	0	0	**17**
<explaining "defining" "managing terminology">	5	4	1	3	0	0	0	0	0	0	**13**
<summary "preview future lecture content">	7	4	1	1	0	0	0	0	0	0	**13**
<explaining "defining" "code-switching">	6	1	4	1	0	0	0	0	0	0	**12**
<summary "preview current lecture content">	8	0	0	0	0	0	0	0	0	0	**8**
<summary "review previous lecture content">	6	0	1	0	0	1	0	0	0	0	**8**
<housekeeping>	4	1	0	0	1	0	0	1	0	0	**7**
<explaining "reasoning" "managing terminology">	3	0	1	1	0	0	0	0	0	0	**5**

<summary "review current lecture content">	4	1	0	0	0	0	0	0	0	0	**5**
<explaining "reasoning" "managing audience behavior">	0	3	0	0	0	1	0	0	0	0	**4**
<explaining "code-switching">	2	0	0	0	0	1	0	0	0	0	**3**
<explaining picture>	1	1	0	0	0	1	0	0	0	0	**3**
<explaining story>	1	1	1	0	0	0	0	0	0	0	**3**
<explaining "defining" "managing the message">	1	0	0	0	0	1	0	0	0	0	**2**
<explaining managing message>	1	0	0	1	0	0	0	0	0	0	**2**
<explaining "defining" "managing the channel">	1	0	0	0	0	0	0	0	0	0	**1**
<explaining "defining" "translanguaging">	1	0	0	0	0	0	0	0	0	0	**1**
<explaining "reasoning" "translanguaging">	0	0	0	1	0	0	0	0	0	0	**1**
Total	**435**	**140**	**102**	**47**	**35**	**32**	**5**	**1**	**2**	**1**	**800**

Table 23. Patterns in the occurrence of disfluencies across various pragmatic functions during the lecture.

The results reveal interesting patterns in the occurrence of DFs across various pragmatic functions during the lecture. Notably, certain functions, such as "explaining" "defining" and "managing terminology", tend to be associated with a higher frequency of disfluencies, including repetitions and unplanned pauses. This suggests that when the lecturer is engaged in explaining or clarifying specific terms, there may be a greater cognitive load or need for precision, which could contribute to more frequent DFs.

The presence of DFs like repetitions and unplanned pauses during these defining moments may indicate moments of hesitation or self-correction, where the lecturer is trying to ensure clarity or accuracy in conveying technical content. In contrast, DFs like fillers and interruptions are less common in these sections, suggesting that the lecturer is more focused on maintaining control over the discourse while delivering key information.

In this respect, research on corpora (see e.g., Brennan & Schober, 2001, p. 275) has demonstrated that speakers tend to interrupt their speech very close to the word causing an issue, as they closely monitor their speech and stop as soon as they detect a problem (Levelt, 1989; Nooteboom, 1980;. Levelt's (1989; as cited in Brennan & Schober, 2001, p. 275) studies also revealed that when an interruption occurs within a word, it is often that specific word which poses the issue, indicating what the speaker does not intend to say. However, Blackmer and Mitton (1991; as cited in Brennan & Schober, 2001, p. 275) suggest that when interruptions are immediately followed by a repair, the issue was likely detected before the interruption. We will explore this further in the next section.

Interestingly, functions like "code-switching" and "managing the channel" exhibit fewer DFs overall. This could imply that when the lecturer shifts between languages or focuses on managing interaction within the classroom environment, the discourse is smoother, possibly because these segments are less cognitively demanding compared to explaining complex concepts.

Overall, the relationship between disfluencies and pragmatic functions provides insight into how the lecturer manages speech production, particularly during more technical or concept-heavy parts of the lecture. The higher frequency of disfluencies in these sections may reflect the increased cognitive demand and the need for accuracy in technical explanations.

Finally, we aimed at identifying the relationship between DFs and the tokens that appear during those moments. As shown in Table 24 below, the tokens that appear most frequently with disfluencies are common function words or markers of speech organization, suggesting that disfluencies tend to occur around routine or connecting words in the lecture.

Tokens	n. of occurrences
the	358
is	142
to	126
you	123
of	108
that	95
a	91
we	71

Table *24*. Tokens most frequently co-occurring with DFs in lecturer's speech.

Instead, the non-function words that most frequently co-occur with disfluencies are listed in Table 25 below.

Tokens	n. of occurrences
have	68
And	55
Flow	47
I	41
It's	40
Because	32
Sand	27
need	25

Table *25*. Content words most frequently co-occurring with DFs in lecturer's speech.

These non-function words, including verbs like "have", "need", and content-specific terms like "flow" and "sand" suggest that disfluencies occur not only with function words but also during moments when the lecturer is explaining or emphasizing content-related information.

4.7 Communicative Strategies

Once we identified DFs markers in lecturer's speech, we counted how many CSs the lecturer performs in the lecture investigated, and identified the specific types of communicative strategies used, as shown in Table 26.

CSs	n. occurrences
SIOR Acknowledgment	1
SIOR Confirmation check	1
SIOR Partial Word/Incomplete Utterance	1
SIOR Request for Confirmation	1
SISR Aborting	**15**

SISR Code-switching		6
SISR Combined Repetition		2
SISR Inserting	4	14
SISR Inserting Paraphrasing	**9**	
SISR Inserting Topic Negotiation	1	
SISR Language Search		2
SISR Parenthetical Remarks Definition	2	6
SISR Parenthetical Remarks Onomatopoeia	1	
SISR Parethetical Remarks Definition	1	
SISR Parethetical Remarks Illustration	2	
SISR Repaired Repetition	**14**	**18**
SISR Repaired Repetition "talking/discussing"	1	
SISR Repaired Repetition_Pronunciation	2	
SISR Repaired Repetition_Vocabulary	1	
SISR Repetition Parallel Phrasing	1	
SISR Replacing	**13**	**15**
SISR Replacing_Verb	1	
SISR Replacing_Tense	1	
SISR Self-Repetition	3	4
SISR Self-Repetition ("while" "wire")	1	
Total		**87**

Table 26. CSs counts.

The data on CSs used by the lecturer reveals a diverse range of strategies employed to manage speech, clarify concepts, and address various issues during the lecture. The most frequently used strategies include *SISR Aborting* (15 occurrences), *SISR Repaired Repetition* (14 occurrences), and *SISR Replacing* (13 occurrences), suggesting that the lecturer frequently needs to adjust or self-correct mid-speech. These strategies indicate moments where the lecturer revises, aborts, or replaces what was initially said, likely in response to realizing a need for greater clarity or accuracy in the information being delivered.

The relatively high occurrence of *SISR Inserting Paraphrasing* (9 occurrences) highlights the lecturer's tendency to rephrase or elaborate on information to ensure understanding, demonstrating an effort to make the content accessible to students by explaining concepts in different ways. Similarly, *SISR code-switching* (6 occurrences) and *SISR Parenthetical Remarks* (used in both definition and illustration) indicate the lecturer's use of contextual shifts or additional explanatory remarks to facilitate comprehension, particularly in instances where additional background or clarification is needed.

Less frequent but notable strategies, such as *SIOR Acknowledgment* or *SIOR Confirmation Check*, suggest that there are occasional moments where the lecturer seeks validation from students or acknowledges their engagement. These strategies help maintain interaction and ensure that the audience remains engaged with the content being delivered.

In summary, the lecturer frequently relies on self-repair mechanisms (repaired repetition, replacing, and aborting), indicating a dynamic and adaptive approach to communication during the lecture. The use of strategies like paraphrasing, parenthetical remarks, and code-switching further reflects an effort to engage students and make the material more comprehensible, especially when dealing with complex or technical content.

We then investigated how CSs align with the pragmatic functions performed during the lecture. To do so, we mapped the CSs against the pragmatic functions previously identified to see which strategies are more likely to occur during specific communicative tasks. Findings show that *SISR Replacing* was commonly associated with functions like "explaining definitions" and "demonstrating". This indicates that the lecturer frequently replaced initial explanations to improve clarity during technical content delivery. *SISR Aborting* occurred in situations where the lecturer was "managing terminology" or providing "reasoning", often indicating a need to halt an explanation and revise it. *SISR Repaired Repetition* frequently co-occurred with explaining functions, such as "defining" or "reasoning". This suggests that the lecturer often repeated key points to reinforce understanding or correct previous statements.

4.8 Gesture Analysis

As Lazaraton (2004, p. 84) pointed out, gestures

> account an "embodied sense" of how vocabulary is explained in L2 classrooms. "Sequences in which vocabulary items are explained seem a logical place to start in examining nonverbal behavior, because, as was pointed out earlier, such behavior, as a communication strategy, is thought to function as a replacement of, a support for, and/or an accompaniment to lexical items or referents in discourse."

In this section we will then focus on the gestures the lecturer performs during the lecture investigated, by "looking at the placement of gestures in multimodal ensembles" in order to grasp how gestures, which "can operate independent of other modes such as speech or be reciprocally related to them" (Bezemer, 2014, p. 361).

To start with, we proceeded analysing the total number of gestures; counting unique gestures based on the annotation "Gesture description"; then we counted each gesture type and function; next we looked at the relationship between gesture type and function, as shown in Table 27 below.

Total number of gestures	1,718
Number of unique gestures	692
Counts for each gesture type	Beats: 466 Deictic: 462 Metaphoric: 322 Iconic: 270 Self-adaptor: 189
Counts for each gesture function	Referential Concept Facilitation (Descriptive): 239 Pragmatic Parse: 186 Referential Concept Facilitation (Spatial): 94 Pragmatic Modal: 92
Relationship between gesture type and function	**Beats** are strongly associated with **pragmatic parse** (179 instances) and **pragmatic modal** (62 instances).
Relationship between gesture type and function	**Deictic gestures** often correspond to **referential concept facilitation** (both **spatial** and **descriptive**).

Table 27. Breakdown of gesture types and functions performed by the lecturer.

Conversely, Table 28 below shows the less common gesture functions found in our dataset.

Less common gesture functions	n. occurrence
Pragmatic Performative	10
Referential Concept Facilitation (Temporal)	9
Interactive	6
Referential Word Facilitation (Descriptive)	5
Referential Concept Compensation (Descriptive)	4
Referential Concept Compensation (Action)	3
Referential Concept Compensation (Spatial)	2
Referential Word Compensation (Descriptive)	2
Referential Word Facilitation (Spatial)	2
Referential Concept Facilitation (Action)	1
Referential Word Compensation (Temporal)	1
Referential Word Compensation (Action)	1
Referential Word Compensation (Spatial)	1

Table 28. Less common gesture functions found in our dataset.

Next, we aimed at identifying co-occurrences of gestures and words. Therefore, we looked at (1) the frequency of co-occurring gesture types and functions with specific words; (2) the duration of gestures and how long they overlap with specific words; and (3) patterns in co-occurrence, such as which gesture types are frequently associated with certain word categories (e.g., verbs, nouns) to assess if gestures correlate more strongly with particular types of speech.

As regards word categories and gesture types, we found that conjunctions/pronouns (e.g., "when", "it") are often associated with beats, deictic, and metaphoric gestures. They appear frequently with gesture types like beats (224 occurrences), deictic (212 occurrences), metaphoric (134 occurrences). Nouns (e.g., "floor", "particle") are most commonly associated with iconic gestures (7 occurrences) and less frequently with deictic and beats. Verbs (e.g., "reaches", "required") have lower counts but are paired with deictic and metaphoric gestures.

As regards word categories and gesture functions, Figure 10 below illustrates the top 10 most frequent gesture types across all word categories, while Figure 11 shows the top 10 most frequent gesture functions. Referential concept facilitation (descriptive) and pragmatic parse are the most common gesture functions, each occurring over 1,200 times. Pragmatic modal and referential concept facilitation (spatial) are also frequent, with a noticeable presence. Other gesture functions, such as referential concept facilitation (action) and referential word facilitation (descriptive), appear less frequently but still play significant roles.

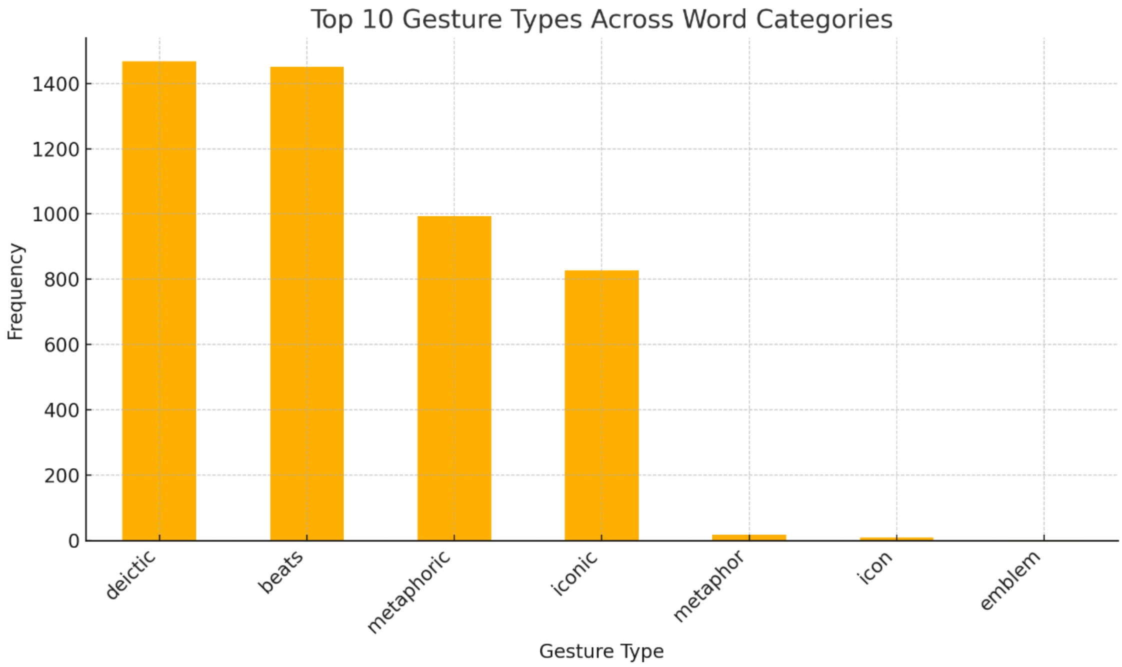

Figure 10. Top 10 most frequent gesture type across all word categories

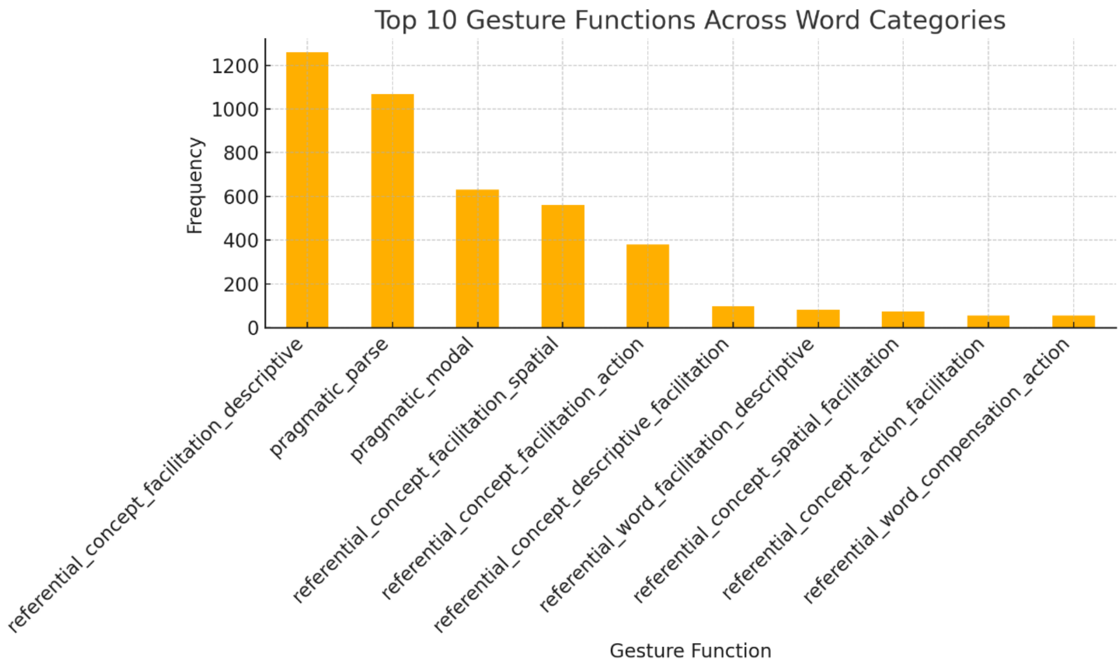

Figure 11. Top 10 most frequent gesture functions across all word categories

The data reveals significant patterns in the use of gestures during the lecture, highlighting the relationship between gesture types, functions, and word categories. The prominence of deictic and beats gestures suggests their integral role in the lecturer's communicative strategy, with each occurring over 1,400 times.

Deictic gestures, which are typically used to point or indicate objects, locations, or concepts, suggest a strong reliance on external references. The lecturer likely uses these gestures to direct attention to visual aids or spatial elements in the environment, such as the projector screen or the blackboard. This is common in educational settings, where deictic gestures help to anchor abstract concepts in the physical world. Their frequent co-occurrence with conjunctions and pronouns (e.g., "it," "when") supports the idea that these gestures often accompany references to previously mentioned or visually accessible elements in the lecture.

Beats gestures, by contrast, are rhythmic hand movements that accompany the natural flow of speech without conveying specific semantic content. Their high frequency indicates the speaker's reliance on non-verbal cues to structure the lecture, emphasize key points, and enhance the rhythmic delivery of speech. This suggests that beats gestures are likely employed to reinforce the verbal structure of the lecture, particularly around transitions and emphasis. Their association with both conjunctions/pronouns and nouns reflects the speaker's use of beats to highlight both function words (which structure sentences) and content words (which convey meaning).

The data also reveals the important role of metaphoric and iconic gestures, which occur frequently but less so than deictic and beats gestures. Metaphoric gestures – which represent abstract concepts – appear around 900 times, indicating the speaker's use of hand movements to help convey intangible or conceptual information. These gestures are critical in contexts where abstract ideas need to be communicated more concretely through visual representation. Their frequent pairing with verbs (e.g., “reaches”, “required”) suggests they are used to support descriptions of actions or processes, particularly when the content involves complex or non-tangible concepts that benefit from visual illustration.

Iconic gestures, which visually depict concrete objects or actions, appear nearly 800 times and are typically associated with nouns (e.g., “floor,” “particle”) and verbs. The use of iconic gestures reflects the speaker’s effort to visually depict physical phenomena or concrete actions, making them an essential tool in explaining processes or examples. This alignment of iconic gestures with content-rich words like nouns and verbs suggests a direct mapping between hand movements and the material being described, further aiding comprehension.

The relatively low frequency of emblems (only 1 occurrence) and icon gestures suggests that the speaker’s use of gestures is primarily illustrative and supportive rather than symbolic.

Overall, the data reflects a structured use of gestures to complement verbal communication in the lecture. Deictic and beats gestures form the foundation of this non-verbal strategy, providing spatial reference points and rhythm to the delivery. Metaphoric and iconic gestures further enhance comprehension by visually representing both abstract and concrete concepts, particularly in relation to action-oriented or descriptive language. This combination of gestures and their alignment with specific word categories highlights the speaker’s careful use of non-verbal cues to facilitate understanding, emphasize key points, and provide visual structure to complex information.

Chapter 5. Conclusions

This study explored the communicative challenges faced by a single NNES lecturer teaching engineering courses through EMI in a multilingual and multicultural academic setting. It analyzed how variations in speech rate, disfluencies, gestures, and communicative strategies intersect to shape the teaching and learning experience. Using data from three primary sources – linguistic analysis of the lecturer's speech from a video-recorded lecture, student feedback surveys, and the lecturer's self-reflections – the study investigated the lecturer's self-perceived challenges, students' reported difficulties, and the dynamics of multimodal communication during the lecture.

The findings highlighted how the lecturer handles linguistic and cognitive challenges during the delivery of technical content, as well as the strategies employed to manage these challenges. The study provides valuable insights into how speech rate, disfluencies, and gestures interact with pedagogical phases and pragmatic functions, focusing on the lecturer's communicative strategies when teaching in EMI, considering its multimodal complexity.

5.1 Key Findings and Implications

The study identified several key areas where the lecturer's communicative strategies intersect with linguistic and cognitive demands, offering insights into how these dynamics shape the teaching process.

5.1.1 Speech Rate and Cognitive Processing

Analysis revealed distinct patterns in the lecturer's speech rate, with variations aligning closely with the complexity of the content being delivered. The lecturer's faster speech rates during foundational or less complex content, such as defining basic terms or describing straightforward procedures, may reflect greater confidence in using technical terminology, likely derived from its frequent use in academic conferences and publications. However, in EMI teaching, where a diverse student body may require additional scaffolding, these faster rates could inadvertently pose challenges for comprehension. This diversity includes students who are non-experts in the field, have varying levels of language proficiency, come from different educational backgrounds, and may be accustomed to different engineering conventions or real-world applications of the concepts being taught.

Conversely, slower speech rates, often accompanied by deliberate pauses, were observed during more complex explanations, indicating an awareness of the need to provide students with additional processing time when tackling abstract or conceptually challenging material.

5.1.2 Disfluencies as Indicators of Cognitive Load

Disfluencies (e.g., unplanned pauses, repetitions, and lengthening) frequently occurred during sections involving high cognitive or linguistic demands, particularly when explaining complex technical concepts. These markers revealed moments where the lecturer appeared to engage in lexical retrieval or self-correction, highlighting the cognitive effort required to balance technical precision with linguistic clarity. For example, repetitions and pauses were often paired with pragmatic functions like defining or reasoning, signaling the lecturer's efforts to ensure accuracy and reinforce understanding. When paired with the observation that the lecturer's speech rate slows down during such explanations, this pattern suggests an awareness that these concepts or definitions require more extensive elaboration. However, it also indicates a potential lack of lexical resources to paraphrase or exemplify the material more effectively, limiting the lecturer's ability to further scaffold understanding for students.

5.1.3 Gestures as Communicative and Diagnostic Tools

Gestures were a significant feature of the lecturer's communication strategy, supporting verbal explanations and bridging gaps in comprehension. Deictic gestures directed attention to visual aids and spatial references, while metaphoric and iconic gestures helped illustrate abstract and concrete concepts, respectively. However, occasional misalignments between gestures and verbal explanations revealed moments of conceptual or linguistic difficulty. These findings underscore the dual role of gestures as both communicative tools and diagnostic indicators of underlying challenges, offering valuable insights into the lecturer's teaching strategies.

5.1.4 Misalignments between Lecturer and Student Perceptions

While the lecturer demonstrated a high level of technical expertise and confidence in his ability to teach in EMI, student feedback revealed recurring challenges related to the clarity of speech, pace of delivery, and pronunciation. These misalignments highlight the need for the lecturer to balance his own familiarity with technical terminology and academic conventions with the diverse linguistic and cognitive needs of his students, particularly in multicultural classrooms.

5.2 Contribution to EMI Research

From a methodological perspective, this research advances the understanding of multimodal communication in EMI by integrating insights from SLA and CA. By incorporating paralinguistic, linguistic, and multimodal cues, this study explores their co-occurrence to uncover patterns that reveal where communicative difficulties arise. This multimodal lens provides a deeper understanding of the nexus between these dimensions, offering valuable insights into how challenges manifest during the delivery of technical content. Through this approach, the study establishes a systematic framework for analyzing EMI instruction that bridges theoretical linguistics with practical pedagogy.

A key contribution of this research is its focus on gestures, not merely as supplementary tools but as diagnostic indicators of underlying linguistic and cognitive challenges. It emphasizes the need to treat gestures and other non-verbal resources as central to the analysis of classroom dynamics, rather than viewing them as secondary to verbal communication. The findings demonstrate that gestures, when examined alongside speech and other communicative cues, can reveal areas where the lecturer faces difficulties, such as lexical retrieval or conceptual formulation. By positioning gestures as integral to the lecturer's multimodal ensemble, the study underscores their dual role: facilitating comprehension and, at times, signaling misalignments that may impede understanding.

This evidence-based methodology, grounded in linguistic analysis, offers an empirical procedure for deconstructing lecturers' implicit beliefs about their teaching strategies. Such an approach resonates with the analytical and systematic mindset typical of engineering disciplines, making the findings more relatable and actionable for content specialists.

Moreover, the research highlights the unique demands of engineering education, where precision in terminology, abstraction of concepts, and practical application of knowledge create distinct challenges for both lecturers and students. Effective instruction in this context requires simplifying technical jargon without sacrificing accuracy, leveraging visuals to clarify complex ideas, and tailoring explanations to accommodate diverse linguistic and cultural backgrounds.

These findings underscore the need for evidence-based strategies that address the specific demands of technical disciplines. By equipping lecturers with such tools, institutions can help them navigate these complexities effectively, fostering a more inclusive and impactful learning environment while advancing the broader field of EMI research.

5.3 Pedagogical Implications

Building on these insights, this research identifies key pedagogical strategies to address the challenges of teaching technical content in multilingual and multicultural classrooms, with particular focus on the unique demands of engineering in EMI contexts. Key recommendations include:

1. *Refining speech rate and managing disfluencies.* Lecturers could be guided to adjust their speech rate according to the complexity of the material. Deliberate pauses, paraphrasing, and exemplification are effective strategies for ensuring that abstract concepts are accessible to students. EMI-specific training programs should emphasize these techniques, challenging the assumption that content expertise alone suffices for effective teaching. For instance, lecturers might rehearse their lectures in advance, practicing breaking down complex explanations into smaller, sequential steps to improve clarity and minimize disfluencies.
2. *Building strategies for lexical retrieval*. Expanding technical vocabulary and developing strategies for managing lexical retrieval challenges can enhance lecture flow. Workshops could focus on skills like using analogies, rephrasing, or integrating glossaries into teaching materials. These approaches encourage lecturers to simplify technical jargon without compromising accuracy, helping students bridge the gap between theoretical and practical knowledge.
3. *Leveraging gestures as communicative tools.* While it may be unrealistic to expect busy content specialists to dedicate significant time to gesture-specific training, reflecting on the multimodal nature of teaching can encourage greater awareness of how gestures complement verbal explanations. EMI trainers could use video

analyses to show how gestures align (or fail to align) with key points, helping lecturers understand the impact of their non-verbal communication. Such reflections should focus on treating gestures as diagnostic tools that reveal both the lecturer's challenges and the students' processing needs. This approach emphasizes gestures as integral components of teaching, rather than mere accessories.

4. *Balancing technical accuracy with linguistic accessibility*. Training programs should encourage lecturers to align their delivery style with students' linguistic proficiency levels. For instance, while advanced students may grasp technical terms quickly, others might benefit from repeated explanations, simpler phrasing, or visual aids that reinforce spoken content. Techniques like scaffolding explanations and gradually increasing complexity can ensure that all students, regardless of their language proficiency, benefit from the instruction.
5. *Developing student-centered strategies*. EMI lecturers must account for the diverse backgrounds of their students. Training programs should encourage strategies like checking for comprehension through targeted questions or allowing pauses for students to ask clarifying questions. Emphasis should also be placed on building inclusive classroom environments where students feel comfortable expressing difficulties.
6. *Collaborative approaches*. Encouraging cooperation between content specialists and language experts is vital to developing comprehensive EMI practices. Collaborative efforts could bridge the gap between technical expertise and linguistic needs, ensuring that instructional strategies are both discipline-specific and student-centered. Moreover, such collaboration could alleviate the burden felt by content specialists in addressing all these challenges alone, offering them support from language experts. This partnership not only enhances the quality of instruction but also positions EMI as an opportunity for professional

development, fostering a more integrated and supportive academic environment.

By integrating these pedagogical approaches, institutions can better equip EMI lecturers to navigate the complexities of teaching technical content in diverse classrooms. These strategies not only enhance the clarity and inclusivity of lectures but also foster a more engaging and effective learning environment.

5.4 Directions for Future Research

This study opens several avenues for future research:

1. *Expanding the sample size*. Replicating this study with a larger sample of lecturers across different engineering disciplines could provide broader insights into the communicative practices of EMI lecturers.
2. *Corpus-based analyses.* Employing corpus-based methods to analyze disfluencies and pragmatic functions could uncover recurring patterns in lecturers' communication, offering data-driven strategies for improving EMI instruction.
3. *Exploring multimodal ensembles*. Future research could further investigate the role of gestures as primary communicative tools, exploring how they reveal unarticulated linguistic or cognitive challenges in technical disciplines.

5.5 Final Reflections

This study highlights the intricate interplay of linguistic, paralinguistic, and non-verbal elements, including speech rate, disfluencies, gestures, and lexis, in shaping the communicative strategies of an experienced EMI lecturer in engineering. The findings underscore the unique challenges posed by technical disciplines, where precision, clarity, and the ability to convey abstract concepts are paramount. Gestures, far from being mere accessories to speech, emerged as meaningful tools that not only facilitated comprehension but also revealed linguistic and cognitive challenges faced by the lecturer. These non-verbal cues offer valuable insights into the lecturer's awareness of their own difficulties and strategies, opening a window into the complex dynamics of EMI instruction.

While the lecturer's strategies demonstrated adaptability and experience, the study also underscores the importance of aligning these strategies with the diverse needs of students in multicultural classrooms. By viewing EMI through a multimodal lens, this research contributes to a deeper understanding of how lecturers address the demands of teaching technical content in a second language, offering practical and theoretical implications for improving EMI practices in engineering and beyond.

Bibliography

Ackerley, K., Guarda, M., Helm, F. (2017). *Sharing Perspectives on English-Medium Instruction*. Lausanne, Switzerland: Peter Lang Verlag. https://doi.org/10.3726/978-3-0343-2538-7

Ädel, A. (2023). Adopting a 'move' rather than a 'marker' approach to metadiscourse: A taxonomy for spoken student presentations. *English for Specific Purposes*, *69*, 4–18. https://doi.org/10.1016/j.esp.2022.09.001

Aguilar, M. (2015). Engineering lecturers' views on CLIL and EMI. *International Journal of Bilingual Education and Bilingualism* 6, 722–735. http://dx.doi.org/10.1080/13670050.2015.1073664

Aguilar, M., & Rodríguez, R. (2011). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism*, *15*(2), 183–197. https://doi.org/10.1080/13670050.2011.615906

Airey, J. (2020). The content lecturer and English-medium instruction (EMI): epilogue to the special issue on EMI in higher education. *International Journal of Bilingual Education and Bilingualism*, *23*(3), 340–346. https://doi.org/10.1080/13670050.2020.1732290

Airey, J. (2012). I don't teach language. The linguistic attitudes of physics lecturers in Sweden. *AILA Review*, *25*(25), pp. 64–79. http://dx.doi.org/10.1075/aila.25.05air

Airey, J. (2011). Talking about teaching in English: Swedish university lecturers' experiences of changing teaching language. *Iberica 22*, 35–54. Retrieved from: https://revistaiberica.org/index.php/iberica/article/view/317

Airey, J., & Linder, C. (2008). Bilingual Scientific Literacy? The Use of English in Swedish University Science Courses. *Nordic Journal of English Studies*, 7, 145–161. Retrieved from: https://uu.diva-portal.org/smash/get/diva2:114031/FULLTEXT01.pdf

Airey, J., & Linder, C. (2009). A Disciplinary Discourse Perspective on University Science Learning: Achieving Fluency in A Critical Constellation of Modes. *Journal of Research in Science Teaching*, *46*(1), 27–49. Retrieved from: https://www.diva-portal.org/smash/get/diva2:114033/FULLTEXT01.pdf

Airey, J., Lauridsen, K.M., Rasanen, A., Salo, L. & Schwach, V. (2017). The expansion of English-medium instruction in the Nordic countries: Can top-down university language policies encourage bottom-up disciplinary literacy goals? *Higher Education*, *73*(4), pp. 561–576. https://doi.org/10.1007/s10734-015-9950-2

Aizawa, I. & Rose, H. (2020). High school to university transitional challenges in English Medium Instruction in Japan. *System*, *95*, 1-14. https://doi.org/10.1016/j.system.2020.102390

Alsop, S. (2016) *The pragmatic annotation of academic lectures: an exploratory study applied to engineering disciplines*. Doctoral thesis. Coventry University. Retrieved from: https://pureportal.coventry.ac.uk/en/studentTheses/the-pragmatic-annotation-of-academic-lectures

Alsop, S. & Nesi, H. (2012). The Pragmatic Annotation of a Corpus of Academic Lectures. In Mello, H. Pettorino, M & Raso, T. (Eds.) *Proceedings of the VIIth GSCP International Conference: Speech and Corpora*, pp. 1560-1563. https://doi.org/10.36253/978-88-6655-351-9

Alsop, S., & Nesi, H. (2013). The summarising function of university Engineering lectures: A cross-cultural perspective. In A. Archibald (Ed.), *Multilingual Theory and Practice in Applied Linguistics: Proceedings of the 45th Annual Meeting of the British Association for Applied Linguistics.* (pp. 11-14). Scitsiugnil Press & British Association for Applied Linguistics.

Alsop, S., Moreton, E., & Nesi, H. (2013). The uses of storytelling in university engineering lectures. In Bondi, M. & Williams, C. (Eds). *ESP Across Cultures, 10, Special issue: Academic English across cultures*, pp. 7-20. Retrieved from: https://edipuglia.it/wp-content/uploads/ESP%202013/Si-n-Alsop-Emma-Moreton-Hilary-nesi.pdf

Anderson, L., & Ciliberti, A. (2002). Monologicità e di(a)logicità nella comunicazione accademica. In C. Bazzanella (Ed.), Sul Dialogo. *Contesti e Forme di Interazione Verbale* (pp. 92–105). Guerini Studio. ISBN: 9788883353093.

Andrade, M. S. (2006). International students in English-speaking universities: Adjustment factors. *Journal of Research in International Education*, *5*(2), 131–154. https://doi.org/10.1177/1475240906065589

Arden-Close, C. (1993). Language problems in science lectures to non-native speakers. *English for Specific Purposes*, *12*(3), 251–261. https://doi.org/10.1016/0889-4906(93)90005-9

Arnó-Macià E. & Mancho-Barés G. 2015, The role of content and language in content and language integrated learning (CLIL) at university: Challenges and implications for ESP, *English for Specific Purposes*, *37*, pp. 63-73. https://doi.org/10.1016/j.esp.2014.06.007

Bagni, M. (2021). Learners' Views of EMI: Non-Native Speaker Teachers' Competence and ELF in an Italian Master's Degree Programme. In Mastellotto, L. & Zanin, R. (Eds.) *EMI and Beyond: Internationalising Higher Education Curricula in* Italy, Bozen University Press, pp. 133-163. Retrieved from: https://hdl.handle.net/11380/1227854

Ball, P., & Lindsay, D. (2013). Language demands and support for English-medium instruction in tertiary education: Learning from a specific context. In A. Doiz, D. Lasagabaster, & J. M. Sierra (Eds.), *English-medium instruction at universities: Global challenges* (pp. 44-61). Bristol: Multilingual Matters.

Basturkmen, H. & Hong, J. (2021). Development and Application of the Framework for the Analysis of Vocabulary Language-Related Episodes. In Costa, F. & Mariotti, C. (Eds.), *Input in English-Medium Instruction*, Routledge Focus on English-Medium Instruction in Higher Education, pp. 26-39. Taylor & Francis. https://doi.org/10.4324/9781003258407-3

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, *19*(2), 151–161. https://doi.org/10.1080/03075079412331382007

Beelen, J., & Jones, E. (2015). Redefining Internationalization at Home. In: Curaj, A., Matei, L., Pricopie, R., Salmi, J., Scott, P. (eds) The European Higher Education Area. Springer, Cham. https://doi.org/10.1007/978-3-319-20877-0_5

Beller-Marino, Y. (2014). *Consonantal voicing effects on vowel duration in Italian-English bilinguals*. Doctoral thesis, City University of New York, New York. Retrieved from: https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1338&context=gc_etds

Beltrán-Palanques, V. & Bernad-Mechó, E. (Eds.) (2024). *Current Trends in EMI and Multimodality in Higher Education*. Routledge Series in Language and Content Integrated Teaching & Plurilingual Education. Routledge.

Bezemer, J. (2014). The use of gesture in operations. In C. Jewitt (Ed), *Handbook of Multimodal Analysis* (second edition). London: Routledge. pp. 354-364.

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, D. & Barbieri, F. (2007). Lexical bundles in university spoken and written registers, *English for Specific Purposes*, 26(3), 263-286. https://doi.org/10.1016/j.esp.2006.08.003.

Björkman, B. (2008). *English as the lingua franca of engineering education*. Paper presented at the School of Electrical Engineering and Computer Science roundtable. Retrieved from: https://www.csc.kth.se/~kristina/utveckling2008/papers/Bjorkman.pdf

Björkman, B. (2011). Pragmatic strategies in English as an academic lingua franca: ways of achieving communicative effectiveness?. *Journal of Pragmatics 43*(4), 950-964. https://doi.org/10.1016/j.pragma.2010.07.033

Björkman, B. (2014). An analysis of polyadic English as a lingua franca (ELF) speech: A communicative strategies framework. *Journal of Pragmatics*, *66*, pp. 122–138. https://doi.org/10.1016/j.pragma.2014.03.001

Bolton, K., & Kuteeva, M. (2012). English as an academic language at a Swedish university: Parallel language use and the 'threat' of English. *Journal of Multilingual and Multicultural Development*, *33*(5), 429-447. https://doi.org/10.1080/01434632.2012.670241

Braine, G. (1989). Writing in science and technology: An analysis of assignments from ten undergraduate courses. *English for Specific Purposes*, *8*, 3–15. https://doi.org/10.1016/0889-4906(89)90003-3

Brennan, S. E., & Schober, M. F. (2001). How Listeners Compensate for Disfluencies in Spontaneous Speech. *Journal of Memory and Language*, *44*, pp. 274-296. https://doi.org/10.1006/jmla.2000.2753

Breeze, R., Roothooft, H. & Meyer, M. (2024). Teaching note-taking in EMI: an experimental study in Economics and Business Administration. *Language Learning in Higher Education*, *14*(1), 5-20. https://doi.org/10.1515/cercles-2023-0033

Broggini, S., and A. C. Murphy (2017) "Metadiscourse in EMI Lectures: Reflections on a Small Corpus of Spoken Academic Discourse", in J. Valcke, A. C. Murphy, and F. Costa (eds.) *Critical Issues in English –Medium Instruction at University, L'analisi linguistica e letteraria 25*, 325–40. Milan: Università Cattolica del Sacro Cuore. Retrieved from: https://www.analisilinguisticaeletteraria.eu/index.php/ojs/article/view/141

Campagna, S. (2016). English as a medium of instruction. A 'resentment study' of a micro EMI context. In Campagna, S., E. Ochse, V. Pulcini & M. Solly (eds.), *'Languaging' In and Across Communities: New Voices, New Identities. Studies in Honour of Giuseppina Cortese*. Bern: Peter Lang, 145–168.

Cenoz, J. (1998). *Pauses and communication strategies in second language speech* (ERIC Document No. ED426630). Educational Resources Information Center. Retrieved from: https://files.eric.ed.gov/fulltext/ED426630.pdf

Chan, S. (2015). Linguistic challenges in the mathematical register for EFL learners: linguistic and multimodal strategies to help learners tackle mathematics word problems. *International Journal of Bilingual Education and Bilingualism*, *18*(3), 306–318. https://doi.org/10.1080/13670050.2014.988114

Chang, Y.-Y. (2012). The use of questions by professors in lectures given in English: Influences of disciplinary cultures. *English for Specific Purposes*, *31*(2), 103–116. https://doi.org/10.1016/j.esp.2011.08.002

Cicillini, S. (2024). Episodes of code-switching in EMI across European contexts. In Dimova, S., Kling, J., & Drljača Margić, B. (Eds.). *EMI Classroom Communication: A Corpus-Based Approach* (1st ed.), pp. 113-144. Routledge. https://doi.org/10.4324/9781003349150-6

Clark, C. (2017). Perceptions of EMI: The students' view of a Master's degree programme. In K. Ackerley, M. Guarda, & F. Helm (Eds.), *Sharing Perspectives on English-Medium Instruction*, pp. 285–308. Bern: Peter Lang.

Costa, F. (2023). "Input Matters" Also in EMI. In Costa, F., & Mariotti, C. (Eds.), *Input in English-Medium Instruction* (Routledge Focus on English-Medium Instruction in Higher Education) (p. 1-12). Taylor & Francis. ISBN 9781003258407.

Costa, F. (2021). Introduction. EMI Stakeholders and Research in the Italian Context. Moving Towards ICLHE? In Mastellotto, L., & Zanin, R. (Eds.), *EMI and beyond: Internationalising higher education curricula in Italy*. bu,press. https://doi.org/10.13124/9788860461827

Costa, F. (2012). Focus on form in ICLHE lectures in Italy: Evidence from English-medium science lectures by native speakers of Italian. *AILA Review* 25, pp. 30–47.

Costa, F., & Mair, O. (2022). Multimodality and pronunciation in ICLHE (Integration of Content and Language in Higher Education) training. *Innovation in Language Learning and Teaching*, *16*, 4-5. https://doi.org/10.1080/17501229.2022.2053130 .

Costa F., & Mariotti C. (2017). Students' profiles and their reception of English-medium Instruction in Italian universities. In Boggio C. & Molino A. (Eds.), *English in Italy: Linguistic, Educational and Professional Challenges*, Franco Angeli, Milano, pp. 160-181. http://hdl.handle.net/10807/154734

Costa, F., & Mariotti, C. (2021). Strategies to Enhance Comprehension in EMI Lectures Examples from the Italian Context. In Lasagabaster, D & Doiz, A (Eds.), *Language use in English-Medium Instruction at university: International perspectives on teacher practice*, p. 80-99. New York, NY ; Abingdon, UK: Routledge. http://dx.doi.org/10.4324%2F9781003134534-4

Crawford Camiciottoli, B. (2008). Interaction in academic lectures vs. written text materials: the case of questions, *Journal of Pragmatics*, *40*(7), 1216-1231. http://dx.doi.org/10.1016/j.pragma.2007.08.007

Crawford Camiciottoli, B., & Fortanet-Gómez, I. (2015). Multimodal analysis in academic settings: From research to teaching. Routledge. ISBN 9781138549074

Csomay, E., & Wu, S. M. (2020). Language variation in university classrooms A corpus-driven geographical perspective. *Register Studies*, *2*(1), 131–165. https://doi.org/10.1075/rs.18002.cso

Curle, S, Jablonkai, R. R., Mittelmeier, J., Sahan, K., Veitch, A., & Galloway, N. (2020). English in higher education – English medium. Part 1: Literature review. In Galloway, N. (Ed.) *English in higher education – English medium*. British Council. Retrieved from:

https://www.teachingenglish.org.uk/sites/teacheng/files/L020_English_HE_lit_review_FINAL.pdf
Dafouz, E. & Smit, U. (2016). Towards a dynamic conceptual framework for English medium education in multilingual university settings. Applied Linguistics, 37(3), 397-415. https://doi.org/10.1093/applin/amu034
Dafouz, E. & Smit, U. (Eds.) (2023). *Researching English-Medium Higher Education. Diverse Applications and Critical Evaluations of the ROAD-MAPPING Framework*. Routledge Focus on English-Medium Instruction in Higher Education. Taylor & Francis Ltd.
Dafouz, E., Nunez B., Sancho C. & Foran D. 2007, Integrating CLIL at the tertiary level: teachers' and students' reactions. In Marsh D. and Wolff D. (Eds.), *Diverse contexts - converging goals. CLIL in Europe*, Peter Lang, Frankfurt, pp. 91-102. ISBN 9783631569054
Dafouz, E., Hüttner, J., & Smit, U. (2018). New Contexts, New Challenges for TESOL: Understanding Disciplinary Reasoning in Oral Interactions in English-Medium Instruction. *TESOL Quarterly*, *52*(3), 540-563. https://doi.org/10.1002/tesq.459
Dang T.N.Y. (2018). A Hard Science Spoken Word List, *International Journal of Applied Linguistics*, *169*(1), pp. 44-71.
Dearden, J., & Macaro, E. (2016). Higher education teachers' attitudes towards English medium instruction: A three-country comparison. *Studies in Second Language Learning and Teaching*, *6*(3), 455–486. https://doi.org/10.14746/sllt.2016.6.3.5
Deroey, K.L.B. (2023). English medium instruction lecturer training programmes: Content, delivery, ways forward, *Journal of English for Academic Purposes*, *62*, 1-16. https://doi.org/10.1016/j.jeap.2023.101223
Deroey, K.L. B., & Johnson, J. H. (2021) Metadiscourse by 'Native' and 'Non-Native' English Speakers: Importance Marking in Lectures, in *Third Metadiscourse Across Genres Conference: Metadiscourse in Intra/Inter/Cross-Cultural Communication*, 27–28 May 2021, Universitat Jaume I de Castelló, Spain. https://hdl.handle.net/10993/47297
Deroey, K., & Taverniers, M. (2011). A corpus-based study of lecture functions. Moderna Språk, 2, 1-22. Retrieved from: https://rb.gy/2g0b58
Derwing, T. M., Munro, M. J., Thomson, R. I. & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, *31*(4), 533–57. https://doi.org/10.1017/s0272263109990015
Doiz, A., Costa, F., Lasagabaster, D., & Mariotti, C. (2019). Linguistic demands and language assistance in EMI courses. What is the stance of Italian and Spanish

undergraduates? *Lingue e Linguaggi*, *33*, 69-85. https://doi.org/10.1285/I22390359V33P69

Dubow, G., Gundermann, S. & Northover, L. (2021). Establishing quality criteria and an EMI certification procedure. In Pun, Jack K.H. & Curle, S. M. (Eds.) *Research Methods in English Medium Instruction*, Routledge Research in Higher Education, (p. 75-91). Taylor & Francis. ISBN 9781003025115

Dalton-Puffer, C. (2007). *Discourse in Content-and-Language-Integrated Learning (CLIL) Classrooms.* Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.20

Efron, D. (1972). Gesture, race and culture; a tentative study of the spatio-temporal and "linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions. The Hague, Mouton.

Eklund, R. (2004). *Disfluency in Swedish human-human and human-machine travel booking dialogues*. Department of Computer and Information Science, Linköping Studies in Science and Technology. Retrieved from: http://www.diva-portal.org/smash/get/diva2:20923/FULLTEXT01.pdf

Ekman, P. and Friesen, W.V. (1969) The Repertoire or Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica*, *1*, 49-98. https://doi.org/10.1515/semi.1969.1.1.49

Emrani, F., & Hooshmand, M. (2019). A conversation analysis of self-initiated self-repair structures in advanced Iranian EFL learners. International Journal of Language Studies, (13)1, pp. 57-76. Retrieved from: https://files.eric.ed.gov/fulltext/ED591531.pdf

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

Feyereisen, P. (2006). How could gesture facilitate lexical access?, *Advances in Speech Language Pathology*, *8*(2), 128-133, https://doi.org/10.1080/14417040600667293

Fortanet-Gómez, I., & Ruiz-Madrid, N. (2014). Multimodality for comprehensive communication in the classroom: Questions in guest lectures. Ibérica, 28, 203–224. Retrieved from: https://www.revistaiberica.org/index.php/iberica/article/view/245

Francomacaro, R. M. (2011). English as a medium of instruction at an Italian Engineering faculty: An investigation of structural features and pragmatic functions. Doctoral thesis. Università degli Studi di Napoli Federico II, Italy. Retrieved from: http://www.fedoa.unina.it/8637/1/francomacaro_mariarosaria_24.pdf

Graziano, M., & Gullberg, M. (2018). When Speech Stops, Gesture Stops: Evidence From Developmental and Crosslinguistic Comparisons. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00879

Gronchi, M. (2023). Language assessment in EMI: unravelling the implicit-explicit dichotomy. *Educational Linguistics*. https://doi.org/10.1515/eduling-2023-0011

Guarda, M. & F. Helm (2016). 'I have discovered new teaching pathways': The link between language shift and teaching practice. *International Journal of Bilingual Education and Bilingualism*, 117. http://dx.doi.org/10.1080/13670050.2015.1125848

Gullberg, M. (2006a). Handling Discourse: Gestures, Reference Tracking, and Communication Strategies in Early L2. *Language Learning*, *56*(1), 155–196. https://doi.org/10.1111/j.0023-8333.2006.00344.x

Gullberg, M. (2006b). Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *International Review of Applied Linguistics*, *44*(2), 103-124. https://doi.org/10.1515/IRAL.2006.004

Hafner, C., & Miller, L. (2018). *English in the Disciplines: A Multidimensional Model for ESP Course Design* (1st ed.). Routledge. https://doi.org/10.4324/9780429452437

Harley, B. (2000). Listening Strategies in ESL: Do Age and L1 Make a Difference? *TESOL Quarterly*, *34*(4), 769. https://doi.org/10.2307/3587790

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

Helm, F. & Guarda, M. (2015). "Improvisation is not allowed in a second language": A survey of Italian lecturers' concerns about teaching their subjects through English. *Language Learning in Higher Education*, *5*(2), 353-373. https://doi.org/10.1515/cercles-2015-0017

Hincks, R. (2010). Speaking rate and information content in English lingua franca oral presentations. *English for Specific Purposes*, *29*(1), 4-18. https://doi.org/10.1016/j.esp.2009.05.004

Holler, J., Turner, K., & Varcianna, T. (2013). It's on the tip of my fingers: Co-speech gestures during lexical retrieval in different social contexts. *Language and Cognitive Processes*, *28*(10), 1509–1518. https://doi.org/10.1080/01690965.2012.698289

Hua, T-L. (2019). Understanding the Learning Challenges of English-Medium Instruction Learners and Ways to Facilitate Their Learning: A Case Study of Taiwan Psychology Students' Perspectives. *Latin American Journal of Content & Language Integrated Learning*, *12*(2), 321-340. https://doi.org/10.5294/laclil.2019.12.2.6

Huang, W., & Fang, F. (2022). EMI Teachers' perceptions and practices regarding culture teaching in Chinese higher education. *Language, Culture and Curriculum*, *36*(2), 205–221. https://doi.org/10.1080/07908318.2022.2115056

Huth, T. (2011). Conversation Analysis and Language Classroom *Discourse. Language and Linguistics Compass,* (5)5, 297–309. https://doi.org/10.1111/j.1749-818X.2011.00277.x

Hyland, K. & Tse , P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly 41*(2), 235–54. http://www.jstor.org/stable/40264352

Jablonkai, R.R. (2021). Corpus linguistic methods in EMI research. A missed opportunity? In Pun, J.K.H., & Curle, S.M., *Research Methods in English Medium Instruction* (Routledge Research in Higher Education) (p. 92-106). Taylor & Francis. https://doi.org/10.4324/9781003025115-7

Jefferson G. (2004). Glossary of transcript symbols with an introduction. In Lerner G. H. (Ed.), *Conversation Analysis: Studies from the First Generation*, John Benjamins, Amsterdam, pp.13-31. https://doi.org/10.1075/pbns.125.02jef

Jenkins, J. (2009). English as a lingua franca: interpretations and attitudes. *World Englishes, 28*, 200-207. https://doi.org/10.1111/j.1467-971X.2009.01582.x

Jensen, C. (2013). Students' Attitudes to Lecturers' English in English Medium Higher Education in Denmark. *Nordic Journal of English Studies*, *13*(1), 87-112. Retrieved from: https://ojs.ub.gu.se/index.php/njes/article/view/1798/1572.

Johnson, J. H., & Picciuolo, M. (2023). Lecturer language: EMI students' experiences on first- and second-cycle degrees. In *Proceedings of the Ninth International Conference on Higher Education Advances (HEAd'23)*, June 19-23, Valencia, Spain. https://doi.org/10.4995/HEAd23.2023.16161

Johnson, J. H., & Picciuolo, M. (2022). *The EmiBO corpus. EMI lecturer discourse across disciplines and lecture modes* (Versione 1.0) [Dataset]. University of Salento. https://doi.org/10.1285/I22390359V53P253

Johnson J.H. & Picciuolo M. (2020). Interaction in spoken academic discourse in an EMI context: the use of questions, in *Proceedings of 6th International Conference on Higher Education Advances* (HEAd'20), June 2-5, 2020, Universitat Politècnica de València, València, pp. 211-219. https://doi.org/10.4995/HEAd20.2020.11018

Kaur, P. (2014). Attitudes towards English as a Lingua Franca. *Procedia - Social and Behavioral Sciences*, *118*, 214–221. https://doi.org/10.1016/j.sbspro.2014.02.029

Kaur, J. (2012). Saying it again: Enhancing clarity in English as a lingua franca (ELF) talk through self-repetition. *Text & Talk*, *32*(5). https://doi.org/10.1515/text-2012-0028

Kaur, J. (2011). Raising explicitness through self-repair in English as a lingua franca. *Journal of Pragmatics*, *43*(11), 2704–2715. https://doi.org/10.1016/j.pragma.2011.04.012

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica*, *62*, 3–28. https://doi.org/10.1515/semi.1986.62.1-2.3

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Kay (Ed.), *The Relationship between Verbal and Nonverbal Behavior*. The Hague: Mouton. Retrieved from: https://rb.gy/r6f3hk

Kim, E.G. & Yoon J.-R. (2018). Korean Science and Engineering Students' Perceptions of English-Medium Instruction and Korean-Medium Instruction. *Journal of Language Identity & Education*, *17* (3), 182-197. https://doi.org/10.1080/15348458.2018.1433539

Kim, J., Tatar B and Choi J. (2014). Emerging culture of English-medium instruction in Korea: Experiences of Korean and international students. *Language and Intercultural Communication, 14*(4), 441-459. http://dx.doi.org/10.1080/14708477.2014.946038

Kırkgöz, Y. (2009). Students' and lecturers' perceptions of the effectiveness of foreign language instruction in an English-medium university in Turkey. *Teaching in Higher Education*, *14*(1), 81–93. https://doi.org/10.1080/13562510802602640

Klaassen, R. (2008). Preparing lecturers for English-medium instruction. *Realizing content and language integration in higher education*, 32-42. Retrieved from: https://rb.gy/ller62

Klaassen, R. G. (2001). The International University Curriculum. Challenges in English-Medium Engineering Instruction. Doctoral thesis. Retrieved from: https://repository.tudelft.nl/record/uuid:dea78484-b8c2-40d0-9677-6a508878e3d9

Klaassen, R. & Räsänen, A. (2006). Assessment and staff development in higher education for English-medium instruction: a question-raising article. In: Wilkinson, R., Zegers, V., & van Leeuwen, C. (Eds.). *Bridging the assessment gap in English-medium higher education*. Bochum: AKS-Verlag, pp. 235-255.

Kornder, L., & Mennen, I. (2021). Listeners' linguistic experience affects the degree of perceived nativeness of first language pronunciation. *Frontiers in Psychology*, *12*, 717615. https://doi.org/10.3389/fpsyg.2021.717615

Kosmala, L. (2021). On the specificities of L1 and L2 (dis)fluencies and the interactional multimodal strategies of L2 speakers in tandem interactions. *Journal of Monolingual and Bilingual Speech*, *3*(1). https://doi.org/10.1558/jmbs.15676

Kunioshi, N., Noguchi, J., Tojo, K., & Hayashi, H. (2016). Supporting English-medium pedagogy through an online corpus of science and engineering lectures. *European Journal of Engineering Education*, *41*(3), 293-303. https://doi.org/10.1080/03043797.2015.1056104

Kunioshi, N., Noguchi, J., Tojo, K., & Hayashi, H. (2014). Identifying Pedagogical Functions in University Science and Engineering Lectures. Paper presented at the *New Perspectives in Science Education* Pixel International Conference, 3rd edition, Florence, 20-21 March, 2014. Retrieved from: https://conference.pixel-online.net/library_scheda.php?id_abs=162

Kuteeva, M., & Airey, J. (2014). Disciplinary Differences in the Use of English in Higher Education: Reflections on Recent Policy Developments, *Higher Education*, *67*(5), 533-549. https://doi.org/10.1007/s10734-013-9660-6

Lasagabaster, D. (2018). Fostering team teaching: Mapping out a research agenda for English-medium instruction at university level. Language Teaching, 51(3), 400-416. https://doi.org/10.1017/S0261444818000113

Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning* 54: 79–117.

Leask, B. (2009). Using Formal and Informal Curricula to Improve Interactions between Home and International Students. *Journal of Studies in International Education*, *13*, 205-221. https://doi.org/10.1177/1028315308329786

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Ablex Pub. Corp. Retrieved from: https://files.eric.ed.gov/fulltext/ED362379.pdf

Levelt, W. (1989). *Speaking*. MIT Press, Cambridge, Massachusetts. https://doi.org/10.7551/mitpress/6393.001.0001

Lickley, R. J. (2015). Fluency and disfluency. In M. A. Redford (ed.), *The handbook of speech production* (pp. 445–74). Hoboken, NJ: John Wiley. https://doi.org/10.1002/9781118584156.ch20

Lim, F. V. (2024). The multimodal turn in higher education. In Beltrán-Palanques, V. & Bernad-Mechó, E. (Eds.), *Current Trends in EMI and Multimodality in Higher Education*. Routledge Series in Language and Content Integrated Teaching & Plurilingual Education, pp. 9-26. Routledge. https://doi.org/10.4324/9781003205517-2

Lim, F. V. (2021). Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, *20*(1), 34–58. https://doi.org/10.1177/1470357218820695

Lim, F. V., O'Halloran, K. L., & Podlasov, A. (2012). Spatial pedagogy: mapping meanings in the use of classroom space. *Cambridge Journal of Education*, *42*(2), 235–251. https://doi.org/10.1080/0305764X.2012.676629

Lin, A. M. Y. (2016). Language across the curriculum and CLIL in English as an additional language (EAL) context: Theory and practice. Springer.

Liu, J., & Nesi, H. (1999). Are we teaching the right words? A study of students' receptive knowledge of two types of vocabulary: "Subtechnical" and "technical." In H. Bool & P. Luford (Eds.), *Academic standards and expectations: The role of EAP* (pp. 141-147). Nottingham, England: Nottingham University Press.

Lo, Y. Y., & Macaro, E. (2012). The medium of instruction and classroom interaction: Evidence from Hong Kong secondary schools. *International Journal of Bilingual Education and Bilingualism*, *15*(1), 29–52. https://doi.org/10.1080/13670050.2011.588307

Lopez-Ozieblo, R. (2022). Cut-offs and co-occurring gestures: Similarities between speakers' first and second languages. *International Review of Applied Linguistics in Language Teaching*, *60*(3), 647-677. https://doi.org/10.1515/iral-2017-0117

Macaro, E. (2020). Exploring the role of language in English medium instruction. *International Journal of Bilingual Education and Bilingualism, 23*(3), 266–276. https://doi.org/10.1080/13670050.2019.1620678

Macaro, E. (2019). Exploring the role of language in English medium instruction. *International Journal of Bilingual Education and Bilingualism*, *23*(3), 263–276. https://doi.org/10.1080/13670050.2019.1620678

Macaro, E, Akincioglu, M & Dearden, J. (2016). English-medium instruction in universities: A collaborative experiment in Turkey. *Studies in English Language Teaching, 4*(1), 51–76. http://dx.doi.org/10.22158/selt.v4n1p51

Mair, O. (2021). EMI Professional Development in Italy: An Assessment Focus. In Mastellotto, L., Zanin, R. (Eds.) *EMI and Beyond: Internationalising Higher Education Curricula in Italy*, pp. 109-132. Bu,press, Bozen-Bolzano University Press. Retrieved from: https://pro.unibz.it/library/bupress/publications/fulltext/9788860461827.pdf

Malmström, H., & Pecorari, D. (2021). Epilogue: Disciplinary literacies as a Nexus for content and language teacher practice. In Lasagabaster, D & Doiz, A (Eds.), *Language use in English-Medium Instruction at university: International perspectives on teacher practice*, p. 213-220. New York, NY; Abingdon, UK: Routledge. http://dx.doi.org/10.4324%2F9781003134534-4

Mariotti, C. (2021). Features of Spoken Input Across the Disciplines and at the Interface Between English-Medium Instruction and Disciplinary Discourse. In Costa, F., & Mariotti, C. (Eds.), *Input in English-Medium Instruction* (Routledge Focus on English-Medium Instruction in Higher Education) (p. 1-12). Taylor & Francis. http://dx.doi.org/10.4324/9781003258407-4

Martinez, R., Machado, P., & Palma, C. (2021). An Exploratory Analysis of Language Related Episodes (LREs) in a Brazilian EMI Context Lecturers' and Students' Perspectives. In Lasagabaster, D & Doiz, A (Eds.), *Language use in English-Medium Instruction at university: International perspectives on teacher practice*, p. 11-33. New York, NY; Abingdon, UK: Routledge. http://dx.doi.org/10.4324%2F9781003134534-4

Martin-Rubió, X., & Diert-Boté, I. (2023). (Dis)fluency and pronunciation accuracy in EMI lectures. In S. Dimova, J. Kling, & B. Drljača Margić, *EMI Classroom Communication* (1ª ed., pp. 36–60). Routledge. https://doi.org/10.4324/9781003349150-3

Matsuura, H., Chiba, R., Mahoney, S., & Rilling, S. (2014). Accent and speech rate effects in English as a lingua franca. *System*, *46*, 143–150. https://doi.org/10.1016/j.system.2014.07.015

Mauranen, A. (2012). *Exploring ELF: Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press, pp. x + 271.

Mauranen, A. (2010a). Discourse Reflexivity - a Discourse Universal? The case of ELF. *Nordic journal of English studies*, *9*(2), 13-40. http://ojs.ub.gu.se/ojs/index.php/njes/article/view/413/401

Mauranen, A. (2010b). Features of English as a lingua franca in academia. *Helsinki English Studies*, *6*, pp. 6–28. Retrieved from: https://blogs.helsinki.fi/hes-eng/files/2010/12/Mauranen_HES_Vol6.pdf

Mauranen, A. (2007). Hybrid voices: English as the lingua franca of academics. In: Flottum, K. (Ed.), *Language and Discipline Perspectives on Academic Discourse*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 243–259.

Maynard, C. & Leicher, S. (2007). Pragmatic Annotation of an Academic Spoken Corpus for Pedagogical Purposes. in *Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse*. ed. by Fitzpatrick, E. Amsterdam: Rodophi, 107-116.

McDonough, J. (2010) English for specific purposes: A survey review of current materials. *ELT Journal* 64, 462 –77. https://doi.org/10.1093/elt/ccq060

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

McNeill, D. (1992). *Hand and mind: What the hands reveal about thought*. Chicago: Chicago University Press.

McNeill, D. (1985). *So you think gestures are nonverbal? Psychol. Rev. 92*, 271–295. Retrieved from: http://www.communicationcache.com/uploads/1/0/8/8/10887248/so_you_think_gestures_are_nonverbal.pdf

McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, *45*(1), 53–78. https://doi.org/10.1002/tea.20201

Miller, L. (2009) Engineering lectures in a second language: What factors facilitate students' listening comprehension? *Asian EFL Journal*, 2(11), 8-30. Retrieved from: https://rb.gy/7csy5u

Molino, A. (2023). Mediation strategies in EMI. Facilitating access to knowledge through language. In Dimova, S., Kling, J., & Drljača Margić, B. (Eds.). *EMI Classroom Communication: A Corpus-Based Approach* (1st ed.), pp. 61-90. Routledge. https://doi.org/10.4324/9781003349150-4

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 & Practice*, *18*(4), 935–56. https://doi.org/10.12738/estp.2018.4.0330.

Moncada-Comas, B., & Sabaté-Dalmau, M. (2024). The role of multimodal competence in 'doing EMI lecturing': Exploring the effectiveness of non-linguistic resources in EMI. In Beltrán-Palanques, V. & Bernad-Mechó, E. (Eds.) (2024). *Current Trends in EMI and Multimodality in Higher Education*. Routledge Series in Language and Content Integrated Teaching & Plurilingual Education (pp. 96-129). Routledge. https://doi.org/10.4324/9781003205517-7

Moniz, H., Batista, F, Mata, AI, Trancoso, I. (2014). Speaking style effects in the production of disfluencies. *Speech communication* *65*, 20–35. https://doi.org/10.1016/j.specom.2014.05.004

Morell, T. (2020). EMI teacher training with a multimodal and interactive approach: A new horizon for LSP specialists. *Language Value*, *12*(1), 56–87. http://dx.doi.org/10.6035/languagev.2020.12.4/

Morell, T. (2018). Multimodal competence and effective interactive lecturing. *System*, *77*, 70–79. https://doi.org/10.1016/j.system.2017.12.006

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*(2), 235–56. https://doi.org/10.1016/j.esp.2005.05.002

Myers, G. (2003). Words, pictures, and facts in academic discourse. *Iberica* *6*, 3–13. Retrieved from https://revistaiberica.org/index.php/iberica/article/view/482

Nesi, H., & Alsop, S. (2021). Stories and scenarios: Lecturers' use of fantastic hypothetical events. *Journal of English for Academic Purposes*, *53*, 101022. https://doi.org/10.1016/j.jeap.2021.101022

Norris, S. (2004). Analyzing multimodal interaction: A methodological framework. Routledge. ISBN 9780415328562

Norte, N. & Morell, T. (2024). English-medium instruction (EMI) multimodal classroom discourse to enhance L2 university students' comprehension. In Beltrán-Palanques, V. & Bernad-Mechó, E. (Eds.) (2024). *Current Trends in EMI and Multimodality in Higher Education*. Routledge Series in Language and Content Integrated Teaching & Plurilingual Education (pp. 29-47). Routledge. https://doi.org/10.4324/9781003205517-4

O'Shaughnessy, D. (1994). Correcting complex false starts in spontaneous speech. *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pp. 349-352. Adelaide, S. Australia. IEEE. Retrieved from: https://files.eric.ed.gov/fulltext/ED356506.pdf

Ollinger, A. (2012). *The good ELF user: A qualitative meta-analysis of strategic language use behaviours in English as a lingua franca.* (Master's thesis). Universität Wien, Austria. Retrieved from: https://rb.gy/ld4ofh

Parkinson, J. (2012). English for Science and Technology. In B. Paltridge and S. Starfield (Eds.), *The Handbook of English for Specific Purposes*, pp. 155-174. https://doi.org/10.1002/9781118339855.ch8

Peltonen, P. (2019). Gestures as fluency-enhancing resources in L2 interaction: A case study on multimodal fluency. *Fluency in L2 Learning and Use*, 138, 111. https://doi.org/ 10.21832/9781788926317-010

Picavet, F., Henderson, A., & De Vries, E. (2023). Storytelling in L2 English-medium engineering lectures: A typology. *Journal of English for Academic Purposes*, *62*, 101214. https://doi.org/10.1016/j.jeap.2022.101214

Picciuolo, M. (2023a). An ELF-oriented corpus-based analysis into the EMI lecturers' use of spatial deixis across two different teaching media. *ELOPE: English Language Overseas Perspectives and Enquiries*, *20*(1), 89-112. https://doi.org/10.4312/elope.20.1.89-112

Picciuolo, M. (2023b). Reconceptualising space in academic lectures: Face-to-face, blended, and online lecturer discourse in the context of English-medium instruction. *ESP Today*, *11*(2), 371–394. https://doi.org/10.18485/esptoday.2023.11.2.9

Picciuolo, M. & Johnson, J.H. (2020). Contrasting EMI lecturers' perceptions with practices at the University of Bologna, in D. R. Miller (Ed.) *Quaderni del CeSLiC. Occasional papers* AlmaDL, Centro di Studi Linguistico-Culturali (CeSLiC) e Alma Mater Studiorum, Università di Bologna: Bologna, 1–23. https://doi.org/10.6092/unibo%2Famsacta%2F6399

Poe, M., Lerner, N., & Craig, J. (2010). Learning to communicate in science and engineering: Case studies from MIT. MIT Press.

Pritchard, R. M. O. & Nasr, A. (2004). Improving reading performance among Egyptian engineering students: Principles and practices. *English for Specific Purposes*, *23*, 425–45. https://doi.org/10.1016/j.esp.2004.01.002

Pueyo, I. G. & Val, S. (1996). The construction of technicality in the field of plastics: A functional approach towards teaching technical terminology. *English for Specific Purposes 15*(4), 251–78. https://doi.org/10.1016/S0889-4906(96)00011-7

Pulcini, V. & S. Campagna (2015). Internationalisation and the EMI controversy in Italian higher education, In S. Dimova, A. K. Hultgren & C. Jensen (eds.), *English-medium instruction in European higher education*. Boston/Berlin: De Gruyter Mouton, 65–87.

Pun, J., & Macaro, E. (2019). The effect of first and second language use on question types in English medium instruction science classrooms in Hong Kong. *International Journal of Bilingual Education and Bilingualism*, *22*(1), 64–77. https://doi.org/10.1080/13670050.2018.1510368

Querol-Julián, M. (2021). How does digital context influence interaction in large live online lectures? The case of English-medium instruction. *European Journal of English Studies*, *25*(3), 297–315. https://doi.org/10.1080/13825577.2021.1988265

Querol-Julián, M. & Amondarain-Garrido, M. (2024). Listening to silence in Spanish and English-medium instruction online lectures. In Beltrán-Palanques, V. & Bernad-Mechó, E. (Eds.) (2024). *Current Trends in EMI and Multimodality in Higher Education*. Routledge Series in Language and Content Integrated Teaching & Plurilingual Education (pp. 71-95). Routledge. https://doi.org/10.4324/9781003205517-6

Roesler, D. (2021). When a bug is not a bug: An introduction to the computer science academic vocabulary list, *Journal of English for Academic Purposes*, *54*, 1-11. https://doi.org/10.1016/j.jeap.2021.101044.

Roth, W.-M. (2001). Gestures: Their Role in Teaching and Learning. *Review of Educational Research*, *71*(3), 365–392. https://doi.org/10.3102/00346543071003365

Roth, W.-M., & Lawless, D. V. (2002). When up is down and down is up: Body orientation, proximity, and gestures as resources. *Language in Society*, *31*(1), 1–28. https://doi.org/10.1017/S004740450200101X

Ruiz Madrid, N., & Valeiras-Jurado, J. (2020). Developing multimodal communicative competence in emerging academic and professional genres. *International Journal of English Studies*, *20*(1), 27–50. https://doi.org/10.6018/ijes.401481

Sahan, K. (2020). ELF interactions in English-medium engineering classrooms. *ELT Journal*, *74* (4), 418-427. https://doi.org/10.1093/elt/ccaa033

Schegloff, E. A. (1992). On talk and its institutional occasions. In P. Drew and J. Heritage, *Talk at work*, (pp. 101–34). Cambridge: Cambridge University Press. Retrieved from: https://escholarship.org/content/qt4365f03g/qt4365f03g.pdf?t=le41dy

Schegloff, E. A. (1997). Whose Text? Whose Context? *Discourse & Society*, *8*(2), 165-187. https://doi.org/10.1177/0957926597008002002

Schegloff, E. A. (2000). When ‘others’ initiate repair. *Applied Linguistics*, *21*(2), 205-243. https://doi.org/10.1093/applin/21.2.205

Schegloff, E. A., Jefferson, G., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language 53*(2), 361–82. Retrieved from: https://icar.cnrs.fr/ecole_thematique/tranal_i/documents/org_seq/scheSacksJeff77_repair.pdf

Secundo, G., Magnier-Watanabe, R., & Heisig, P. (2015). Engineering knowledge and information needs in Italy and Japan: Bridging the gap between theory and practice. *Journal of Knowledge Management*, *19*(6), 1310–1334. https://doi.org/10.1108/JKM-01-2015-0029

Seedhouse, P. (2004). The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective. *Language Learning, 54*(Suppl1), x–300. https://doi.org/10.1111/j.1467-9922.2004.00266.x

Seyfeddinipur, M. (2006). *Disfluency: Interrupting Speech and Gesture* .. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from: https://repository.ubn.ru.nl/handle/2066/41397

Shriberg, E. E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Doctoral thesis, University of California, Berkeley. Retrieved from: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c0ca94051f549f08e0bb4be7694540460fd47f1b

Simpson-Vlach, R. & Leicher, S. (2006). *The MICASE Handbook: A Resource for Users of the Michigan Corpus of Academic Spoken English*. Ann Arbor: University of Michigan Press. ISBN 10: 0472030728

Sinclair, J., & Coulthard, M. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press.

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2010). The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Human Brain Mapping*, *32*(4), 520–533. https://doi.org/10.1002/hbm.21041

Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning* (Vol. 2). John Benjamins Publishing Company. https://doi.org/10.1075/gs.2

Suviniitty, J. (2012). *Lectures in English as a Lingua Franca: Interactional Features*. Doctoral thesis. University of Helsinki. Retrieved from: https://helda.helsinki.fi/bitstreams/bfa1210e-a23c-493f-a468-b28080067492/download

Swales, J. M. (1995). The role of the textbook in EAP writing research. *English for Specific Purposes*, *14*(1), 3–18. https://doi.org/10.1016/0889-4906(94)00028-C

Tange, H. (2010). Caught in the Tower of Babel: University lecturers' experiences with internationalisation. *Language and Intercultural Communication*, *10*(2), 137–149. https://doi.org/10.1080/14708470903342138

Tatzl, D. (2011). English-medium masters' programmes at an Austrian university of applied sciences: Attitudes, experiences and challenges. *Journal of English for Academic Purposes*, *10* (4), 252–270. https://doi.org/10.1016/j.jeap.2011.08.003

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, *65*(1), 71–79. https://doi.org/10.1093/elt/ccq020

Thai, K. W.H. (2021). Researching translanguaging in EMI classrooms. In Pun, J.K.H., & Curle, S.M. (Eds.) *Research Methods in English Medium Instruction* (Routledge Research in Higher Education), pp. 119-132. Taylor & Francis. https://doi.org/10.4324/9781003025115-9

Thøgersen, J., & Airey, J. (2011). Lecturing undergraduate science in Danish and in English: A comparison of speaking rate and rhetorical style. *English for Specific Purposes*, *30*(3), 209–221. https://doi.org/10.1016/j.esp.2011.01.002

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1–27). Oxford University Press.

Uchihara, T., & Harada, T. (2018). Roles of Vocabulary Knowledge for Success in EnglishMedium Instruction: Self-Perceptions and Academic Outcomes of Japanese Undergraduates. *TESOL Quarterly*, *52*(3), 564–587. https://doi.org/http://dx.doi.org/10.1002/tesq.453

Valcke, J., & Pavón, V. (2015). A Comparative Study on the Use of Pronunciation Strategies for Highlighting Information in University Lectures. In R. Wilkinson, and M. L. Walsh (Eds.), *Integrating Content and Language in Higher Education from Theory to Practice*, pp. 323–341. Frankfurt: Peter Lang.

Vinke, D. (1995). *English as the medium of instruction in Dutch Engineering Education*. Doctoral thesis. Delft: Delft University Press. Retrieved from: https://repository.tudelft.nl/file/File_513d687d-300b-44a3-a98d-4260654f00fe

Wächter, B., & Maiworm, F. (2014). *English-Taught Programmes in European Higher Education. The State of Play in 2014*/Bernd Wächter, Friedhelm Maiworm (eds.)– Bonn: Lemmens Medien GmbH. Retrieved from: https://www.lemmens.de/dateien/medien/buecher-ebooks/aca/2014_english_taught.pdf

Ward, J. (2009). EAP reading and lexis for Thai engineering undergraduates. *Journal of English for Academic Purposes, 8*, 294–301. https://doi.org/10.1016/j.jeap.2009.10.002

Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes*, 6(), 18–35. https://doi.org/10.1016/j.jeap.2006.10.001

Ward, J. (2001). EST: Evading scientific text. *English for Specific Purposes 20,* 141–52. http://dx.doi.org/10.1016/S0889-4906(99)00036-8

Ward, J. (1999). How large a vocabulary do EAP engineering students need? *Reading in a Foreign Language*, *12*, 309–23. Retrieved from:

https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/186ef026-ade4-407c-9966-0dbc0a4ed2bf/content

Wellington, J. J., & Osborne, J. (2001). *Language and literacy in science education*. Open University. Buckingham · Philadelphia.

Wright, M., Bergom, I., & Bartholomew, T. (2019). Decreased class size, increased active learning? Intended and enacted teaching strategies in smaller classes. *Active Learning in Higher Education*, *20*, 51-62. https://doi.org/10.1177/1469787417735607.

Yang, H. (1986). A new technique for identifying scientific/technical terms and describing scientific texts. *Literary and Linguistic Computing*, *1*(2), 93-103. https://doi.org/10.1093/llc/1.2.93

Young, L. (1995). University lectures – macro-structure and micro-features. In J. Flowerdew (A c. Di), *Academic Listening* (1ª ed., pp. 159–176). Cambridge University Press. https://doi.org/10.1017/CBO9781139524612.013

Zhang, L. and Lo, Y. Y. (2012). EMI teachers' use of Interactive metadiscourse in lecture organization and knowledge construction. In Lasagabaster, D. & Doiz, A. (Eds.), Language use in English-Medium Instruction at university: International perspectives on teacher practice, p. 56-79. New York, NY; Abingdon, UK: Routledge. http://dx.doi.org/10.4324%2F9781003134534-4