Linguistics and Education 91 (2026) 101479

Contents lists available at ScienceDirect

Linguistics and Education

journal homepage: www.elsevier.com/locate/linged

Knowledge-building in EMI STEM lectures: A constellation and sequencing analysis

Nashwa Nashaat-Sobhy [a,*], Tom Morton [b]

[a] Universitat Politècnica de València, Departamento de Lingüística Aplicada, Carrer del Paranimf 1, 46730, Gandia, València, Spain
[b] Universidad Autónoma de Madrid, Departamento de Filología Inglesa, Campus de Cantoblanco, 28049, Madrid, Spain

ARTICLE INFO

Keywords:
English-medium instruction (EMI)
Explanations, Code theory (LCT)
Knowledge building
Semantic density
STEM

ABSTRACT

This study examines how a lecturer in an online English-medium instruction (EMI) setting navigates the complexity of presenting the MATLAB programming environment to undergraduate aeronautical engineering students. While much EMI research focuses on linguistic challenges, this paper turns attention to the structuring of disciplinary knowledge in teacher discourse. Drawing on Legitimation Code Theory (LCT), the analysis uses two complementary tools: constellations and sequencing. Constellation analysis shows how sets of ideas, practices, or stances are selected from a broader semantic field and arranged into a particular pattern. The sequencing tool examines how short passages of discourse are combined into longer, cumulative sequences that enable knowledge-building. Together, these tools are used to show how the lecturer organises and layers complex meanings relating to the MATLAB environment across extended discourse. The analysis indicates that the lecturer adopts an aggregating strategy, clustering related ideas around key terms or categories rather than linking them through linear assembling sequences. This approach is enacted through sequencing moves, most notably cumulative sequencing, which condenses and integrates dispersed meanings into progressively stronger knowledge, in contrast to segmental sequencing, which partitions ideas without cumulative integration. The paper highlights how sequencing choices shape the pedagogic trajectory through a constellation of meanings, and how these strategies contribute to making complex disciplinary knowledge more accessible to students in EMI settings

1. Introduction

English-medium instruction (EMI) is broadly defined as the use of English to teach academic subjects in settings where it is not the majority language (Macaro, 2018). It is a growing phenomenon in international higher education that is often framed as a strategy to internationalise the curriculum for all students (Beelen & Jones, 2015). While EMI supports strategic goals, it also places considerable demands on lecturers, who must communicate complex disciplinary knowledge through a language that is often not their own first language, nor that of their students (Deignan & Morton, 2022). Much of the existing literature has focused on linguistic and interactional challenges, with less attention being paid to a central pedagogical concern: how disciplinary knowledge is structured in discourse in ways that support understanding (Tsui, 2024). As students are expected to engage with unfamiliar conceptual material in a foreign language, the way knowledge is sequenced, linked, and accumulated over time in classroom talk becomes critical. This study addresses that issue directly by examining how disciplinary meanings are constructed through the lecturer's discourse in English during an online EMI lecture, with a focus on knowledge building strategies used by the lecturer to organise and layer complex content.

A significant body of EMI research has analysed the discourse strategies lecturers use to enhance clarity and accessibility, highlighting the role of metadiscourse, signposting, reformulation, and repetition in guiding students through lectures (e.g., Doiz & Lasagabaster, 2022; Sánchez-García, 2020; Zhang & Lo, 2021). While such work has shed light on how lecturers manage the flow of delivery, it often stops short of engaging with how knowledge itself is structured in discourse. Indeed, as Airey (2020) points out, if applied linguists only focus on their linguistic concerns, they risk their work being seen as irrelevant by disciplinary content specialists.

To address this limitation, some recent EMI scholarship is shifting the focus from linguistic performance to epistemic meaning-making. This "epistemic turn" in EMI research reflects an interest in how language is

* Corresponding author.
E-mail addresses: nash@idm.upv.es (N. Nashaat-Sobhy), tom.morton@uam.es (T. Morton).

https://doi.org/10.1016/j.linged.2025.101479
Received 11 February 2025; Received in revised form 13 November 2025; Accepted 17 November 2025
Available online 24 November 2025

used not just to convey information clearly, but to build and connect disciplinary concepts in ways that support cumulative learning. Studies grounded in disciplinary literacy (Airey, 2024), variation theory (Tsui, 2024), and knowledge profiling (González-Mujico & Lasagabaster, 2024b) have all pointed to the need to analyse how classroom discourse supports students in accessing disciplinary ways of knowing. However, few studies have examined how this knowledge-building unfolds in real-time lecturer discourse, and even fewer have used analytical tools that reveal the epistemological structure of what is being taught.

This study contributes to this line of work by applying tools from the sociological framework of Legitimation Code Theory (LCT) to analyse how knowledge is structured and sequenced in online teaching of the programming environment MATLAB to students on an aeronautical engineering degree. LCT provides a set of conceptual and analytical dimensions and tools for examining how knowledge is organised, legitimated, and made accessible in education. Using the Semantics dimension of LCT, the analysis draws on two specific tools: *constellation*, which maps how key disciplinary concepts are introduced and connected, and *sequencing*, which traces how discourse unfolds to support cumulative knowledge building, understood as the process by which knowledge claims are integrated with previous ones and refined over time (Maton, 2014, p. 107).

Three sessions focusing on introducing and providing practice with the MATLAB environment were selected for analysis. The analysis examines how the lecturer handles complex disciplinary content through discourse, specifically, how ideas are introduced, connected, and built over several turns. By analysing the lecturer's use of constellation-building and sequencing strategies, the study sheds light on how lecturer discourse can support the intelligibility of technical knowledge. In doing so, the study contributes to ongoing efforts in EMI research to place knowledge-building at the centre of pedagogic attention, rather than just language performance.

2. From rhetorical strategies to knowledge building: engaging with disciplinary pedagogy in EMI lecture discourse

Emerging lines of inquiry in EMI, in both secondary and tertiary education, have begun to shift attention away from surface delivery and instructional talk as linguistic and rhetorical performance. Instead, they focus on how instructional talk supports access to epistemic knowledge structures (e.g., key concepts, conceptual relationships, and explanatory frameworks). This conceptual shift draws on different theoretical perspectives to examine knowledge building in EMI. One perspective draws on variation theory and the concept of structure of awareness (Lo & Marton, 2011), which views learning as the process of noticing and focusing on what is conceptually important. Drawing on variation theory, Tsui (2024) shows how questioning can enable secondary school students to discern critical aspects of physics content, making them more salient, and that such questioning was less effective in EMI than in L1. In another study Tsui et al. (2024) adopt a discourse-analytic approach to question the application of an SLA lens to EMI, specifically the use of *Language-Related Episodes* (LREs) alone as an analytical tool in EMI. They argue that LREs may or may not contribute to disciplinary learning, and that their presence alone fails to capture the effectiveness with which teachers address language issues during content learning. Complementing these perspectives, a disciplinary literacy perspective emphasises the importance of socializing students with the ways of reasoning and communicating that are valued within specific academic disciplines (Moje, 2008). From this standpoint, language is not simply a tool for transmitting facts but is central to constructing and accessing disciplinary knowledge. Airey (2024) similarly argues that EMI teaching should focus on helping students access the ways of knowing characteristic of the discipline rather than aim for native-like English, and that English should be regarded as the medium of epistemic access, not the object of instruction. Taken together, these studies reinforce the idea that metadiscourse, clarity, and fluency, while necessary, are not sufficient for effective disciplinary teaching, converging on a common concern: that language in EMI should be understood not simply as a vehicle for delivery, but as a resource for knowledge-building.

Within this broader epistemic turn, LCT as a framework, developed by Karl Maton (2014) for exploring and interpreting forms and issues of knowledge building, has been gathering momentum in studies across different fields in schooling and higher education. LCT builds on central questions explored by Bourdieu, who examined how knowledge is shaped by power relations and social fields (e.g., Bourdieu, 1990), and by Bernstein, who investigated how knowledge is transformed into pedagogic communication through curriculum and discourse (e.g., Bernstein, 2000). LCT focuses on the underlying, often implicit, organizing principles that determine what counts as legitimate knowledge, and how it is communicated. It has been widely used to explore curriculum design, classroom discourse, assessment practices, and the organisation of expertise. In the context of EMI, LCT offers a way to go beyond surface-level language and explore how knowledge itself is selected, structured, and communicated.

Recent EMI studies have drawn on the Semantics dimension of LCT, which is concerned with complexity (semantic density) and context-dependence (semantic gravity) of meaning. For example, in a study focusing on an academic English programme supporting EMI students, Brooke (2020) demonstrates how teachers can use semantic waving, moving between abstract and concrete meanings, to support students' conceptual understanding. Importantly, this approach shows how epistemic concerns can inform language-focused pedagogy in ways that are meaningful for disciplinary lecturers. Also using the Semantics dimension of LCT, Argüelles-Álvarez and Morton (2023) show how shifts in semantic gravity and semantic density enable lecturers to move between foundational explanations and more abstract representations in live spoken discourse. They show how these shifts guide student attention and manage conceptual progression, offering insight into how discourse supports or inhibits knowledge-building. González-Mujico and Lasagabaster (2024a) extend this work to the visual mode. Using LCT Semantics as well, they analyse PowerPoint slides in EMI and L1-medium STEM lectures, showing that EMI slides often display flatter semantic profiles, meaning there is less movement between abstract and concrete meanings, and fewer conceptual connections. In another study, González-Mujico and Lasagabaster (2024b) examine the effect of discipline and multimodality in knowledge building practices. They demonstrate how both modalities can either support or constrain the semantic flow of meaning-making in EMI, particularly when lecturers rely heavily on pre-prepared materials.

Other studies have drawn on the Autonomy dimension of LCT, which examines how knowledge building practices and purposes can be more or less isolated from each other. In EMI, Bozbiyik and Morton (2024) examine how examples are used to navigate these boundaries, showing how lecturers discursively position knowledge relative to disciplinary norms. Together, these studies constitute a distinctive strand within the epistemic turn in EMI. They use LCT to uncover how disciplinary knowledge is constructed, layered, and sequenced, whether through speech, visual representation, or boundary positioning. This body of work offers a way to move beyond the question of how lecturers speak English, toward the more fundamental question of how they build meaning in and through English. The present study builds on this emerging LCT-informed tradition by applying two tools, constellation and sequencing, to examine the real-time unfolding of disciplinary knowledge in an EMI STEM lecture. While previous EMI studies using LCT have focused on visual materials such as slides (González-Mujico & Lasagabaster, 2024a), or the use of examples (Bozbiyik & Morton, 2024), this study analyses how disciplinary concepts are introduced, related, and cumulatively developed in moment-by-moment teaching. The next section introduces these two tools and explains how they were used in the analysis.

3. Analysing knowledge-building in EMI discourse: constellations and sequencing

At the heart of teaching is the act of explaining: making ideas meaningful by selecting, relating, and organising them in ways that learners can follow. Explanations typically involve not just isolated facts or definitions but networks of related concepts that must be developed and navigated in real time. The following subsection outlines the analytical approach adopted to capture this process of knowledge building in this study.

3.1. Overview of the analytical approach

The choices lecturers make in selecting and arranging concepts into particular patterns and sequences reflect their stance, understood here as a set of ideas, beliefs, and practices that shape decisions about which ideas to emphasise, how to connect them, and how to guide learners through them. For example, in explaining why tides occur, one teacher might begin with the gravitational pull of the Moon, while another might start with the Earth's rotation or the distribution of land and sea (Maton & Doran, 2021). These choices reflect different stances and result in different entry points and pathways into and through the topic for building knowledge. One way of tracing these pathways is provided by constellation analysis, which is part of LCT's broader cosmological analysis (Maton & Doran, 2021). This tool makes it possible to trace how meanings are introduced, revisited, and integrated across an explanation according to the lecturer's organising logic.

At the same time, because explanations unfold over time, building cumulative and complex meanings through discourse, it is also necessary to examine how that conceptual complexity is developed. In LCT, this process is captured through the concept of semantic density, which refers to how much meaning is condensed in discourse. In classroom practice, lecturers often shift between more and less dense meanings as they move between abstract principles, concrete examples, and contextual explanations. These shifts are particularly important in EMI settings, where complexity must be carefully managed through the medium of a second language.

A type of semantic density that is of particular relevance for such settings is epistemic semantic density (ESD). ESD is the condensation of disciplinary meanings and a means of understanding complexity in discourse. To account for this condensation and its contribution to complexity, different LCT tools can be used to analyse ESD (Maton & Doran, 2017a; 2017b), depending on the level of discourse: within terms (using the wording tool), across phrases or clauses (using the clausing tool), or across longer stretches of talk (using the sequencing tool). This study focuses exclusively on the sequencing tool, which is used to examine how ESD develops across extended discourse, as ideas are introduced, connected, and elaborated across a sequence of turns to complete an explanation.

Constellation analysis and sequencing are used together in this study to examine how a complex foundational topic in an EMI programming course is both conceptually structured and discursively sequenced. The next two sub-sections introduce each tool in more detail before explaining how they were applied in the analysis.

3.1.1. Constellation analysis

In this study, Maton and Doran's (2021) constellation analysis is used to trace the epistemological structure of a lecture: how units of meaning are selected, grouped, and related in order to build understanding of complex disciplinary content. These units of meaning, referred to as *nodes* in LCT, can be a single idea, entity, or concept, ranging from specific terms to abstract notions. Constellation analysis examines how nodes are arranged into *clusters*, which are groupings of meanings that are conceptually or functionally related, and how those clusters are combined into larger structures or *constellations*, carrying increasing layers of meaning as the explanation unfolds.

In constellation analysis, the distinction between nodes and clusters is fluid, depending on whether one zooms in on the individual nodes or zooms out to view them as a collective cluster. As Maton and Doran (2021) explain, when a group of related ideas is tightly integrated and referred to collectively as a single concept, "clusters or even whole constellations are condensed into a new node" (p. 53). For instance, in an explanation of tides, a teacher may first introduce gravity, the Moon, and the Sun as separate nodes. At a later point, the teacher may bundle these ideas together under the single concept of "gravitational forces," treating them as one integrated unit of meaning. That cluster, now condensed and functioning as a single node, can then be related to other clusters (such as *water movement* or *tidal patterns*) within the broader constellation of the explanation.

Maton and Doran (2021, pp. 54–55) identify three key distinctions that guide constellation analysis. Together, these distinctions show ways in which lecturers may link ideas, layer core and supporting concepts, and choose whether to build knowledge progressively or separately before connecting parts. These distinctions involve:

1. Independent vs Dependent Links: *Independent links* connect nodes that express relatively complete meanings on their own. For example, statements such as "The Moon has gravity", or "The Earth's surface includes water and land" are meaningful independently; each can stand alone without requiring the other to make sense. Independent links are like standalone explanatory strands. A node with many independent links is therefore one that helps initiate or anchor conceptual understanding (as seen with the Moon's pull in the tides example).

   Dependent links, in contrast, involve nodes whose full meaning only emerges through their connection to others. For instance, linking "The Moon has gravity" with "The Earth is covered in water" creates a new, dependent meaning that explains the phenomenon of tides. Neither statement alone explains tides; it is their combination that constructs the fuller understanding. Analysing these links helps trace how lecturers build complex meanings by integrating relationships over time.
2. Base vs Supplementary Clusters: A *Base Cluster* forms the backbone of an explanation. It is an essential set of nodes that must be introduced first to create a coherent understanding of the topic. For example, that (1) gravity is a force, (2) it pulls objects toward the Earth, and (3) this is why things fall when dropped are three nodes form the base cluster needed to understand what gravity is.

   A *Supplementary Cluster*, in contrast, expands or refines the base, adding detail, examples, or elaboration. For example, once the students understand gravity as the base cluster, a teacher may then add that gravity also keeps the Moon orbiting Earth. This distinction is not about categorising knowledge as more or less important, but how explanations are layered, with some clusters providing the conceptual foundation and others supporting or extending it.
3. Assembling vs Aggregating Structures: In *Assembling*, nodes and clusters are built progressively. One concept leads to another, forming a chain of ideas that grow from a common base, a process we compare to a plant unfolding from a seed. In *Aggregating*, separate clusters are developed independently and brought together later, more like assembling a bouquet from distinct flowers. These two strategies reflect different ways of building complexity. Assembling supports tightly integrated explanations; aggregating presents knowledge in discrete segments that may or may not be clearly connected.

In addition to these three distinctions, Maton and Doran (2021) also distinguish between two kinds of constellations that help us understand variation in how knowledge is organised: *the schematic constellation* and *the pedagogic constellation*. The schematic constellation refers to the ideal structure of a topic as it appears in official materials, such as a textbook or syllabus. It lays out the core concepts and relationships that are expected to be covered. The pedagogic constellation, by contrast, captures how a particular lecturer chooses or happens to navigate that structure in practice. This includes decisions about which nodes to emphasise, how to link them, and how to order their presentation.

While constellation analysis displays knowledge building as structures of nodes and clusters, with their links creating the overall architecture, sequencing analysis examines whether knowledge is progressively built through discourse connections, realising stronger ESD, or remains compartmentalised through weaker ESD.

3.1.2. Sequencing types and ESD

As Maton and Doran (2017b) explain, sequencing analysis considers how meanings accumulate, condense, and interrelate across passages of talk or text. Sequencing has different types and subtypes, which correspond to varying degrees of epistemological condensation (EC). Their types fall along a continuum of knowledge condensation strengths, with stronger EC (EC+) indicating more knowledge complexity and weaker EC (EC-) reflecting lower levels of complexity, and distinctions can be made among and within sequencing types.

At one end of the continuum, cumulative sequencing promotes knowledge-building by explicitly linking meanings across discourse. These links enable new understandings to emerge, creating conceptual progression that exceeds the sum of its parts. Cumulative sequencing includes both vertical and horizontal forms. Vertical sequencing refers to moments when earlier meanings are picked up and bundled into more condensed expressions, or "transporters", often through nominalisation or referential terms such as "this process" or "those ideas." Depending on the span of prior discourse being integrated, this can be either subsumptive (drawing from a single prior segment) or integrative (gathering multiple earlier meanings). Horizontal sequencing, by contrast, builds cohesion by linking adjacent passages through logical or temporal relationships. It includes consequential links such as "because", "so", or "if…then" and sequential links that structure procedures or timelines, using expressions like "then" or "after that".

Lower on the continuum of EC is segmental sequencing, in which no explicit linking of terms from related passages occurs. Here, discourse proceeds through discrete segments that are not explicitly epistemologically connected, reducing opportunities for semantic development. Segmental sequencing can be sedimental or compartmental. Sedimental sequencing involves reiteration, repeating similar ideas or terms either with variation (reiterative) or verbatim (repetitive). While such repetition may serve to emphasise or maintain attention, it adds little in terms of building complexity. Compartmental sequencing occurs when unrelated ideas are placed side by side without meaningful linkage. This may seem coherent, in that topics are topically related, or incoherent when ideas shift abruptly with no evident connection.

Together, these distinctions offer a toolkit for analysing how knowledge unfolds in classroom discourse. They help identify points where complexity is built through meanings being actively accumulated and integrated, and can potentially identify passages where attempts to build complex knowledge break down. Table 1 summarises these types and subtypes, along with examples and notational conventions. Vertical sequencing is marked with <<double angle brackets>> around the passages which are integrated, and ⟨single angle brackets⟩ around the transporter. Horizontal sequencing is indicated with {{double curly brackets}} around the passages which are linked together and {single curly brackets} around linking markers. Sedimental sequencing is indicated with double vertical bars ||, while compartmental sequencing remains unmarked.

We consider sequencing analysis to be a complementary lens to constellation analysis, with the potential to provide a richer account of knowledge-building in teacher discourse. As Maton and Doran (2021, p. 53) observe, "clusters or even whole constellations are condensed into a new node, which can then be constellated with more nodes", highlighting how sequencing transforms and reshapes the architecture of constellations. Moreover, they conclude their analysis of constellating science by encouraging drawing on other LCT concepts, especially epistemological condensation tools, to examine their attributes (p. 73).

Our aim is to investigate how the lecturer in our case organises and connects key concepts during instruction. Some topics may lend themselves to an assembling approach (built step by step) and others may be better suited to an aggregating approach (developed independently

Table 1
Sequencing levels of delicacy and epistemic semantic density and condensation.

ESD/EC Levels	Sequencing Types	Subtype	Sub-subtypes	Examples from our data with script annotation
ESD+/ EC+	**Cumulative**	**<<Vertical>>:** Condenses meanings from one or more passages into a term in another passage, creating a dense and integrative link that builds complex knowledge.	**Integrative sequencing** condenses meanings from multiple passages into a new conceptual unit. It builds conceptual integration across extended discourse	After explaining *different classifications of data types* over several turns, the phrase <<*That is the idea*>> encapsulates the broader explanation of data types introduced in previous passages.
			Subsumptive sequencing condenses meanings from a single passage into a summarising phrase. It packs ideas into a condensed term or phrase.	Referring to <<These operations>> after naming a number of computational operations.
		{{Horizontal}}: Links passages together by conjunctive means, adding relational or causal meanings without condensing terms, enhancing logical understanding.	**Consequential sequencing** adds causal, conditional, or contrastive links between passages. It supports logical argumentation or explanation	{{{If} you see the use of the word arrays and not matrices is {because} in many computing languages the normal name for this kind of structures is arrays.}}
			Sequential sequencing adds temporal or procedural order to passages and helps structure procedural or narrative flow	{{First we calibrate the sensor, then we run the experiment}}
ESD-/ EC-	**Segmental**	**||Sedimental||:** Repetitive use of terms or concepts without adding new meaning or complexity, maintains basic understanding and reinforces existing knowledge.	**Reiterative sequencing** repeats similar meanings with slight variation. It emphasises key ideas through varied repetition	||This function *is helpful*. It is *extremely useful*||
			Repetitive sequencing repeats ideas verbatim without adding new meaning. It reinforces ideas without developing them further	||It is a *good tool*. For sure, it is a *good tool*||.
		Compartmental: Organizes knowledge into distinct but potentially related categories, may either bind well or remain conceptually separated	**Coherent sequencing** maintains topical continuity across separate ideas. It preserves focus but limits integration	MATLAB is used in aeronautical engineering. It's also used in videogaming.
			Incoherent sequencing Juxtaposes unrelated or disruptive topics without epistemic link. It disrupts conceptual flow and coherence.	MATLAB is good for numerical simulations. It's widely used in academia. Don't forget next week's task.

before being linked). Furthermore, the schematic constellation (e.g., a textbook or syllabus structure) sets expectations about what is to be taught, but the pedagogic constellation reflects the teacher's actual explanation, which may diverge.

Finally, as lecturers execute the pedagogic constellation, their discourse resources vary over the stretch of instruction. Sequencing analysis of their discourse can show if shifts in the strength of EC occur in alignment with the constellation's structure, revealing whether knowledge is progressively cumulated through cumulative forms of sequencing (EC+) or remains compartmentalised through segmental sequencing (EC-).

The research questions (RQs) guiding the study are thus:

RQ1. What choices does the lecturer make in charting a pathway through the targeted knowledge constellation?
RQ2. How is complexity reflected in the lecturer's discourse through sequencing patterns?

RQ1 addresses how the lecturer organises nodes and ideas when constructing the pedagogic constellation. RQ2 considers whether there is evidence of EC+ and thus higher knowledge complexity in the discourse, or EC- where knowledge remains separated or recycled. While both EC+ and EC- are examined, the central focus is on EC+ being the key indicator of progressive knowledge building.

By addressing how EMI lecturers can support learning by managing the accumulation and linking of meanings over time, the study demonstrates the use of LCT tools to understand better how teacher discourse shapes the organisation and potential accessibility of disciplinary knowledge. In doing so, the study contributes to a growing body of EMI research that views knowledge-building, not just language use, as central to effective teaching.

4. Methodology

4.1. The context

The study is situated in the first year of an aeronautical engineering bachelor's degree at a Spanish university. As part of their training, students are introduced to MATLAB, a software platform widely employed across universities in Spain and internationally for both basic and advanced programming tasks relevant to STEM disciplines. MATLAB is not only foundational in aeronautical engineering but also used in other degree programs at the institution, such as audiovisual communication, and in many other engineering and science degrees across Europe. Its functions in algorithm development, data visualization, data analysis, and numerical computation make it a representative tool for entry-level STEM education. Importantly, the medium of instruction for the course was English, and the students had not previously studied programming in a structured program or used MATLAB. This combination of a widely generalizable disciplinary tool and a first exposure to programming in EMI creates a particularly salient context for examining knowledge-building.

For background, the lecturer had over 20 years of teaching experience and subject expertise in programming and educational technologies. Although this was his first experience of teaching in English, the lecturer is a researcher with publications in English and had received prior training in integrating content and language in higher education.

4.2. Data collection and analysis

The data analyzed for the study consists of three online lectures conducted via Microsoft Teams during the COVID-19 pandemic in 2021. These 90-minute lectures, along with related materials, were provided by the lecturer to facilitate the research. The recordings were transcribed then analysed using Maton and Doran's (2021) approach to constellation analysis and Maton and Doran's (2017b) sequencing tool.

Informed consent was obtained from the lecturer to analyse the data of these lectures. The study did not involve prospective recruitment, social interventions, the collection of sensitive information, or the use of identifiable participant data, limiting ethical considerations to obtaining informed consent from the lecturer and ensuring the confidentiality of the provided materials and recordings.

The data was first divided into episodes following the tradition of discourse analysis, with segmentation based on the field (topic) under discussion. Each episode was assigned a sub-objective as part of familiarizing students with MATLAB and its operations. In total, 60 episodes were identified, spanning from the introductory sections to students performing operations in MATLAB. We focused on a specific teaching objective to ensure feasibility in charting the pathway taken by the lecturer. For this purpose, we selected a pivotal point: the presentation of the MATLAB environment along with its basic functions and operations.

The episodes analysed (30–34), representing this objective, were situated mid-way through the recorded sessions and lasted approximately 40 min. They were selected for detailed analysis as they feature the lecturer actively introducing, explaining, and connecting discourse typically associated with teaching presence (cf. Garrison, 2011).

For constellation analysis, following Maton and Doran (2021a), we formulated a 'schematic constellation' of MATLAB's datatypes content. The formation of this constellation relied on information highlighted by the teacher and supplemented by resources available on the MATLAB developers' website (Mathworks), used in tandem with artificial intelligence tools to unravel complex content. This served to provide a depiction of pivotal nodes in the topic of datatypes and their logical interconnections. Subsequently, we developed a 'pedagogic constellation' based on teacher's actual instruction on MATLAB datatypes.

For the visual representation of datatypes, we identified nodes (mentions of new terms or processes) at the clause or turn level. A node was marked as an independent link if it was stated as a stand-alone notion, concept, or term, and as a dependent link if its meaning relied on a relationship with another node (e.g., "is a type of"). Base clusters were classified as those sets of nodes that anchored the topic and provided the minimal explanation necessary for the discourse to hold together. Supplementary clusters were identified as additional sets of nodes that added nuance, detail, or extension (see Section 3.1). The figures were redrafted until a stable consensus was reached.

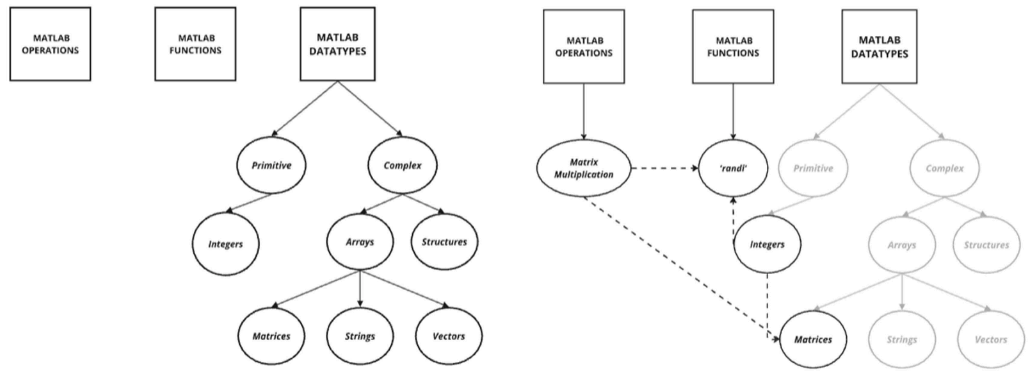
For sequencing analysis, we followed Maton and Doran (2017b) in identifying types and subtypes that realise stronger or weaker epistemic condensation (EC). We jointly developed a codebook, refined it iteratively on a data sample, then independently coded the focal dataset and resolved discrepancies through discussion until full consensus was reached.

5. Results

We begin by describing a generic schematic constellation of datatypes in MATLAB, showcasing a comprehensive assembly of common nodes constituting this topic. We then describe the pathways taken by the teacher when approaching the topic of datatypes, or the pedagogic constellation.

5.1. Schematic constellation

In the following textual description, lines A to L summarise a possible schematic constellation (idealised version based on MATLAB materials). The three major clusters in this constellation are *MATLAB Datatypes, MATLAB Functions*, and *MATLAB Operations* (Fig. 1). Among these, *Datatypes* is the central cluster and point of entry to the explanation, serving as the base cluster. *Functions* and *Operations* are introduced later as supplementary clusters, since they act on *Datatypes* and become meaningful only once datatypes are established. In this schematic, the three clusters also function as independent links, as each can initially be



(a) Three main clusters: MATLAB Datatypes as the base cluster (b) Linking supplementary and base clusters

Fig. 1. Schematic Constellation of MATLAB components. From left (a) to right (b), the figure displays the three main clusters with MATLAB Datatypes as the base cluster (a), and the linking of supplementary clusters—MATLAB Functions and MATLAB Operations—back to the base (b). Independent links are shown in squares, and dependent links are shown in circles.

explained on its own before being related to the others.

In the description below, the three **clusters** are shown in **bold**, while the *dependent nodes* (e.g., *arrays, matrices, vectors*) are shown in *italics*:

Line A: **MATLAB operations, functions**, and **datatypes** are fundamental components of MATLAB to manipulate and analyse data.
Line B: **MATLAB datatypes** are organized into '*primitive*' and '*complex*' categories.
Line C: *Primitive* datatypes include *integers*, among *other subtypes*.
Line D: *Complex* datatypes include *arrays* and *structures*.
Line E: An *array* is a general term for a collection of elements, with subforms such as *matrices, vectors*, and *strings*.
Line F: A *matrix* is a two-dimensional *array* and is the default name for this type of *array* in MATLAB.
Line G: A *vector* is a one-dimensional *array* in MATLAB.
Line H: *Structures* are containers to store data under named fields, useful for managing and organizing complex data.
Line I: *Strings* are a special type of *array* that represent text as character *vectors*.
Line J: **MATLAB functions** create or transform datatypes; for example, '*randi*' generates numeric arrays of random integers.
Line K: **MATLAB operations** compute over datatypes; for example, *matrix multiplication* takes *matrices* as inputs and returns a *matrix*, which requires the '*randi*' function.

The initial sentence (Line A) introduces learners to three primary clusters in MATLAB—Operations, Functions, and Datatypes—laying the foundation for further exploration. Among these, Datatypes is developed first as the base cluster, since further explanation depends on establishing the categories of information it provides. This base cluster is expanded progressively and linearly through assembling (Lines A–I), with each node (e.g., arrays, matrices, vectors, strings) gaining meaning through dependent links to the overarching cluster. Functions and Operations, by contrast, are introduced later as supplementary clusters. They are initially independent clusters, but their role becomes clear only when dependent links connect them back to Datatypes. An aggregating move occurs when Lines J–K link these clusters, for instance, moving from generating integer arrays (Function) to computing with matrices (Operation). The schematic constellation therefore shows a progression: beginning with what the entities are and then demonstrating how they work together. In doing so, the explanation combines assembling (Lines A–I) with aggregating (Lines J–K) to constellate knowledge of this section of MATLAB.

5.2. Pedagogic constellation

Turning to the pedagogic constellation, the pathway taken by the lecturer for the same section of MATLAB is illustrated below in Fig. 2 (a, b & c), where the grey parts denote the points not mentioned in each turn. The teacher first presents MATLAB Datatypes independently of the other two main clusters (Fig. 2a), Functions and Operations. Here, datatypes serve as the entry point to the explanation and represents the base cluster, comprising dependent taxonomic nodes that classify datatypes into primitive and complex and name their subtypes.

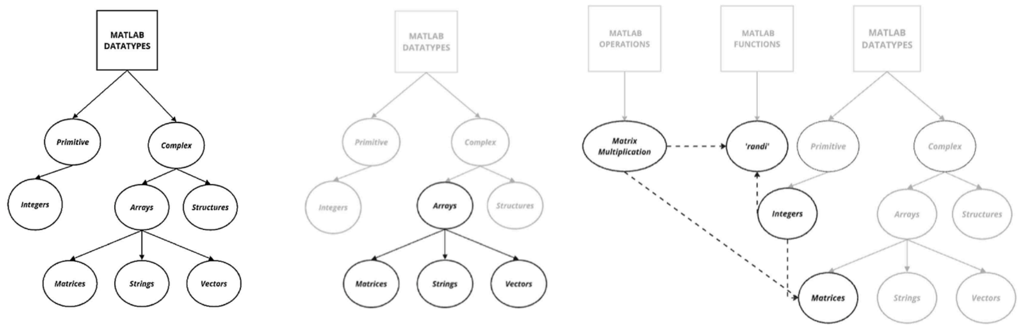
In a subsequent turn (Fig. 2b), unlike in the schematic constellation (Fig. 1), 'Arrays' are elaborated further, with attention to subforms such as matrices, vectors, and strings. This is an assembling move that unpacks arrays in more detail, yet it suggests that the students received only part of the information in each turn.

In a third turn (Fig. 2c), the teacher introduces the function 'randi', which theoretically links functions, datatypes and operations by bringing together the need to generate random integers, populate matrices, and perform matrix multiplication (creating a cross-cluster chain of dependent links). However, the other two main clusters, Functions and Operations, are not named. As a result, '*randi*' and '*matrix multiplication*' appear as unanchored dependents for students, who may not recognise which cluster they pertain to.

Overall, the constellation unfolds as follows: (a) an initial assembling move that lays out the base taxonomy; (b) a supplementary/aggregating move that deepens the relatively smaller arrays cluster; (c) and a cross-cluster move that links functions, datatypes, and operations via aggregation, creating a dependent chain. In contrast to the schematic constellation, the pedagogic version creates partial depth within the base cluster, leaving cross-cluster relations unanchored. Because the two supplementary clusters (Functions and Operations) are never explicitly named in the explanation, their connections remain opaque to students, reducing the perceived coherence of the constellation for novices, despite being analytically traceable. In comparison to the schematic constellation, the pedagogic constellation thus appears more fragmented.

5.3. Sequencing in the lecturer discourse

Turning to the second research question, we provide extracts from the lecture discourse to show how the lecturer uses sequencing to build the pedagogic constellation. Extract 1 illustrates the sequencing approach in the initial presentation of basic datatypes in MATLAB.



(a) Base Cluster: MATLAB Datatypes (b) Arrays in MATLAB (c) ‘*randi*’, integers, matrices, and m. multiplication

Fig. 2. Illustration of the sequence in the teacher’s pedagogical constellation of MATLAB. From left (a) to right (c), the figure displays independent links within the squares, and the dependent links are in circles. Elements in grey represent those not mentioned to denote the lecturer’s emphasis in each turn.

Extract 1

So, I think okay all as always, you have the sources within programming, you have the theory -data types, array, strings. I have this information here.
{{This unit is the unit addressed to explain the use of data types, sorry it was… the basic data types, {because} there are other data types as Carlos commented in the session.}}
>>And there are the first classification of data types, I will describe later, between primitive and more complex, like arrays for example or matrices.
{{{If} you see the use of the word arrays and not matrices is {because} in many languages many computing languages the normal denomination or the normal name for this kind of structures is arrays.}}
{{In C Programming C language programming or PHP is for web, most of the programming languages uses the word use the word arrays as the general term and {then} matrices can be - the concept is the same, it’s very similar, but matrices is how arrays are considered in MATLAB.}}
{{And {then} later a string is similar {because} they are implemented in MATLAB and set the basic types.}}
All is implemented using matrices or arrays it’s the same. For example, be careful about the denomination for vectors. Vector is an array or a matrix in one dimension.
<That is idea> about the basic data types.>>
And we are mainly focused on these arrays and strings for the current session.
Remember that the classification of the data types are divided into two basic parts or sides the primitives or scalar diskette types as floats, integer, Boolean, true or false, character and enumerated and these are simple types.
{{{So}, the second part is about the composites the more complex including the arrays, string and structures.}}
And there are several types of structures.
{{This is very fast or an overview of the main data types, but the structures even can be divided in cells in tables, {even so}, they are different or they are more data types, besides the types named here, or just mentioned here.}}

The passages that are linked using both sequential (“then”) and consequential sequencing (“because”, “so”, “even so”) represent horizontal sequencing (EC+). In this way the lecturer builds up relations between elements in the constellation, both in a linear stepwise order (e.g., “and then later a string…”), stating what comes “next”, and through cause–effect relations (e.g., “…arrays not matrices because…”). At the same time, an example of sedimental sequencing is also present (EC–) in the repetition “arrays or matrices … it’s the same”, which maintains attention but does not build new complexity. This form of sedimental repetition helps consolidate terminology for novices without helping them integrate the concept into the wider constellation. Overall, the more frequent use of horizontal sequencing here helps make relations between different nodes in the constellation explicit, rather than leaving the students to work out for themselves how they are connected.

There is one possible example of integrative vertical sequencing (EC+), where the lecturer bundles up what has been explained about basic data types in previous short passages using a transporter with the unorthodox wording “that is idea”. This allows the preceding details about datatypes to be condensed into a portable block of meaning. However, because the terms arrays, matrices, and strings are not securely tied back to the overarching Primitive vs. Complex taxonomy, the strength of this integrative sequencing is weakened. ‘Strings’ slips between being described as “basic types” and as “complex types.” This inconsistency is likely a slip, but it could be confusing, as it prevents students from clearly locating strings in the broader classification. Out of the portable block of data types, only a limited number will be the focus of the current session, as seen in the next line where the lecturer says that “we are mainly focused on these arrays and strings for the current session”.

As the lecturer progresses through the pedagogical constellation, nodes mentioned previously are revisited for further elaboration and new ones are added (Fig. 2). Thus, terms like ‘matrices’, ‘strings’, and ‘vectors’ are taken up again, and expanded with related details.

Extract 2 shows how the node “strings” is elaborated and integrated across Datatypes and Operations. In this extract, the lecturer returns to the node strings, identifying them as a “special type of array”, which is also “stored as a matrix”, thus realising EC+ through horizontal consequential sequencing, aggregating and integrating these three nodes (strings, arrays, matrices). He also adds information about the purpose of strings (“to store characters”), further consolidating their role.

Extract 2

Okay, then, we’re going to manipulate strings. The string is a special type of array that is used to store characters, or to store streams of character strings, that is the denomination for this data type. But it’s also stored as a matrix or as an array.
{{{So} the next objectives are related to the performing operation, with strings or with matrices, mathematical operations, solving equation, for example.}}
<<In general, generating random numbers.
{{<That> will be the first operation I’m going to do {in order to} select the first student to do the first task or the first to solve the first exercise.}}>>

The lecturer goes on to use causal horizontal sequencing (EC+) with “so” (“So the next objectives are related to the performing operation…”) to extend the constellation from data types into operations. This move links strings not only to arrays and matrices but also to mathematical operations. In the final part of the extract, he employs subsumptive vertical sequencing (EC+), using the transporter “that” to compress the idea of generating random numbers into a single unit. This condensed unit can then be carried forward as the basis for the next classroom activity (selecting a student), showing how sequencing turns a stretch of explanation into a portable block of meaning.

At the same time, weaker sequencing is also evident (EC–). The reference to “generating random numbers” is not embedded into the developing set of mathematical operations but instead framed in terms of classroom management (selecting a student). As a result, the new element is juxtaposed with, rather than integrated into, the conceptual progression of operations. This reflects compartmental sequencing (EC–), as the operation is added without being fully connected to the larger structure of the constellation. For novices, this juxtaposition may lead them to treat the operation as a classroom routine rather than as part of a coherent operations cluster, temporarily maintaining fragmentation in their understanding of MATLAB processes if the underlying conceptual link remained unnoticed.

Overall, Extract 2 shows how the lecturer extended the constellation by linking datatypes with operations (EC+), while also leaving parts of the conceptual structure compartmentalised (EC–).

In Extract 3, which occurs a few moments later, the lecturer turns his attention to matrices.

Extract 3

<<Okay and the same for the creation of matrices
{{you know that there are different function for example, {to} create a zero matrix.}}
A matrix composed by two rows and three columns, but only composed by zero values on the same with all ones that I think you are familiar or you know, how to create this this kind of matrix. And another aspect to remark is the way to access the components of a matrix
{{you have different ways to access, using the brackets, the round brackets, {in order} to select the different elements.}}
Okay, I think all <this information> is I suppose, I guess that is already known>>
<<and how to manipulate or how to work with different matrices how to perform mathematical operation, addition over matrices or how to multiply matrices.
{{We will review the difference for example between a vectorial multiplication {or} matrix multiplication regarding integer with multiplication or scalar multiplication,}}
There are different kinds of operation about the use of matrices.
{{Okay, <that> is only an overview>> about the {because}along the the different exercises, you will be able to show how you how are you solving the exercises that involve the use of ⟨this function⟩ {or} the manipulation of matrices of the operation that can be performed.}}

This extract shows how the lecturer continues with the aggregation strategy by returning to matrices, linking them across the constellation to functions, with a focus now on the creation of matrices. He adds further information to the nodes and uses horizontal sequential sequencing (EC+) to describe different ways of accessing the components of matrices. He completes this block of the pedagogical constellation by using integrative vertical sequencing (EC+) ("this information") to bundle up and summarise what has been covered in the preceding discourse, before moving on to the node of operations in the left-hand side of the constellation, using horizontal consequential sequencing (EC+) with "or" to identify alternative types of multiplication with matrices (e.g., vectorial or scalar).

He ends this block of aggregation by again using integrative vertical sequencing (EC+) to wrap up the whole bundle of information presented as an "overview". He also uses horizontal consequential sequencing (EC+) to causally link the overview to what the students will be doing in the exercises – that is, it is enough to have only an overview as they will get a chance to practice in the exercises.

There is also an indication of subsumptive vertical sequencing (EC+) here, in the use of "this function", which may refer back to the randi function he mentioned earlier, and horizontal sequencing ("or") is used to identify the two elements from the left-hand side of the constellation (functions and operations) which will be worked on.

The sequencing analysis reveals that both horizontal and vertical moves (EC+) built links between clusters and compressed stretches of discourse into portable units, yet sedimental repetition and compartmental juxtapositions (EC–) created, or at least sustained, separation between clusters. Taken together, the results indicate that while some complexity was progressively built within clusters, broader integration remained partial, leaving the overall pedagogic constellation fragmented.

6. Discussion

This study examined how a lecturer charted a pathway through the topic of MATLAB datatypes and how epistemic semantic density was managed through sequencing in English. Our analysis compared a schematic constellation, or an idealised map of the topic based on a summary of different materials, with the pedagogic constellation, or how the same knowledge was actually unfolded in lecture discourse.

In the schematic constellation (Fig. 1), Datatypes functioned as the base cluster, progressively elaborated through assembling (i.e. linear expansion step by step: primitive vs. complex, arrays, matrices, vectors, strings). Functions and Operations were supplementary clusters: initially independent but later integrated with Datatypes through dependent links (e.g. the function randi generating arrays; the operation of matrix multiplication acting on those arrays). In this way, the schematic constellation combined assembling with aggregating, bringing separately developed clusters together.

In the pedagogic constellation (Fig. 2), however, the lecturer left implicit both Functions and Operations: they surfaced only in the form of related examples (randi, matrix multiplication) without being named as main clusters. Therefore, while the initial assembling within Datatypes was clear and strong, the aggregating moves across clusters appeared unanchored (see the grey areas in Fig. 2). Students were given fragments of a larger structure without always being shown how they fit together, echoing Maton and Doran's (2021) observation that aggregating strategies often provide "bits" before the "whole". In this case, it was precisely the reliance on aggregating moves, involving the separate development of clusters without naming or tying them back to the schematic base, that directly produced the fragmented pedagogic constellation observed in our analysis.

Sequencing resources in the pedagogic constellation reinforced this pattern. Horizontal sequencing (EC+), realised through connectives such as then, because, and so, linked successive nodes in cause-and-effect or temporal relations: "arrays not matrices because …", "and then later a string …". These moves built local coherence by explicitly relating items in the constellation. Vertical sequencing (EC+), through transporters such as "that is idea" or "this information", bundled stretches of detail into portable meaning blocks, allowing the discourse to move forward. In Maton and Doran (2017b) terms, horizontal sequencing "unpacks" complexity into smaller steps, while vertical sequencing "repacks" them into condensed knowledge.

Weaker forms of sequencing were also present. Sedimental sequencing (EC–) appeared in a few repetitions ("arrays or matrices … it's the same"), which helps maintain attention but did not add new meaning. Compartmental sequencing (EC–) occurred when new elements (e.g. random numbers for classroom management) were introduced without integration into the constellation of operations, leaving them juxtaposed rather than connected. While these weaker moves do not undermine the explanation, they do not appear to support the cohesion and the integration of the presented topic.

Taken together, the analyses suggest that the pedagogic constellation (Fig. 2) allowed students to develop local complexity but only partial global integration. This pattern resonates with Maton and Doran's (2017) finding that sequencing is not simply about linking sentences but about shaping how knowledge complexity is managed. In this case, assembling moves and horizontal sequencing supported stepwise access to detail, but the aggregating move that revisited 'Arrays' and the absence of stronger integrative moves across clusters could have weakened epistemic access to the overall structure of MATLAB. This fragmentation would make it particularly difficult for novice students to reconstruct the broader disciplinary logic. Pedagogically, the implication is not that the lecturer was ineffective, but that students' grasp of disciplinary knowledge depends on whether relations among clusters were clear to them. Novice students, for example, who are less able to infer cross-cluster links independently, are therefore especially likely to have struggled.

In connecting these findings to previous EMI research, our findings build on the work of Argüelles-Álvarez and Morton (2023) by extending their focus on the role of varying semantic density to scaffold comprehension. Specifically, while their study emphasised shifts in density, we add that the lecturer's constellation-building choices, whether favouring aggregation or assembling, also crucially shape how complex disciplinary concepts are introduced and related.

Furthermore, our study aligns with Tsui et al. (2024), who highlight how disciplinary access depends not only on linguistic clarity but also on how concepts are patterned and linked over time. This supports our claim that structuring disciplinary knowledge is a central pedagogic matter, not simply a linguistic one, and that metadiscourse strategies do not occur in a linguistic vacuum. In practical terms, this means that language choices and metadiscursive strategies (e.g., Doiz and Lasagabaster, 2022; Zhang & Lo, 2021) should be understood in relation to broader knowledge-building practices. Rather than viewing these approaches as separate, lecturers could benefit from integrating both: using metadiscursive cues to signal relationships between concepts while deliberately sequencing disciplinary content.

7. Conclusion

For EMI practice, the present analysis suggests three areas that may be especially fruitful for teachers and trainers to consider when planning instructional pathways and related metadiscourse. First, a useful step could be to reflect on whether relations among instructional ideas are made visible. One way of doing so is by sketching a pedagogic constellation in advance, explicitly identifying the main clusters in the topic and deciding which nodes need to be treated as dependents of a cluster. Second, closer attention to deliberate sequencing cycles may support coherence. Horizontal sequencing can be used to open up complexity step by step (*unpack*), and vertical sequencing to close it down into portable summaries (*repack*), with transporters pre-planned to signal what is being bundled (e.g., "this is our definition of arrays; that procedure belongs to the *Functions* cluster"). Third, it may be helpful to anticipate and address pressure points, understood here as moments where cumulative knowledge-building is most vulnerable, such as when taxonomic categories are blurred (e.g., a concept could be both basic and complex, depending on the situation), or when cross-cluster links are implicit by default and need clarifying, specifically requiring linguistic transporters to signal integration. These are not prescriptive steps but possible avenues for strengthening epistemic access in EMI.

A key insight of this study is that such a "non-linguistic" approach, despite its primary focus on knowledge practices rather than language itself, can paradoxically lead to more targeted and meaningful English language development for EMI lecturers. By identifying pressure points in pathways through constellations—such as where transporters fail to convey how complex knowledge is bundled and integrated—applied linguists and academic developers can pinpoint where language interventions are most needed. This complements a focus on behavioural pedagogy or metadiscourse by drawing systematic attention to how language use directly impacts knowledge-building. Rather than treating language development as a separate or incidental concern, this approach promotes an integrated perspective where linguistic accuracy serves disciplinary learning outcomes. For EMI lecturers, this alignment can be empowering, as it ties language development to their core objective of effectively communicating disciplinary knowledge.

For using LCT in EMI, constellation and sequencing offer a way of engaging with disciplinary knowledge-building that remains accessible to content specialists. As Maton and Doran (2017a) emphasise, the typologies "are not linguistic concepts, … but are "sociological concepts for revealing an attribute of knowledge." This makes the framework appealing to lecturers invested in how their subject knowledge is structured and conveyed who may not engage with linguistic concepts.

At the same time, our analysis shows that the effectiveness of these knowledge-building approaches is still influenced by how lecturers use English in real-time classroom discourse. While linguistic competence was not the focus of this study, it is a legitimate concern for EMI support. Future research could explore how specific acts of sequencing are realised through linguistic resources (lexis, syntax, prosody, semantics) and in interaction with students. Instances of "awkward" expression (Smiskova, Verspoor & Lowie, 2012), such as slips in vertical sequencing, suggest that even minor errors can have significant pedagogical consequences if they occur at pressure points. Ambiguity or weak transporters at such moments may prevent students from linking elements of the constellation, constraining their ability to build knowledge cumulatively.

For LCT more broadly, the study contributes a pedagogic use case that highlights how local EC+ resources may be strong while global integration remains fragile. This echoes Maton's (2014) distinction between cumulative and segmented knowledge-building. It also opens avenues for connecting constellation and sequencing with other LCT tools (e.g., clausing, autonomy) and using multimodal data, since slides, gestures, and code may either reinforce or obscure the relations made in talk (see González-Mujico & Lasagabaster, 2024a, 2024b).

Limitations and future directions should be acknowledged. This is a single-lecturer, single-topic case in an online format, and both the subject matter and delivery mode are likely to have shaped the knowledge-building strategies observed. Broader comparisons are needed across disciplines, particularly between STEM and the humanities, across face-to-face and blended settings, and across topics with different epistemological structures.

Despite these limitations, constellation and sequencing analyses provide both analytic clarity for researchers and a reflective lens for practitioners. They can help EMI lecturers to tie language directly to disciplinary knowledge-building, offering concrete ways of making relations explicit and systematic. In doing so, they also help bridge applied linguistics and subject expertise, opening pathways for collaborative pedagogy that strengthens students' epistemic access.

Funding

Open access funding was provided by the Universitat Politècnica de València (UPV) through the agreement between Elsevier and the CRUE Spanish Universities-CSIC Alliance.

CRediT authorship contribution statement

Nashwa Nashaat-Sobhy: Writing – review & editing, Writing – original draft, Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tom Morton:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The authors do not have permission to share data.

References

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.

Airey, J. (2024). Disciplinary literacy and English-medium instruction: Multilingual, multimodal and multifunctional considerations. In A. B. M. Tsui, & E. Macaro (Eds.), *Language issues in english medium instruction* (pp. 48–59). Routledge.

Argüelles-Álvarez, I., & Morton, T. (2023). Using legitimation code theory to investigate English medium lecturers' knowledge-building practices. *Journal of English for Academic Purposes, 65*, Article 101285. https://doi.org/10.1016/j.jeap.2023.101285

Beelen, J., & Jones, E. (2015). Redefining internationalization at home. In A. Curaj, L. Matei, R. Pricopie, J. Salmi, & P. Scott (Eds.), *The european higher education area: Between critical reflections and future policies* (pp. 59–72). Springer.

Bernstein, B. (2000). *Pedagogy, symbolic control, and identity, 5*. Bloomsbury Publishing.

Bourdieu, P. (1990). *The Logic of Practice*. Stanford University Press.

Bozbıyık, M., & Morton, T. (2024). Lecturers' use of examples in online university English-medium instruction: A micro-analytic classroom interaction and knowledge-building perspective. *Language and Education, 38*(3), 344–363. https://doi.org/10.1080/09500782.2023.2173528

Brooke, M. (2020). Guiding teacher talk in the content and language integrated learning classroom using semantics from legitimation code theory. *Teaching in Higher Education, 25*(7), 812–824. https://doi.org/10.1080/13562517.2019.1598968

Deignan, T., & Morton, T. (2022). The challenges of English medium instruction for subject lecturers: A shared viewpoint. *ELT Journal, 76*(2), 208–217. https://doi.org/10.1093/elt/ccab084

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

González-Mujico, F., & Lasagabaster, D. (2024a). PowerPoint presentations in L1-medium and english-medium lectures: Examining knowledge-building practices in multimodal slide content. *Language and Education, 39*(2), 380–401.

Garrison, R. (2011). *E-learning in the 21st century* (2nd ed.). Routledge.

González-Muiico, F., & Lasagabaster, D. (2024b). Similarities and differences in EMI disciplinary knowledge construction. In A. B. M. Tsui, & E. Macaro (Eds.), *Language issues in english medium instruction: Theoretical orientations and cases from disciplinary practitioners*. Routledge.

Lo, M. L., & Marton, F. (2011). Towards a science of the art of teaching: Using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies, 1*(1), 7–22.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching, 51*(1), 36–76.

Maton, K. (2014). *Knowledge and knowers: Towards a realist sociology of education.* Routledge.

Maton, K., & Doran, Y. J. (2021). Constellating science: How relations among ideas help build knowledge. In K. Maton, J. R. Martin, & Y. J. Doran (Eds.), *Teaching science: Knowledge, language, pedagogy* (pp. 49–75). Routledge.

Maton, K., & Doran, Y. J. (2017a). Semantic density: A translation device for revealing complexity of knowledge practices in discourse, part 1—Wording. *Onomázein*, 46–76.

Maton, K., & Doran, Y. J. (2017b). Condensation: A translation device for revealing complexity of knowledge practices in discourse, part 2—Clausing and sequencing. *Onomázein*, 77–110.

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy, 52*(2), 96–107.

Sánchez-García, D. (2020). Mapping lecturer questions and their pedagogical goals in Spanish- and English-medium instruction. *Journal of Immersion and Content-Based Language Education, 8*(1), 28–52. https://doi.org/10.1075/jicb.18016.san

Smiskova, H., Verspoor, M. H., & Lowie, W. M. (2012). Conventionalized ways of saying things (CWOSTs) and L2 development. *Dutch Journal of Applied Linguistics, 1*(1), 125–142. https://doi.org/10.1075/dujal.1.1.09smi

Tsui, A. B. M. (2024). Linguistic mediation of learning processes in content classrooms. In A. B. M. Tsui, & E. Macaro (Eds.), *Language issues in english medium instruction* (pp. 29–47). Routledge.

Tsui, A. B. M., Wang, C., Tseng, J. J., Lo, M. L., & Liu, M. M. D. (2024). Teachers' language awareness and language related episodes (LRE) revisited: The case of EMI in computer engineering. In A. B. M. Tsui, & E. Macaro (Eds.), *Language issues in english medium instruction* (pp. 100–120). Routledge.

Zhang, L., & Lo, Y. Y. (2021). EMI teachers' use of interactive metadiscourse in lecture organisation and knowledge construction. In D. Lasagabaster, & A. Doiz (Eds.), *Language use in english-medium instruction at university: International perspectives on teacher practice* (pp. 56–79). Routledge.