



HAL
open science

Summarizing and storytelling in English-medium engineering education

Francis Picavet

► **To cite this version:**

Francis Picavet. Summarizing and storytelling in English-medium engineering education. Education. Université Grenoble Alpes [2020-..], 2024. English. NNT : 2024GRALH021 . tel-04918327

HAL Id: tel-04918327

<https://theses.hal.science/tel-04918327v1>

Submitted on 29 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : SHPT - Sciences de l'homme, du Politique et du Territoire

Spécialité : Sciences de l'Éducation

Unité de recherche : Laboratoire de Recherche sur les Apprentissages en Contexte

Synthèse et récit dans l'enseignement de matières d'ingénierie en anglais

Summarizing and storytelling in English-medium engineering education

Présentée par :

FRANCIS PICAUVET

Direction de thèse :

Erica DE VRIES

PROFESSEURE DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES

Directrice de thèse

Alice HENDERSON

PROFESSEURE DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES

Co-directrice de thèse

Rapporteurs :

Stéphanie ROUSSEL

PROFESSEURE DES UNIVERSITES, Université de Bordeaux

Alex BOULTON

PROFESSEUR DES UNIVERSITES, Université de Lorraine

Thèse soutenue publiquement le **17 octobre 2024**, devant le jury composé de :

Patricia MARZIN-JANVIER,

PROFESSEURE DES UNIVERSITES, Université de Brest - Bretagne Occidentale

Présidente

Erica DE VRIES,

PROFESSEURE DES UNIVERSITES, Université Grenoble Alpes

Directrice de thèse

Alice HENDERSON,

PROFESSEURE DES UNIVERSITES, Université Grenoble Alpes

Co-directrice de thèse

Stéphanie ROUSSEL,

PROFESSEURE DES UNIVERSITES, Université de Bordeaux

Rapporteuse

Alex BOULTON,

PROFESSEUR DES UNIVERSITES, Université de Lorraine

Rapporteur

Jean-François BOUJUT,

PROFESSEUR DES UNIVERSITES, Grenoble INP - UGA

Examineur

Patrick STUDER,

PROFESSEUR, Universität Zürich

Examineur



UNIVERSITÉ GRENOBLE ALPES

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : SHPT – Sciences de l'éducation

Présentée par

Francis PICALET

**Summarizing and storytelling in
English-medium engineering education**

*Synthèse et récit dans l'enseignement
de matières d'ingénierie en anglais*

Thèse soutenue publiquement le **jeudi 17 octobre 2024**,
devant le jury composé de :

Madame Erica de VRIES

Professeure des Universités, Université Grenoble Alpes, Directrice de thèse

Madame Alice HENDERSON

Professeure des Universités, Université Grenoble Alpes, Co-directrice de thèse

Madame Patricia MARZIN-JANVIER

Professeure des Universités, Université de Bretagne Occidentale, Présidente

Monsieur Jean-François BOUJUT

Professeur des Universités, Institut Polytechnique de Grenoble – UGA, Examineur

Monsieur Patrick STUDER

Professeur des Universités, Université des Sciences Appliquées de Zurich, Examineur

Madame Stéphanie ROUSSEL

Professeure des Universités, Université de Bordeaux, Rapporteur

Monsieur Alex BOULTON

Professeur des Universités, Université de Lorraine, Rapporteur

Perhaps the best test of a man's intelligence is his capacity for making a summary.
Lytton Strachey

*What unites people? Armies? Gold? Flags? — Stories!
There's nothing in the world more powerful than a good story.
Nothing can stop it. No enemy can defeat it.*
Tyrion Lannister (George R. R. Martin)

*A caller burn o' siller sheen
Ran cannily out ovr the green
And whan our gutcher's drouth had been to bide right sair
He loutit down and drank bedeen a dainty skair*
Robert Fergusson, Cauler Water

À Véronique, Éloïse, Océane, Nicolas, Rémy, Mathieu.

Acknowledgements

A doctoral thesis is never a single person's endeavor. First and foremost, one needs to find a host organization and a research supervisor, which is not always an easy task when one is undertaking research after many years outside university circles. Therefore, the first person I would like to thank is Prof. Pascal Pansu, director of the *Laboratoire de recherche sur les apprentissages en contexte* (LaRAC), the education science department of the Université Grenoble Alpes (UGA), for welcoming me; for giving credence to this thesis project in the very first place; for making it possible for me to take part in international conferences; most of all, for inviting me to submit the project to Prof. Erica de Vries, soon to become my all-wise thesis supervisor at the LaRAC for the next five years. Thus, thank you, Erica, for channeling me through my lyrical and so often inappropriate wanderings, each time patiently bringing me back on the straight and narrow path of rigorous science: it has been an honor to work under your supervision. My infinite gratitude also goes to Prof. Alice Henderson, who co-supervised this thesis and vouched for its linguistic accuracy and integrity. Thank you so much, Alice, for your eternal good humor, constant enthusiasm, warm encouragement, and these countless hours spent reviewing my scribble. I would also like to thank LaRAC's administrative staff, and in particular Sandrine Lescure and Géraldine Garcia, for their kind and efficient help on so many occasions.

My gratitude goes to the jury members—Profs. Patricia Marzin-Janvier, Stéphanie Roussel, Jean-François Boujut, Alex Boulton, and Patrick Studer—for accepting to evaluate my work and giving me the opportunity to discuss my research in depth.

My employer, Grenoble INP-UGA, receives my special thanks for supporting this project throughout and allowing me to work as a part-time student on this thesis since its inception, back in 2019. In particular, I would like to thank Prof. Mireille Jacomino (VP-CEVU, Student Life and Studies Board), Prof. Frédéric Dufour (*Conseil Scientifique*, Scientific Advisory Board). I have additional thoughts for my local administration at Grenoble INP-Ense3— Profs. Yves Maréchal and Delphine Riu (Directors), and Céline Aubert (Administrative Director)—for facilitating the practical side of things, as well as for my closest language colleagues—Rebecca Guy, Sophie Bélanger, and Christine Schmidt—who supported me in this foolish undertaking, while sharing some of my former administrative responsibilities between them; for this, I sincerely apologize.

I will not thank COVID-19 for costing me a whole year's work on oral interaction between lecturers and students, my original subject. However, as every cloud has a silver lining, this change of subject gave me the opportunity to turn to the fourteen engineering lecturers who so willingly agreed to take an active part in this study, and whose voice recordings constitute the gold mine that is the Corpus of French Engineering Lectures (CoFEL), the core of the present research. Unfortunately, their names I cannot mention here, for reasons of anonymity, but they will recognize themselves. My heartfelt thanks also go to the students of these lecturers, who welcomed me into the intimacy of their lecture theaters.

Life is full of chance happenings, and one has to seize an opportunity when it comes. Such is my meeting with Prof. Hilary Nesi of Coventry University, UK, on the occasion of a corpus linguistics symposium in Grenoble in 2019, when I was starting to build CoFEL. It was then that, over a cup of tea, Professor Nesi introduced me to their recently collected Engineering Lecture Corpus (ELC), as well as to her former doctoral student, Siân Alsop, who had completed her thesis on the ELC a couple of years before. Benefiting from Professor Nesi's start-up boost and from Siân's experience, in the early stages of my own thesis, was a such a methodological beacon and a time saver that I cannot thank them both enough.

My gratitude goes to Isabelle Rousset for her guidance with the ELAN annotation software and to Prof. Olivier Kraif for those numerous hours spent on corpus analysis methodology. My special thanks also go to the Integrating Content and Language in Higher Education association (ICLHE), for being such a think tank with rich, positive ideas and an inexhaustible source of inspiration for a better educational future, and to Verena Grau, from the University of Vienna, Austria, for managing our ICLHE Special Interest Group of international PhD students so gracefully.

Last, but most certainly not least, my loving thanks go to my three sons, Mathieu, Rémy, Nicolas, and my daughter-in-law Dr. Myriam Grillot, for their kind help, and above all to my wife Véronique who, with infinite patience, has been my anchor during these long, long thesis years.

Thank you all.

Abstract

This thesis reports on the study of summarizing and storytelling in an English-Medium Instruction (EMI) context, as found in an original corpus of French engineering lectures. The pragmatic function of summarizing is to structure a lecture with episodes of summation and prediction, called summaries, whereas storytelling's function is to favor the appropriation of knowledge. Engineering lecturers may use stories to refer to professional issues, the evolution of techniques or technology, famous inventors and historical landmarks, or amusing incidents and serious accidents. Engineering in France is mainly driven by graduate schools of engineering, the 'Grandes Écoles'. Students have good reasons to enroll in engineering programs: these programs are geared towards scientific innovation, and are open to the world of business, research and international relations. Internationalization is at the heart of the strategy of most European higher education institutions. Although Europe promotes multilingualism and multiculturalism, the Bologna Process (1999), through EMI and despite encouraging student and staff mobility, has largely contributed to the sole spread of Englishization, particularly at Master's level. EMI focuses on content: language learning is neither planned nor assessed, even though content is delivered in English. In such contexts, English becomes the vehicular language of communication and may diverge from the L1-English norm in its phonology, syntax, or lexis; researchers refer to it as English as a Lingua Franca (ELF). The main research questions addressed are: How can summaries and stories be identified in EMI lectures? How can they be categorized? What is the amount of summarizing activity in lectures? For these purposes, a 28-hour corpus of EMI lectures was created: the Corpus of French Engineering Lectures (CoFEL). The corpus was built to study English-medium engineering education in a French context. Fourteen lecturers of mixed nationalities volunteered to participate over the 2019-2021 period. For lecturers and students alike, English was an additional language. To structure their lectures, lecturers use summarizing. Beyond summation and prediction, summaries further fall into four subcategories: reviews of previous, and of current lecture content; previews of current, and of future lecture content. First, an analysis grid based on this typology was used to identify summaries, and three independent raters identified 104 summaries. Next, a quantitative data analysis was carried out for the overall amount of summarizing activity; whether summary duration varied according to type; whether summary types depended on location. Finally, non-quantitative data was investigated to examine pragmatic and linguistic variation. Results show that all the lecturers use summarizing, but in different places throughout the lecture. With regard to storytelling, a typology of stories

was established, and a framework designed for their identification, categorization and description. Eighteen stories were found in the corpus, falling into six types: Narrative, Recount, Exemplum, Anecdote, Scenario, and Story-Like. The results suggest that stories help lecturers make sense of their own experience or that of their communities, and they allow the lecturer and the student audience to break away from routine while conveying vital information. To conclude, this thesis highlights the presence of summaries and stories in EMI-ELF lectures and pleads for their full pedagogical potential to be exploited. This could be achieved by improving the integration of content and language. If institutions enhanced their EMI teacher training by taking these findings into consideration, both lecturers and students could benefit.

Key words: English-Medium Instruction (EMI); English as a Lingua Franca (ELF); engineering lectures; summarizing; storytelling; corpus

Résumé

Cette thèse porte sur l'étude de la synthèse et du récit dans l'Enseignement de matières en anglais (EMA), à partir d'un corpus inédit de cours d'ingénierie français. La fonction pragmatique de la synthèse est de structurer un cours avec des épisodes de récapitulation et de prédiction, appelés synthèses, alors que la fonction du récit est de favoriser l'appropriation des connaissances. Les professeurs d'ingénierie peuvent utiliser des récits pour évoquer l'évolution des techniques ou de la technologie, des incidents drôles ou des accidents graves. En France, l'ingénierie est principalement portée par les Grandes écoles d'ingénieurs. Les formations sont très demandées : elles sont orientées vers l'innovation scientifique et ouvertes sur le monde de l'entreprise, de la recherche et des relations internationales. L'internationalisation est au cœur de la stratégie de l'enseignement supérieur européen. L'Europe promeut le multilinguisme et le multiculturalisme, cependant le Processus de Bologne (1999) a largement contribué à l'anglicisation des formations, surtout au niveau master. L'EMA se concentre sur le contenu ; l'apprentissage de la langue n'est ni planifié, ni évalué. L'anglais devient la langue véhiculaire et peut différer de la norme de l'anglais L1 dans sa phonologie, sa syntaxe ou son lexique ; les chercheurs le nomment « Anglais comme lingua franca » (ALF). Les principales questions de recherche abordées ici sont : Comment identifier les synthèses et les récits dans les cours d'EMA ? Comment les catégoriser ? Quelle est l'importance de l'activité de synthèse dans les cours ? Pour y répondre, un corpus de 28 heures de cours d'EMA a été créé : le Corpus de cours d'ingénierie français (CoFEL). Quatorze enseignants de diverses nationalités se sont portés volontaires pour participer sur la période 2019-2021. Pour les enseignants, comme pour leurs étudiants, l'anglais était une langue étrangère. Par la synthèse, les enseignants structurent leurs cours. Au-delà de la récapitulation et de la prédiction, les synthèses se répartissent en quatre sous-catégories : des revues du contenu des cours précédents ; des revues du contenu de la session en cours ; des aperçus du contenu de la session en cours ; des aperçus du contenu des cours futurs. La réalisation d'une grille d'analyse a permis l'identification de synthèses dans CoFEL et trois évaluateurs indépendants en ont trouvé plus de cent. Ensuite, une analyse quantitative des données a été effectuée : sur le volume global de l'activité de synthèse, pour déterminer si la durée de la synthèse variait en fonction du type, et si les types de synthèse dépendaient de leur position dans le cours. Enfin, des données non-quantitatives ont été étudiées pour leurs variations pragmatiques et linguistiques. Les résultats montrent que tous les enseignants ont recours à la synthèse, mais à différents endroits du cours. Pour le récit, une typologie a été établie et une grille d'analyse a été conçue pour l'identification des récits, leur

catégorisation et leur description. Dix-huit récits ont été trouvés, répartis en six types : Narratif, Descriptif, Exemplum, Anecdote, Scénario, Quasi-Histoire. Les résultats suggèrent que les récits aident les enseignants à donner un sens à leur propre expérience ou à celle de leur communauté et qu'ils permettent aux enseignants et à leurs étudiants de rompre la monotonie tout en transmettant des informations essentielles. Pour conclure, cette thèse met en évidence la présence de synthèses et de récits dans les cours d'EMA et plaide pour que leur potentiel pédagogique soit mieux exploité. Cela pourrait se faire en améliorant l'intégration de la langue et du contenu. Si les institutions tenaient compte de ces résultats dans la formation de leurs enseignants en EMA, cela pourrait bénéficier aux enseignants, ainsi qu'aux étudiants.

Mots-clefs : enseignement de matières en anglais (EMA) ; anglais comme lingua franca (ALF) ; cours d'ingénierie ; synthèse ; récit ; corpus

Table of Contents

Acknowledgements	1
Abstract.....	3
Résumé	5
List of Tables.....	10
List of Figures	11
List of abbreviations.....	13
1 Introduction	15
1.1 Overview of the research topic	15
1.2 Thesis overview	19
2 English-medium instruction (EMI).....	21
2.1 Distinguishing EMI from CLIL.....	21
2.2 EMI, internationalization and Englishization.....	24
2.3 EMI in France.....	27
2.4 English as a lingua franca.....	34
2.4.1 <i>Framing and naming plural English-es</i>	35
2.4.2 <i>Engineering lecturing in English as a lingua franca</i>	36
2.4.3 <i>Changing the paradigm</i>	36
2.4.4 <i>Instructors as communicators</i>	37
3 Corpus construction	39
3.1 Setting.....	39
3.2 Participants	40
3.2.1 <i>Participating lecturers</i>	40
3.2.2 <i>MEng student audiences</i>	42

3.3	Data collection.....	46
3.4	Corpus analysis tools.....	47
3.4.1	<i>The ELAN software</i>	47
3.4.2	<i>Hesitation markers</i>	52
3.4.3	<i>Sketch Engine</i>	53
3.5	CoFEL: corpus description.....	55
4	Summarizing.....	57
4.1	Conceptual and research background.....	58
4.1.1	<i>Deixis</i>	59
4.1.2	<i>Discourse markers</i>	61
4.1.3	<i>Information packaging and enumeration</i>	62
4.1.4	<i>Salient features of summative summarizing</i>	62
4.1.5	<i>Salient features of predictive summarizing</i>	63
4.1.6	<i>Scaffolding the lecture: from micro to macro structures</i>	65
4.2	Research questions.....	66
4.3	Method.....	67
4.3.1	<i>Building an analysis grid</i>	67
4.3.2	<i>Identifying candidate summaries</i>	70
4.3.3	<i>An example: identifying RevPREs</i>	71
4.3.4	<i>Categorization of summaries</i>	74
4.3.5	<i>Summary data visualization</i>	74
4.4	Results: quantitative summarizing data in CoFEL.....	80
4.4.1	<i>Summarizing activity</i>	80
4.4.2	<i>Type of summarizing activity in lectures</i>	82
4.4.3	<i>Token length and duration of episodes as a function of summarizing type</i>	83
4.4.4	<i>Summarizing type as a function of location in the lecture</i>	84
4.5	Pragmatic and linguistic variation in CoFEL summaries.....	86
4.5.1	<i>A trait of lecture beginnings: RevPRE</i>	86
4.5.2	<i>Another trait of lecture beginnings: PreCUR</i>	91
4.5.3	<i>A trait of lecture midsections: RevCUR</i>	95
4.5.4	<i>A trait of lecture endings: PreFUT</i>	99
4.5.5	<i>Chaining of summaries</i>	104
4.6	Summarizing in CoFEL: conclusion and perspectives.....	108
5	Storytelling.....	111
5.1	EMI, engineering lectures and storytelling.....	111

5.2	Defining the terms ‘story’ and ‘storytelling’	112
5.3	Research questions	112
5.4	Method.....	113
	5.4.1 <i>Building an analytical framework for studying storytelling</i>	113
	5.4.2 <i>Identifying stories in CoFEL</i>	114
	5.4.3 <i>Categorizing stories in CoFEL</i>	116
5.5	Exemplification of the typology: stories found in CoFEL	118
	5.5.1 <i>Narratives – The Drôme Farmhouse</i>	118
	5.5.2 <i>Anecdotes – The Untimely Airbag</i>	122
	5.5.3 <i>Exempla – The Malpasset Dam Disaster</i>	123
	5.5.4 <i>Recounts – The Seven Bridges of Königsberg</i>	125
	5.5.5 <i>Story-likes – The Bubbles (a cable-car)</i>	126
	5.5.6 <i>Scenarios – The behavior of unsaturated soil & electromagnetic interference</i>	129
5.6	Storytelling in CoFEL: conclusion and perspectives	131
6	Discussion	134
	6.1 Main findings.....	134
	6.2 Corpus patterns, relationships, and features	136
	6.3 Implications	137
	6.4 Limitations of the study	140
	6.5 Further research and conclusion	140
	References	143
	Appendices	169

List of Tables

Table 1	A paradigm shift from English language classes to EMI lectures.....	37
Table 2	Graduate school of engineering: list of the 12 participating lecturers.....	41
Table 3	Department of energy, civil and environmental engineering: countries of origin of MEng students.....	46
Table 4	Key to abbreviations and codes used in CoFEL excerpts and transcripts.....	49
Table 5	Fourteen engineering lectures of CoFEL, in chronological order of recording.....	56
Table 6	Analysis grid for retrieving candidate summaries from CoFEL, according to the 4 summary types.....	69
Table 7	Examples of a 1-gram KWIC concordance extraction from three CoFEL engineering lectures.....	72
Table 8	Examples of a 2-gram KWIC concordance extraction from three CoFEL engineering lectures.....	73
Table 9	Examples of a KWIC concordance extraction (3- to 4-grams) from four CoFEL engineering lectures.....	73
Table 10	Summarizing activity and summarizing episodes in CoFEL lectures.....	81
Table 11	Type of summarizing activity in CoFEL lectures.....	82
Table 12	Length of the summarizing episodes expressed in tokens and in seconds.....	83
Table 13	Summarizing type as a function of the location of summaries in lectures (number of episodes).....	86
Table 14	CoFEL: list of summary chains.....	104
Table 15	Selected search terms to identify potential story passages in CoFEL.....	115
Table 16	Selected search terms relating to incidents or accidents in engineering.....	115
Table 17	Selected search terms relating to tenses other than the Present.....	115

List of Figures

Figure 1	The language-content continuum of learning outcomes.....	22
Figure 2	Language & content: the steps from no integration (Non-CLIL) to full integration (CLIL)	23
Figure 3	Department of energy, civil and environmental engineering: origin of MEng students (2013-2022)	44
Figure 4	Department of energy, civil and environmental engineering: student headcount trend (2013-2022)	45
Figure 5	ELAN: typical workspace window (pre-COVID recordings).....	48
Figure 6	Screenshot of CoFEL tier dependencies (annotation grid) in ELAN.....	51
Figure 7	Screenshot of a Sketch Engine ‘word sketch’ example from CoFEL, with the lemma “power” as a search entry	54
Figure 8	Sketch Engine: screenshot of an example 3-gram concordance search result in CoFEL	55
Figure 9	Breakdown of summarizing into two main functions and four types.....	58
Figure 10	The contract established between the lecturer and the students when summarizing	59
Figure 11	The 3-stage process to identify and delineate the four summary types in CoFEL	71
Figure 12	Example of an Annotation Density Plot (ADP) in the ELAN software.....	76
Figure 13	Three examples of SumVIZ single-summary graphs	77
Figure 14	Examples of two multiple-summary display possibilities for the same lecture	78
Figure 15	SumVIZ: example of a fine-grain summary distribution display per type.....	78
Figure 16	SumVIZ: overview of the summarizing activity in CoFEL	79
Figure 17	CoFEL - Summarizing overview by type: whole lectures with the three-part boundaries.....	85
Figure 18	Summarizing in CoFEL: beginning part of the 14 lectures.....	87
Figure 19	Position and duration of the shortest RevPRE summary in CoFEL.....	88
Figure 20	Position and duration of a RevPRE summary in initial position in CoFEL	89
Figure 21	Position and duration of the longest summary in CoFEL (RevPRE type)	91
Figure 22	Position and duration of a short PreCUR summary in CoFEL.....	92
Figure 23	Position and duration of the longest PreCUR summary in CoFEL.....	93

Figure 24	Summarizing in CoFEL: midsection of the 14 lectures	96
Figure 25	Position and duration of the shortest summary in CoFEL (RevCUR type)	97
Figure 26	Position and duration of the 2 nd longest summary in CoFEL (RevCUR type)	98
Figure 27	Summarizing in CoFEL: ending part of the 14 lectures	100
Figure 28	Position and duration of a PreFUT summary: ending of SRM lecture L02 ..	101
Figure 29	Position and duration of a PreFUT summary: ending of lecture L04, Soil & Rock Mechanics	102
Figure 30	The six story types and their features	117

List of abbreviations

CEFRL	Common European Framework of Reference for Languages
CoFEL	Corpus of French Engineering Lectures
DM	Discourse Marker
EAL	English as an Additional Language
EAP	English for Academic Purposes
ELF	English as a Lingua Franca
EME	English-Medium Education
EMEMUS	English-Medium Education in Multilingual University Settings
EMI	English-Medium Instruction
ESP	English of Specific Purposes
HE	Higher Education
SE	Standard English

1 Introduction

This chapter presents the central theme of this thesis: the pragmatic functions of summarizing and storytelling as used by lecturers, teaching future engineers, and teaching in English—as an additional language—to international student audiences. This chapter indicates the scope of the situation and the purpose and design of the thesis.

1.1 Overview of the research topic

According to the French ministerial agency for information on education and professions¹, whose main purpose is to support young people in building their career plans, students have good reasons to enroll in graduate schools of engineering: their programs are geared towards scientific innovation, and are open to the world of business, research and international relations. Engineering as an academic discipline offers profession-oriented training within a research environment, with optimal professional integration and high salaries, and a guarantee to move up the career ladder easily. A number of projects throughout the engineering curriculum introduce students to real-life issues, project management and teamwork. Twenty per cent of courses taught by practicing professionals and a minimum of six months of mandatory internships provide an opportunity to apply the knowledge acquired in the lecture hall and to refine one's career plan. All engineering schools place their students in a research environment linked to their laboratories, some of which are partnered with universities, public research institutes or private companies. Engineering graduates have access to a wide range of business sectors and employment opportunities, as they can join any type of company, from small businesses to large multinationals. In France, having an engineering degree almost guarantees employment: one year after graduation, nine out of ten engineers looking for a job have found one². Additionally, the international dimension of engineering training has developed rapidly over the last twenty years. Engineering schools prepare their students for an international professional environment based on the globalization of trade, multiple locations of production, and English-speaking research groups. To this end, English was made compulsory in all French engineering schools in 2006, and students have to validate a minimum level of proficiency with an external certification in order to obtain their degree (CTI, 2006).

¹ *Office national d'information sur les enseignements et les professions (ONISEP)* [Accessed May 15, 2024]. <https://www.onisep.fr>

²Source: ONISEP [Accessed May 15, 2024] <https://www.onisep.fr>

Today, an engineering degree course necessarily includes an experience abroad, either as an intern or as a student. Opening courses to an increasing number of foreign students and involving professors from partner universities is meant to promote a cultural mix. However, although engineering lecturers around the world may use a common language to deliver the same kind of courses on the same subjects, engineering lecturing remains context-, culture-, and lecturer-specific (Alsop, 2015).

Internationalization is at the heart of the strategy of most graduate schools of engineering today, in France³, as well as in the rest of the world. Driving this internationalization process are the worlds of business, finance, and politics, as well as those of science and engineering. To meet the demand for well-trained and competent international field actors, universities have made international development one of their major strategic priorities. In Europe, to promote the mobility of staff and students, the employability of graduates, and the European presence in higher education, an initiative was launched to create the European Higher Education Area in the early 2000s (Curaj et al., 2020). Seeking to improve their international influence and visibility, universities have increasingly anchored their training activities in strong international institutional partnerships and collaborative degree programs. For example, most current curricula have an international dimension, like spending a semester or a year studying abroad, or graduating with a double degree, such as the European Erasmus Mundus Joint Master's⁴ degree. To improve their position in international rankings⁵, to build their reputation, and to increase their global visibility, institutions have had to adapt their approaches to international student recruitment and strategic partnerships. However, some voices have warned about the numerous caveats of these rankings and their side-effects (European University Association, 2023). The fact that the English language plays a dominant role in a globalizing environment, for better or for worse, is indisputable, and one adaptation of importance that has taken place is the worldwide Englishization of curricula, often to the detriment of national languages and plurilingualism. This has led to the spread of English-medium instruction (EMI). France, one of the early signatories of the Bologna Process (European Commission, 1999), is no stranger to the EMI phenomenon, and French higher education institutions have followed suit, after a rather slow start.

³ C.f., the international web pages of the *Commission des Titres d'Ingénieur* (CTI) and the *Conférence des Grandes Écoles* (CGE) [retrieved May 16, 2024] (CTI, n.d.; CGE, n.d.).

⁴ <https://erasmus-plus.ec.europa.eu/opportunities/opportunities-for-individuals/students/erasmus-mundus-joint-masters>

⁵ E.g., Times Higher Education's World University Rankings, QS World University Rankings, Shanghai Ranking, etc.

In the French higher education landscape, one finds graduate schools of engineering, the *Grandes Écoles d'ingénieur-e-s*, which are elite educational institutions. They, too, have also thrown themselves wholeheartedly into the enrollment of international students and the Englishization of their courses, often as leaders in the process, as the international dimension of engineering is a vital issue (Borri et al., 2007). The combined phenomena of internationalization and Englishization are most evident at Master's level, as these courses develop in parallel with the traditional *diplôme d'ingénieur-e* courses. The majority of local students, who have done their entire non-linguistic training in French up to Bachelor's level, suddenly discover EMI at Master's level: they find themselves in mixed groups of local and international students, and most, if not all, of their remaining two-year curriculum is taught in English. With their lecturers heavily involved in research and international communication, institutions generally assumed, overtly or covertly, that the switch from French to English for course content, and from local to international audiences, could be made almost overnight. However, the reality of EMI on the ground has proved to be different and a number of issues have rapidly arisen, like the question of faculty and student language proficiency, the linguistic and cultural aspects of pedagogy, or the integration of local and multicultural students, among others. Some of these issues have been addressed by stakeholders, others overlooked or ignored.

Thus, English has become the dominant medium of delivery of engineering content in graduate schools of engineering, at Master's level, in France. But which English? Although some native English speakers may be found among the French engineering schools' resident faculty, these are far more likely to be language teachers than content lecturers⁶. Students' exposure to native speakers in their courses (outside their language classes) may take place through some costly, short-term guest lecturers from English speaking countries, or, more likely, through videoconferencing platforms or filmed lectures as part of MOOCs⁷. Even then, there is a good chance that many of them are non-native speakers of English working in a British or American university. In reality, engineering students will be mostly exposed to local lecturers who speak English as an additional language. To describe the language used in a communication situation when neither of the speakers share the same first language and English is their only option, we will use the term "English as a lingua franca" (Seidlhofer, 2013). Over

⁶ There do not seem to be any official statistics available on the number of native English-speaking engineering lecturers in France. Therefore, this claim is based on my personal experience of teaching in French engineering schools for 30 years.

⁷ Massive Open Online Courses.

the years, the term has supplanted “English as an international language”, or “Global English⁸”, especially when it is used to refer to a vehicular language in which non-native speakers take an active role in its development (Modiano, 1999). English as a lingua franca (ELF) has become the main medium of communication in the EMI context, a context in which the intention to teach and learn the English language is absent, as it is neither planned for nor assessed (Brown & Bradford, 2017).

In writing this thesis, my research objective has been to study two pragmatic functions of EMI engineering lectures: summarizing and storytelling (Alsop, 2015; Nesi et al., 2014). Summarizing can be defined as the episodes that lecturers use to signpost what they have done, or will be doing in their lectures; or to signal or recapitulate a change of section or topic that has taken place, or will take place. Storytelling happens when lecturers tell stories to their students to illustrate or highlight a point of the lecture, with real-life or imaginary examples, from their own experience or that of others. A typology of both summaries and stories is elaborated, and a precise inventory of the occurrences of the two pragmatic functions found in the Corpus of French Engineering Lectures, that I built for this purpose, is given (CoFEL, Picavet, 2022). A selection of examples from CoFEL illustrates and concludes the analyses of both pragmatic functions.

With the steady rise of EMI in the past twenty years, some voices proposed the idea that the integration of content and language be improved in higher education⁹. It has been argued that separating language and content denies students the opportunity to focus on language at the precise moment when their motivation to learn specific content may be strongest (Lightbown, 2014). Content and Language Integrated Learning (CLIL) has been widely encouraged for primary and secondary education at the European level since the 2000s (Eurydice, 2006), but such a policy at the tertiary level seems to be making a much slower start. In CLIL, the approach is dual-focused, and the additional language is used for the teaching and learning of both content and language (Frigols-Martin et al., 2011). Many advantages are claimed for such integration: it develops proficiency in both the content subject and the language in which this is taught (Eurydice, 2006); it offers the parallel development of professional competences and domain-relevant communication skills; students are exposed to authentic, goal-directed and pragmatically pertinent interaction with the language; it better

⁸ “Globish” is mostly derogatory, tending to infer a simplified English with less value (Henderson, 2021).

⁹ A leading voice is the Integrating Content and Language in Higher Education (ICLHE) association. <https://www.iclhe.org/>

prepares students for mobility and life in Europe and elsewhere (Tudor, 2009); it fosters interdisciplinary collaboration between content and language specialists (Lyster, 2017). Today, an evolution of the term English-Medium Instruction (one-way teaching) towards the more inclusive English-Medium Education (encompassing both teaching and learning) is increasingly advocated (Dafouz-Milne & Smit, 2016, 2020). A growing number of field actors and researchers argue for more interdisciplinary and holistic approaches to exploit the potentialities lying at the intersection of language and content, in the context of the internationalization of higher education (Studer & Smit, 2021). It is within this multi-faceted context of internationalization, Englishization, and English-medium instruction that this thesis was written.

In the graduate school of engineering where this study is set, and arguably in other such institutions, lecturers and students are under pressure to perform, to deliver and learn content effectively. If summarizing and storytelling can contribute to this effectiveness, it would be useful to explicitly recognize the importance of paying attention to both language and content, and to support this by providing training on how best to achieve such integration. Change may be imminent, and it might come from the lecturers themselves. When asked if they would agree to take part in this study, where I would observe of their lecturing, all fourteen engineering lecturers responded with an enthusiastic “yes”. This was a pleasant surprise, as it might express that they felt the current EMI situation could be improved, and that they were happy to contribute to changing it by participating in this study.

1.2 Thesis overview

This thesis comprises three main parts: an introduction to the context of the study (Chapter 2), a description of the collection and preparation of study data (Chapter 3), and a presentation of the analysis of this data, both the process and results (Chapters 4 & 5). Chapter 2 introduces the broad context in which this study is set: English-medium instruction (EMI). First, some terminological considerations are explained, distinguishing EMI from Content and language Integrated Learning (CLIL): although the two concepts have certain similarities in terms of the transmission of content knowledge via a foreign language, the former does so without regard to language, while the latter places content and language on an equal footing in terms of learning outcomes and assessment. Second, the concepts of internationalization and Englishization of higher education are examined, from an EMI perspective, before EMI in

France is focused on. Finally, the use of English as a lingua franca, the lecturers' medium of instruction and communication, is investigated.

Chapter 3 is devoted to the construction and description of the Corpus of French Engineering Lectures (CoFEL), the data source for this study. It opens with a description of the setting where data in the form of audiovisual recordings was collected. Then, the participants—the lecturers who volunteered to take an active part in this study—are presented, as well as their international student audiences. Afterwards, the conditions under which the data was collected are described. A rapid description of the tools that were used in the corpus transcription-annotation and management-analysis processes follows, namely, Max Planck Institute's ELAN software (ELAN, 2023) and Lexical Computing's Sketch Engine online platform (Kilgarriff et al., 2014). To close, an overview is provided of CoFEL and its contents.

Chapters 4 and 5 represent the core of this thesis, with the in-depth analysis of two pragmatic functions found in CoFEL, respectively: summarizing and storytelling. Chapter 4 opens with the conceptual and research background on which the analysis of the summarizing pragmatic function is built, with such concepts as deixis, discourse markers, and information packaging. Summarizing is divided into two sub-functions, summative and predictive, and some of the key features of each are examined; given the scaffolding role of summaries, the structure of lectures will be examined. Then, the research questions are listed, and the analysis methodology is presented: how the analysis grid was built (to identify summaries in the corpus), and an example of how it was used. The summary categorization process is described, as well as a summary visualization tool. Quantitative results then give a comprehensive overview of the amount of summarizing activity that can be found in the engineering lectures studied, as well as information on summary duration, distribution, or frequency. The chapter closes with a qualitative investigation of the pragmatic and linguistic variation in CoFEL summaries.

Chapter 5 deals with the storytelling pragmatic function and follows a similar but not identical pattern. After a presentation of the theoretical framework based on the concepts of story and storytelling in engineering lectures, research questions are given. The method used to identify and categorize stories according to the story typology is then presented, followed by an in-depth analysis of one example of each of the six story types.

Finally, Chapter 6 examines the main findings of this thesis, and their theoretical and practical implications are addressed.

2 English-medium instruction (EMI)

In the wake of the Bologna Process of the late 1990s and the creation of the European Higher Education Area (EHEA, European Commission, 1999), more and more institutions at tertiary level around the world have adopted courses and whole curricula taught in English, moving from a national to a foreign language for content delivery, usually English (Costa & Mastellotto, 2022). Such teaching is commonly referred to as English-medium instruction (EMI) and is done primarily by lecturers for whom English is an additional language (EAL). Researchers have tracked the spread of EMI in tertiary education across Europe, in countries where English is not the national language (Wächter & Maiworm, 2002, 2008, 2014)¹⁰. Their results show that from 725 such programs in 2002, numbers rose to 2,389 in 2008, and to 8,089 in 2014 (+1116% compared to 2002). EMI is found predominantly at Master's level and in the fields of sciences, business, and engineering, with other subjects (humanities, health, arts, etc.) lagging far behind (Alsop, 2015; Wächter & Maiworm, 2014). In a world in which native speakers of English are outnumbered by EAL speakers (Crystal, 2003, 2018; Ethnologue, 2023), more of the latter have experience of EMI, as students, lecturers, or other university staff.

2.1 Distinguishing EMI from CLIL

Our context is higher education (HE), so it is important to distinguish between the two acronyms most pertinent to our research context: CLIL¹¹ and EMI. The idea of Content and Language Integrated Learning (CLIL), has its origins in the 1980s, but the term was coined in the mid-1990s. According to the Council of Europe, probably its greatest promoter, CLIL is “a dual-focused educational approach in which an additional language is used for the learning and teaching of content and language” whose objective is to promote “both content and language mastery to pre-defined levels” (Frigols-Martin et al., 2011, p. 11).

A great variety of CLIL practices across institutions, teachers, and levels (Dalton-Puffer et al., 2010; Mehisto et al., 2008). However, they can be categorized into two major types of bilingual or immersion education, according to the status of the languages used to teach the subjects. “CLIL type A” designates “provision where different subjects are taught in a language

¹⁰ In their studies, Wächter & Maiworm refer to ‘English-taught programs’ (ETPs) in lieu of EMI.

¹¹ The French translation of CLIL is *EMILE (Enseignement d'une matière par l'intégration d'une langue étrangère)*.

designated in the top-level curriculum as a foreign language”; “CLIL type B” designates “provision where different subjects are taught in a regional or minority language” (Birch et al., 2023, pp. 147–148). Thus, it is CLIL type A that is relevant to this study, as an approach to learning content through English as an additional language. The dominant language in CLIL is English (Dalton-Puffer & Smit, 2013), reflecting the widespread notion that mastery of English is seen as an essential feature of literacy worldwide.

The degree to which language is explicitly considered as an objective for a content class is of crucial importance. Some researchers propose to look at the language-content degree relation in HE as a continuum, as shown in the graph in Figure 1 (Airey, 2016). This continuum places CLIL in the middle, where equal attention is paid to language and content learning outcomes. On the far left of the diagram are English for Academic Purposes courses, where only language is taught, with mainly language learning outcomes; the main objective is to train students in the knowledge and skills of a language appropriate for study in an HE institution. On the far right is EMI, where learning outcomes only focus on the acquisition of content: simplistically, English is just a pragmatic tool for delivering content. It is assumed that the language poses no problems for teachers or students and that it does not affect the content of teaching or learning. Although useful as a starting point, Airey finds the division artificial, similarly to Halliday and Martin (1993) who argue that content and language cannot be separated in such a way; language is actively engaged in bringing conceptual structures into being, thus content and language are intertwined.

Figure 1

The language-content continuum of learning outcomes



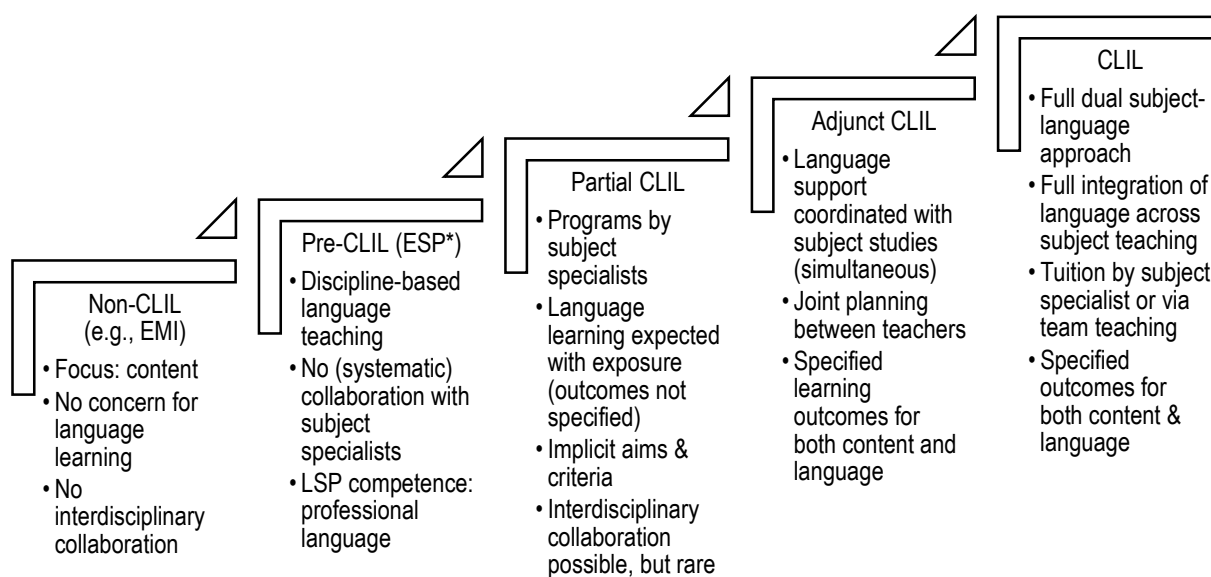
Note. Adapted from Airey (2016, pp. 72–73). *EAP: English for Academic Purposes.

Concerning the EMI-CLIL continuum in higher education, Räsänen (2011) proposed a more nuanced view of the various degrees of learning outcomes, criteria and interdisciplinary collaboration, as shown in Figure 2. Reading the graph from left to right, the case of EMI is straightforward: the focus is only on content, delivered by L1 or EAL lecturers, and language learning is not planned. Pre-CLIL practices are much more common than semi or fully integrated approaches: content courses are delivered by either L1 or EAL lecturers and the

dominant practices are English for academic purposes or English for specific purposes (EAP and ESP respectively), with rare interdisciplinary collaboration between content and language specialists. An example of an ESP course would be ‘English for engineers’, given to students to help them learn the language of their specialization. In the case of Partial CLIL, the focus is not usually on language improvement, but there may be an awareness of the role of linguistic accuracy in effective communication; some language support may be offered to students; some interdisciplinary coordination may exist, but learning outcomes remain assessed separately. Adjunct CLIL involves some coordination between language and content specialists, such as joint planning or team teaching; learning outcomes and assessment criteria are specified for both language and content. Finally, CLIL is entirely focused on the duality of the subject-language approach: there is full integration of language into content instruction; tuition is delivered by the content teacher or through team teaching; content and language learning outcomes are specified.

Figure 2

Language & content: the steps from no integration (Non-CLIL) to full integration (CLIL)



Note. Adapted from Räsänen (2011). *ESP: English for Specific Purposes.

It has been argued that true CLIL is an authentically communicative activity, as it creates a situation of genuine need and desire to communicate through English as an additional language (Tudor, 2009). Students are engaged in a variety of communicative tasks which have a clear pragmatic objective: assimilate the knowledge and the related language skills related to the subject studied. A full range of communicative activities is then de facto involved, comprising active listening, interacting, or writing.

Looking at the present and future of CLIL in higher education, recent CLIL developments promote the inclusion of all levels of education (primary, secondary, tertiary, vocational and adult) as “the overall major challenge”, in its “integrative nature”, in the “development and implementation of a teacher education curriculum” (Frigols-Martin et al., 2011, p. 5). Early CLIL research has focused more on primary and secondary education than on tertiary education, our subject of choice, but many researchers argue for, and work for a greater integration of content and language at tertiary level as well (e.g., Doiz et al., 2019; Gabillon, 2020; Lasagabaster & Doiz, 2021; Pérez Cañado, 2021; Räsänen, 2011; Vilkanienė, 2011). In this context, the term increasingly used is ICLHE: integrating content and language in higher education (Smit & Dafouz-Milne, 2012). The term was introduced to reflect the different interests of compulsory education and higher education establishments.

The definition of EMI is protean (e.g., Airey, 2016; Lasagabaster & Sierra, 2010; Smit, 2023), but we will adopt the following, as it is one of the most complete and corresponds to the approach of this study: “The use of the English language to teach academic subjects (other than English itself) in countries or jurisdictions where the first language (L1) of the majority of the population is not English.” (Dearden, 2014, p. 2; Macaro et al., 2018, p. 37). Most importantly, language teaching and learning is neither planned for nor assessed (Brown & Bradford, 2017), which is clearly different to CLIL, and students are expected to master the language skills required to deal with the subject before starting their studies (Graddol, 2006).

2.2 EMI, internationalization and Englishization

EMI exists within the worldwide processes of internationalization and Englishization¹² (e.g., Bowles & Murphy, 2020; de Wit et al., 2015; Dimova et al., 2015; Galloway et al., 2017; Lasagabaster, 2022a; Richards & Pun, 2023; Robson & Wihlborg, 2019; Wilkinson & Gabriëls, 2021, among the most recent). One of the most advanced definitions of internationalization is as follows:

¹² The term “glocalization” can also be found in the literature. It could come from a Japanese business word meaning “global localization” (Robertson, 1992, pp. 173–174).

the intentional process of integrating an international, intercultural or global dimension into the purpose, functions and delivery of post-secondary education, in order to enhance the quality of education and research for all students and staff, and to make a meaningful contribution to society. (de Wit et al., 2015, p. 281)

Tertiary education is where EMI has spread the most (Wächter & Maiworm, 2014) for two main reasons. First, EMI is seen as one way to increase foreign student enrolment and attract funds¹³, while potentially improving the international-intercultural skills of the non-mobile domestic students. The latter aspect has been referred to as ‘internationalization at home’ (Almeida et al., 2019; Beelen & Jones, 2015; Robson et al., 2017; Robson & Wihlborg, 2019), or ‘curricular internationalization’ (Wächter, 2018). At European Union’s level, this focus on domestic students is integrated into the Erasmus +¹⁴ program called ATIAH, for Approaches and Tools for Internationalization at Home (Almeida, 2018). Second, EMI can improve a university’s standing in international rankings, which can also have financial benefits (Baltaru et al., 2022; Benito et al., 2019; Marginson & van der Wende, 2011).

Since the late 2010s, the development of internationalization in universities has been called into question because of the dilemma it often provokes: on the one hand, the need to prepare graduates for interconnected societies and systems with shifting borders with an international vision of the world; on the other, the need to take account of national cultural priorities (Bowles & Murphy, 2020). In some countries, the status of national languages can be seriously challenged by the use of English (Gabriëls & Wilkinson, 2020). At all levels of education, the omnipresence and status of English can mean unequal access to university programs, resulting from – and perpetuating – an uncritical attitude towards the quality of the educational experience (Lanvers & Hultgren, 2018). The Bologna process negotiators have thus been criticized for failing to foresee the dominance of the English language in its founding principles, with internationalization too rapidly becoming synonymous with Englishization (via English-medium higher education), and truly bilingual degrees or practices which promote multilingualism have often been set aside (Kirkpatrick, 2011; Phillipson, 2006, 2008). However, counter examples can be found, like the University of Bergen language policy, in Norway (UiB Board, 2019), or the University of Jyväskylä, in Finland (JyU Board, 2015). The University of Jyväskylä “promotes active and systematic internationalisation of its students and

¹³ Non-EU students pay more for a master’s program than EU students : € 243 (US\$ 264) for EU (unchanged since 2019), and € 3,770 (US\$ 4,093) for non-EU, since 2019 (Campus France, 2023).

¹⁴ Erasmus: EuRoPean Action Scheme for the Mobility of University Students. Originally for higher education, the “+” was added in 2014, creating a single overarching program for education, training, youth and sport. <https://erasmus-plus.ec.europa.eu/about-erasmus/history-funding-and-future>

staff” (2015, p. 1), and Finland has even made the internationalization of English-language universities a national policy, with Finnish and English considered “self-evident primary languages of the universities” (Saarinen & Rontu, 2018, p. 97).

Overall, there seems to be a broad consensus that EMI has been driven primarily by market forces and the commodification of the products of research and higher education rather than educational ones (Lasagabaster, 2022a; Le Lièvre, 2021). Many researchers, language practitioners and proponents of national language and culture therefore see the worldwide growth of EMI as much as a threat as an opportunity (Block & Khan, 2021; Chapple, 2015; Hugonnier et al., 2020; Macaro et al., 2018). In the scope of this PhD, one of the most commonly cited pedagogical shortcomings of EMI is that users of English as an additional language cannot speak with as much nuance as in their first language, especially in high-stakes subjects (Kirkpatrick, 2011; Vinke, 1995; Wilkinson & Gabriëls, 2019). Two other limitations are a lack of necessary language skills in students and the idea that students may suffer from higher cognitive load when trying to master field content (Breetvelt, 2018; Roussel et al., 2017; Truchot, 2010).

Since the 2010s EMI has been developing towards a more positive and qualitative concept, for example:

- a CLIL-like orientation in tertiary education (LanQua, 2010; Räsänen, 2011), including attempts to raise faculty awareness of disciplinary literacy goals and to require language-learning outcomes in curricula (Airey et al., 2017);
- the implementation of student-centered teaching methodology, tailored to the specific context, that compensates for the additional cognitive load imposed on both learners and teachers (Airey, 2012; Lasagabaster & Doiz, 2021; Wilkinson, 2013);
- the fostering of interdisciplinary collaboration and team-teaching, with language teachers working alongside content lecturers (Lasagabaster, 2018, 2022a);
- encouragement of adequate, critical institutional planning and the professional development of lecturers (Macaro et al., 2018).

One new perspective even posits that all EMI lecturers are language teachers, even in monolingual L1 settings, and even if most of them would not identify themselves as such (Airey, 2012, 2020; Block & Moncada-Comas, 2019; Lasagabaster, 2022a; Malmström & Pecorari, 2021). Finally, higher education institutions, both national and local, have a role to play: developing language awareness and a coherent language policy and guidelines will have an impact on the effectiveness of CLIL/ICLHE, as well as on the institution's image; appropriately funded pedagogical and administrative support structures are needed; quality assessment indicators must be established, combined with quality enhancement approaches (LanQua, 2010; Taillefer, 2013).

2.3 EMI in France

France is no stranger to the fast spread of EMI in tertiary education in Europe. From 33 programs in 2002, numbers rose to 499 in 2014, an increase of 1512% (Wächter & Maiworm, 2002, 2014), with France ranking fourth after the Netherlands (1,078), Germany (1,030) and Sweden (822). Wächter and Maiworm also found that, in the academic year 2013-2014, Nordic countries (Denmark, Finland, Iceland, Norway, Sweden) concentrated 61% of all European EMI programs at Bachelor's and Master's levels, representing 20% of all programs and 5% of student enrollment. At the other end of the spectrum, their Southwest Europe group of countries including France (with Italy, Portugal, and Spain) presented the lowest figures at the time, respectively: 17% of all European EMI programs, representing 2.8% of all programs and 0.5% of student enrollment. Updated figures for 2022 from the official Campus France agency¹⁵ announce a sharp increase in their EMI offerings since 2014, reaching 1,328 referenced programs in 2022 (Campus France, 2022). The number of programs reached 1,718 in 2023-2024 (652 in engineering and technology, or 38%), of which 1,435 taught exclusively in English (Campus France, 2024b). To attract an ever-increasing number of foreign students, including those who might be put off by a poor command of French, the invitation is clear:

¹⁵ Campus France is a national agency responsible for promoting French higher education abroad, managing French and foreign government grants and welcoming international students. It is jointly supervised by the *Ministère de l'Europe et des Affaires étrangères* (Ministry of Europe and Foreign Affairs) and the *Ministère de l'Enseignement supérieur et de la Recherche* (Ministry of Higher Education and Research). It was created in 2010. <https://www.campusfrance.org>

For students currently learning French **courses in English are a great introduction to the French language and culture**. These courses are a **great complement to trainings taught in French**, and are designed for a different audience, people who don't speak French well enough to study in French, but who **already caught the "I'm into France" virus**. (Campus France, 2022; original emphasis)

Thus, EMI is now part of the French higher education system and the standard for prestige courses at Master's level. However, it is also an example of how national language policy is being challenged, with CLIL or ICLHE struggling to be accepted and regular backlashes against an 'all English' tertiary education. Going back to 2004, Smith had already listed the difficulties that may arise when adopting EMI (Smith, 2004; in Coleman, 2006), and France is no exception. Among such accurately foreseen difficulties he listed:

- for both faculty and students, inadequate language proficiency, hence the need for training;
- unpreparedness of lecturers to teach through English;
- for local students, a lack of interest;
- for international students, cultural integration difficulties;
- assessment equity between local and international students (to which we can add the choice of language of assessment);
- for institutions, organizational or administrative (and linguistic) issues.

As with most matters concerning EMI, Nordic countries (comprising the Netherlands, in this case) took a head start in addressing those issues, engineering training included, leaving France behind, among others, with relevant questions, but often without pragmatic answers. For example, in 2001, Klaassen paved the way by addressing the subject of the relationship between effective lecturing behavior and language proficiency in EMI, in an engineering training context (Klaassen, 2001).

Related to France, some studies concerning the spread of EMI and its domestic consequences (examples below). However, they will also remark that France does not seem to have had much of a presence in international scientific literature and conferences on EMI (Taillefer, 2009, 2011). Moreover, there has been little discussion of the impact of the institutional initiatives on the development of EMI in the French context (Lasagabaster & Pagèze, 2018; Truchot, 2005). For example, the fledgling Journal of English-Medium

Instruction¹⁶ has not seen a single article about EMI in or from France since its first publication, back in 2022. The same appears to apply to international scholarly associations, such as the growing, European ICLHE¹⁷, which promotes the integration of content and language in higher education and regularly supports special interest research groups, organizes symposia and conferences, now on a global scale, from Brazil to Japan. The reasons for France's rather low profile on the international EMI research scene are hard to pin down, in a country in which 39% of the population speaks English¹⁸, either as L1 or EAL (Patel et al., 2023). In France, English can be perceived as a threat, even though the language is less widely used than in many other European countries, like Germany, the Netherlands, or all Nordic countries, and the country has often been uncomfortable with its multilingualism (Le Lièvre, 2008). However, there have been signs of a change, with increasing numbers of publications on EMI in English in the past decade (e.g., Blattès, 2018a, 2018b; Deneire & Benmokhtar, 2024; Le Lièvre, 2021; Taillefer, 2013).

Another reason why EMI in France has been under-researched is that France is a latecomer to the worldwide EMI trend (Blattès, 2018b). Offering EMI programs has not been imperative for attracting students from abroad: if it is the world's sixth-largest host country for international students (Campus France, 2024a), it ranked third up to the 2020s, and French is the fifth most widely-used language in the world and the fourth on the Internet¹⁹. Together with a top-down policy, it may be one of the reasons why since the 2000s the development of EMI has provoked much reflection, heated debates, and even legal actions²⁰ (similar to Italy²¹), among the political and academic circles (e.g., Blattès, 2018b, 2018a; Graziani, 2014; Grin, 2013; Héran, 2013; Le Lièvre & Forlot, 2014; O'Connell & Chaplier, 2014; Truchot, 2005, 2010). The strong, uncritical spread of Englishization has drawn fierce criticism (Frath, 2014, 2017; Graziani, 2022), with voices wanting to see plurilingualism encouraged instead, like that of the European Observatory for Plurilingualism (EOP, 2024; Tremblay, 2022). Another criticism is that France lacks language awareness (Taillefer, 2011), and that language issues are

¹⁶ Journal of English-Medium Instruction (JEMI), John Benjamins.

<https://www.jbe-platform.com/content/journals/26668890>

¹⁷ Integrating Content and Language in Higher Education association. <https://www.iclhe.org/>

¹⁸ Some figures for a few other EU countries (Patel et al., 2023, pp. 336–341, Appendix 5): Belgium: 52%; Germany: 56%; Italy: 34%; Spain: 22%; Netherlands: 90%. *The Future of English: Global Perspectives*, British Council (see Patel et al., 2023, p. 160, for methodology).

¹⁹ Source: <https://www.francophonie.org/>

²⁰ Caccivio, A. (2022, January 7): “Masters in English: The European Observatory for Plurilingualism's appeal against the University of Burgundy is rejected” (*Masters en anglais: Le recours de l'Observatoire européen du plurilinguisme contre l'université de Bourgogne est rejeté*) - Dépêche n° 665342. <https://www.aefinfo.fr>

²¹ See, for example: “Italian court pushes back on the race towards English”, University World News (2018, February 3). <https://www.universityworldnews.com/post.php?story=2018020304275939>

dealt with without prior study, without debate, without evaluation by the relevant public authorities, and without alternative solutions being sought (Truchot, 2010, 2013). Finally, some authors link the criticisms of Englishization to the French humanist tradition, pregnant with questions of quality, identity and diversity (Le Lièvre, 2021).

However, these debates have not hindered the development of EMI in France, still ongoing today (Beacco, 2019; Beacco et al., 2022; Le Lièvre et al., 2018). The promulgation of what is known as the Fioraso law²² in 2013, authorizing teaching in a foreign language, mostly English, was a catalyst: the official Campus France agency's motto is "In France, you can also study in English!", while their opening statement reads "Since 2015, and the law on higher education and research, all French higher education institutions, including universities, have the right to teach classes in English." (Campus France, 2022). It was preceded by the 2007 law on the liberties and responsibilities of universities²³, which some authors consider as the true start of the French engagement in the European Bologna process (Deneire & Benmokhtar, 2024). Somewhat paradoxically, the process also led to a re-nationalization, and sometimes an instrumentalization, as national objectives were promoted (Musselin, 2009). For instance, the graduate school of engineering in which this study is set has retained its status as a *Grande École d'ingénieur-e-s*, like most similar French institutions. It recruits the vast majority of its elite students after two years of dedicated preparatory courses (*classes préparatoires*) and highly competitive examinations, and it does not operate on the Bachelor's-Master's model (three + two years); it still awards an engineering degree (*diplôme d'ingénieur-e*) instead of a Master's degree²⁴, after three years of specialist engineering training. Despite this, it has remained a public institution and was recently absorbed, with other HE schools and colleges, in the local, growing public university. The move is part of a "university excellence initiative"²⁵, whose official aim is to improve international rankings, and whose result is a growing number of EMI programs. This is an example of how some authors argue that excellence and

²² Law of July 22, 2013, on higher education and research (*Loi du 22 juillet 2013 relative à l'enseignement supérieur et à la recherche*): "Teaching in a foreign language will be authorized when provided under an agreement with a foreign or international institution." <https://www.vie-publique.fr/loi/20671-enseignement-superieur-recherche-langue-etrangere-anglais-cours-en>

²³ *Loi n° 2007-1199 du 10 août 2007 relative aux libertés et responsabilités des universités (loi LRU)*. <https://www.legifrance.gouv.fr/>

²⁴ Students who leave the engineering school before the end of their training may obtain an equivalent Bachelor's or Master's degree, depending on their departure date. Musselin (2009, Section 4.1) further explains that a subtle legal and terminological difference has been kept in France between the *grade de Master* and the *diplôme d'ingénieur-e*.

²⁵ *IDEX: Initiatives d'Excellence. (Loi n° 2013-660 du 22 juillet 2013 relative à l'enseignement supérieur et à la recherche)*. <https://www.info.gouv.fr/conseil-des-ministres/compte-rendu-du-conseil-des-ministres-du-08-03-2017>

internationalization had almost become synonyms (Cosnefroy et al., 2020; de Ketele et al., 2016).

Two recent examples of developing French EMI science and engineering programs point to other solutions for helping students and lecturers best benefit from EMI. The first one was launched in 2015 at Institut national des sciences appliquées (INSA)²⁶, in Toulouse, and is called English-Medium Instruction at INSA (EMINSA, 2024; Moore & Détourbe, 2019). Over the years, the language department team have been able to convince science and engineering module teams of the interest of interdisciplinary collaboration. Through developing scientific teaching in English, their objective is both to meet the needs of French-speaking students, faced with the increasing internationalization of the engineering profession, and to raise the graduate school of engineering 's international profile and appeal among foreign students. The multi-disciplinary EMINSA team, made up of INSA science lecturers and English teachers, guides and supports teaching teams in setting up science courses in English, while designing appropriate learning situations. The project-based scheme aims to integrate active or blended learning modes to ease the transition into English, for both students and lecturers. Building on its success, since 2020 the team has been offering other INSA establishments a hybrid version of the Toulouse project. Another recent instance of large-scale spread of EMI in science and engineering is to be found at the Université de Montpellier²⁷ and is called Inter-Disciplinary and In-Lab graduate program since 2022 (IDIL, 2022). Comprising eleven Master's English-taught programs, its objective is to familiarize students with the world of research in laboratories attached to the University of Montpellier, and encourage interdisciplinarity around three main themes: Nourishing, Caring and Protecting. However, apart from an Effective Communication optional module for students to develop public speaking, body language and reading skills, if any language-and-content interface exists, no mention seems to be made of it.

Concerns about the quality of EMI programs has raised the unavoidable and politically sensitive issues of EMI staff competence, instructional designs, expected learning outcomes, and qualification assessment (Klaassen & Räsänen, 2006; Räsänen & Klaassen, 2006). As a consequence, tests have been developed to certify EMI lecturers and training programs. There does not seem to be any dedicated EMI test in France, whereas some European universities developed internal assessment methods based on the language for specific purposes model

²⁶ National Institute of Applied Sciences, a university network of seven French engineering establishments located in major French regional cities: Lyon, Rennes, Rouen, Strasbourg, Toulouse, Valenciennes, Blois, and Bourges. It is the largest engineer training group in France.

²⁷ University of Montpellier, IDIL Graduate Program. <https://idil.edu.umontpellier.fr/en/>

(Dimova & Kling, 2018). For example, two tests, the Test of Oral English Proficiency for Academic Staff (TOEPAS) from the University of Copenhagen (CIP-KU, 2009; Dimova, 2017, 2021; Kling, 2012), and the Certificate in EMI Skills, from Cambridge University (Cambridge English, 2021)²⁸, have been available both locally and worldwide. O’Dowd’s survey of the training and accreditation practices in seventy European Universities (2014-2015 period) revealed a highly diverse range of approaches in both criteria, raising the issue, among others, of whether enough lecturers are able to teach the content of their subject through the English medium (O’Dowd, 2018).

Some universities have addressed the question of lecturer training and professional development. For example, the Polytechnic University of Madrid developed the Teacher Target Model, with the aim of helping content lecturers deliver EMI more effectively (Sancho Guinda, 2013). Another format is inclusive programs such as the ‘English for staff’ at Germany’s Hochschule RheinMain, aimed at training teachers, researchers and administration members, and to which three English trainers are dedicated (Schutte et al., 2020). In France, for example at the university where this study was carried out, comprehensive English skills development courses are organized throughout the year for all categories of personnel, at various levels, in individual or group sessions. Typical topics covered include: interaction with foreign students, staff, and partner institutions; e-mailing and telephoning skills; language for research purposes, seminars, conferences and presentations; lecturing in English; day-to-day business situations abroad. However, among the lecturers participating in this study, few have benefited from these courses: the time they have available in their lecturing schedule is generally absorbed by one or more heavy responsibilities, pedagogical meetings, and a miscellany of administrative tasks²⁹. Perhaps one of France's most advanced programs is Université de Bordeaux’s ‘Défi International’ initiative, which aims to develop the range of courses offered in English and international degrees, through the implementation of pedagogical and linguistic support, and training for teaching in English (Pagèze, 2017). The full-fledged program offers, on the lecturers’ side: tailor-made support for teaching teams during the transition to, and implementation of, international courses and training modules to prepare for teaching in English, with support and follow-up; on the students’ side: help and support for students taking international courses. In 2018, Lasagabaster and Pagèze undertook a thorough survey of Défi

²⁸ The test was retired in 2023.

²⁹ This information was gathered during interviews prior to the recording of their lectures.

International, and the findings revealed a number of instructive points about EMI in France, and in general. Below is a selection of some of these points (Lasagabaster & Pagèze, 2018).

- Transforming HE on the ground takes time.
- The French context highlighted how cultural and political issues were a complicating factor for EMI.
- The development of quality teaching in English imposes an additional workload on teachers.
- Content lecturers are willing and able to manage and organize classroom communication to support EMI learning, if given the tools and means.
- Lecturers are more likely to invest in EMI if they see gains in the lecture hall.
- Lecturers changed their teaching practices and made their classes more student-centered.
- The shift in perspective reduced the stakes for lecturers: they began to see English as a ‘lingua franca’, which helped to allay some of their fears as non-native speakers.
- Défi International is changing perspectives on the role of language in learning, but it is still a long way from CLIL/ICLHE.

The internationalization of higher education is subject to the constraints of the local context, but it can also become a driving force for change within that context (Lasagabaster & Pagèze, 2018). While considering the particularities of the French cultural and educational contexts, some researchers have suggested approaching the CLIL-ICLHE topic using a SWOT analysis matrix (strengths, weaknesses, opportunities, and threats) to set realistic goals, between teachers of English for specific purposes, content teachers and institutional representatives (Taillefer, 2013). Obviously, one of the first decisions to be made is to decide upon the degree of integration that is the objective as on the EMI-CLIL continuum (Figure 1). In the French graduate school of engineering where this study was carried out, the “I” in CLIL-ICLHE (which represents *integrated*, *integrating*, or *integration*) has not been implemented, apart from the optional staff training courses mentioned earlier: engineering courses and language-and-

communication courses coexist, side by side, with little or no interdisciplinary pedagogical collaboration, team teaching or co-constructed learning outcomes.

To conclude, Deneire & Benmokhtar (2024) surveyed a representative sample of the three principal stakeholders in the French EMI venture using questionnaires and interviews: university and program administrators, EMI instructors, and students in EMI programs. They found that the stakeholders have fully embraced the EMI values, as defined by international rankings, namely quality and employability. For the time being, administrators have privileged the international profile of their institution and the employability of their students, so far to the detriment of intercultural or European values. While most lecturers agree that a balance must be preserved between English and French, they too find that EMI contributes to improve employment opportunities for students, but also that it builds their international linguistic skills and self-confidence. However, the faculty favor international cooperation and collaboration over ranking competition. Students, too, value the wider employability, the professional and general culture, and the travelling opportunities that the practice of EMI has given them. Nonetheless, Deneire and Benmokhtar point out that these rather flattering results from their EMI survey should not distract attention from the overall picture and certain failings of the French higher education system. Rigorous selection processes exclude students who have had less opportunity to develop their English language skills and who may become victims of discrimination. The implementation of EMI may, thus, act as a barrier for many students, increasing inequalities within the French tertiary education system. This goes against the Council of Europe's four goals, as expressed in the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2024, homepage): "to protect linguistic and cultural diversity, promote plurilingual and intercultural education, reinforce the right to quality education for all, and enhance intercultural dialogue, social inclusion and democracy". For Deneire and Benmokhtar (2024), it seems that twenty-five years on, the Bologna process (European Commission, 1999) in France has thus engendered quite the opposite.

2.4 English as a lingua franca

The rapid expansion of EMI over the last 20 years has been accompanied by the growth of research into its vehicular language, English as a Lingua Franca (ELF)³⁰. Up to the 1990s, a

³⁰ The term 'lingua franca' (or 'Frankish tongue') has its origins in the mixture of languages which, between the 16th and 19th centuries, served as a widely shared means of communication around the Mediterranean and beyond.

lingua franca could be defined as “any lingual medium of communication between people of different mother tongues, for whom it is a second language” (Samarin, 1987, p. 371). The modern spread of English as an international language has generated more refined definitions of ELF, like “any use of English among speakers of different first languages for whom English is the communicative medium of choice, and often the only option” (Seidlhofer, 2013, p. 30). In our context, it is defined as the medium of communication used by academics and their students who are speakers of English as an Additional Language (EAL). ELF may diverge from the L1-English norm, or Standard English, in its phonology, syntax, lexis, as well as in its interdiscursive or social aspects (Björkman, 2008b; Kappa, 2016; Newbold, 2023). Consequently, the concept of a native speaker has been seriously challenged (Davies, 2003, 2013; Paikeday, 1985; Seidlhofer, 2000) and progress has been made toward descriptive rather than prescriptive approaches (Jenkins, 2014, 2018). Moreover, divergence³¹ is becoming legitimized, as ELF is gaining its own idiomaticity (Björkman, 2013; Briggs Baffoe-Djan & Smith, 2017), and this development impacts EMI in terms of learning outcomes as well as teaching practices. It is thus necessary to see how ELF has arisen and what the implications are in terms of its existence within institutions of Higher Education.

2.4.1 *Framing and naming plural English-es*

Kachru's oft-cited model of three concentric circles is the starting point for an overview of international English around the world (Kachru, 1985). Kachru divided what he termed World English into three circles: the inner circle (English-speaking countries), the outer circle (post-colonial regions; English is used as an additional and often official language) and the expanding circle (countries where English is learned as a foreign language). This model sparked much debate, like McArthur's constructive critique of Kachru's model (McArthur, 1987). Today, this fixed-region model is seen as simplistic and no longer adapted to the growing use of English in a world of near-instantaneous communication technologies and unlimited mobility (Wingate, 2022). The terms International English and English as an International Language (EIL) have been used, and from the 2000s on ELF caught on, notably by Jenkins (Jenkins, 2000a, 2007; Cagliero & Jenkins, 2011; Archibald et al., 2011) and Seidlhofer (Seidlhofer, 2000, 2004, 2005; Pitzl et al., 2006). The implications of English as a lingua franca has also

Although there are few written records of it, one example can be found in Molière's play *Le Bourgeois gentilhomme*, Act 4, Scene 5. (Dakhli, 2008)

³¹ The term ‘divergence’ as used in this study is not to be confused with its use in the interactive process of ‘accommodation’ and its theory, as proposed by Giles et al. (1991), in which ‘divergence’ is opposed to ‘convergence’, as understood from a social-psychological perspective.

been operationalized for language teachers by researchers such as Walker (2021) and Patsko (Patsko, 2013, 2014; Dewey & Patsko, 2017). Since the mid-2010s, ELF has clearly overtaken EIL as the more commonly used abbreviation (Henderson, 2021). The concept of a single core variety was questioned and expanded to the plural World Englishes by some of authors (e.g., Baratta, 2019; Onysko, 2021). The idea was to pluralize English to work towards an egalitarian approach and recognize Englishes as enfranchised codes worldwide, a move that Kachru himself also made (e.g., Kachru, 1990, 1997). Thus two conceptions of English research came to coexist: the varieties of English, or plurality from within, and plurality among the Englishes of the world, as separate entities (Onysko, 2021).

2.4.2 Engineering lecturing in English as a lingua franca

In general, monologic lecture discourse tends to comprise much longer and complete utterances than dialogic speech. However, this is commonly not a direct consequence of the lecturer's linguistic proficiency, but rather a generic characteristic (Björkman, 2008a). Research into linguistic variation in classroom interaction has found that, unlike higher education subjects such as business or education, engineering lectures are teacher-led and feature little or no student-led interactive discourse, or collaborative efforts in which lecturers and students are equally involved (Csomay, 2007). In these lectures, ELF users are not learners or teachers of English, but speakers. Thus, such terms as interlanguage or language errors are not applicable in this context. In the early 2000s, Seidlhofer and Jenkins were already pointing out that it was necessary to study what ELF users actually did when they communicated, rather than assuming what they needed or should do (Seidlhofer & Jenkins, 2003). One speaker largely dominates the lectures, and interaction tends to be limited to a few questions or comments from other speakers. In such situations, utterances tend to be longer and more complete than in interactive discourses and expert content is delivered by highly experienced specialists in the relevant field (Björkman, 2008a). However, lecturers do not always produce 'expert' or 'nativelike' English, as discussed in the next section.

2.4.3 Changing the paradigm

Higher education institutions who seek to prepare their students for the real world are taking the development of ELF into account (Henderson, 2021). With regard to EMI engineering lectures, it is crucial to note the move away from the normative concept of language

error or mistake, based on the native speaker model, to that of ELF divergence from the norm (Björkman, 2008b, 2008a). Table 1 shows some of the aspects of this shift.

Table 1
A paradigm shift from English language classes to EMI lectures

Aspects	English language classes	EMI lectures
English	Object of study	Functional tool
Students	Learners	Users/Speakers
Focus	Form/Language	Content/Science
Target model	Nativelike norm	Intelligible speech
Goal	Improved language skills	Effective content transmission
Language approach	Prescriptive: language errors = mistakes	Descriptive: divergent forms = features

In this emerging ELF perspective on EMI lectures, English moves from being an object of study to a functional tool for conveying content knowledge. Students evolve from being the learners of the language to the users-receivers-decoders of a linguistic medium. In the lecture hall, the pedagogical focus shifts from the language itself to the content conveyed through the vehicular tongue, and the target model, from a theoretical nativelike norm to making the information intelligible. Intelligibility becomes the key factor in transmitting course content effectively, without concern for improving students' language skills. Linguistic errors evolve from mistakes to be corrected to widely accepted divergent forms. Such typical cases can include the omission of the third person singular "s" ending in verbs in the present tense, for example, or the use of the Present Perfect instead of the Simple Past, when French speakers mirror the *Passé Composé* structure of their L1.

2.4.4 *Instructors as communicators*

In the context of this study, the mainly monologic EMI lecture discourse is delivered via English spoken as an additional language. The lecturers' primary objective is to deliver their engineering content to their student audiences effectively, but they must simultaneously cope with the extra cognitive load inherent in expressing themselves in a foreign language. Extensive research has been dedicated to the cognitive load of learners, learning and acquisition (e.g., Chanquoy et al., 2007; Hulstijn & Laufer, 2001; Paas et al., 2010; Roussel et al., 2017; Roussel & Tricot, 2017; Sweller, 2016, 2018; Sweller et al., 2011), but less has been done on the cognitive load of instructors.

For lecturers, it is possible to posit that the situation-induced cognitive load, together with other factors such as intermediate levels of EAL knowledge and skills, or end-of-session fatigue, may combine to disturb fluency and coherence of discourse, engender sequences of hesitations, or generate misunderstandings and communication breakdowns. Divergence from the L1-English ‘native speaker’ norm, or Standard English, may concern interdiscursive or social aspects in ELF oral exchanges, like politeness, appropriateness, humor, or turn-taking (e.g., Alsop, 2016; Kappa, 2016; Newbold, 2023; Tzoannopoulou, 2016), or phonology, at segmental and suprasegmental levels (e.g., Derwing & Munro, 2009a, 2009b; Frost et al., 2021; Henderson, 2019, 2021; Jenkins, 2000b; Richter, 2019; Trofimovich & Isaacs, 2012). As with all types of divergence, an ELF perspective can make ‘errors’ acceptable, yet the major question is whether it disturbs the communication situation (Björkman, 2008a, 2008b). More importantly, she also suggests that, because engineering is more of a nomothetic³² discipline, like economics or linguistics, in which language reports results (contrary to idiographic disciplines like philosophy or literature), it may reduce disturbance in communication. However, I would argue that engineering lectures are not solely about reporting procedures and results, and that they do combine nomothetic and idiographic orientations. Indeed, I posit that pragmatic functions such as summarizing and storytelling operate as a bridge between types of knowledge, thus reflecting the lecturers’ desire and ability to make their high-stakes content accessible. Analyzing those pragmatic functions makes them a useful locus to observe how ELF divergences can coexist with good, or at least well-intentioned, lecturing practices. It may therefore be instructive to create a corpus of engineering courses in order to observe whether, and how, summaries and stories contribute to this nomothetic-idiographic bridge. Such a corpus is presented in the next chapter.

³² Cf., Windelband (1998, 1894 translated reprint), German philosopher (1848-1915) mainly remembered for introducing the terms *nomothetic* and *idiographic*.

3 Corpus construction

The present study is based on the investigation of a corpus of original recordings of French engineering lectures delivered in English: the Corpus of French Engineering Lectures, or CoFEL (Picavet, 2022). The present chapter describes the theoretical aspects and practical steps involved in building and using CoFEL: the setting (3.1), the participants (3.2), the collection of the data (3.3), the corpus management tools (3.4), the compilation of the corpus, and its storage (3.5). The choice was made from the outset to preserve the anonymity of the participants and the various institutions involved. Therefore, their identities are either coded or not mentioned.

3.1 Setting

The data was collected at a graduate school of engineering in France, chosen for the diversity of its international undergraduate programs. The graduate school comprises eight departments offering thirty-four training programs in various branches of engineering, and a graduate school of management. There are fifty-three Master's (MEng) programs, thirty-six of which are 'international' and taught in English). The range of engineering fields taught is very broad, and the following list is not exhaustive: physics and materials science; mechanical and energy engineering (renewable energy, bioenergy, fluid mechanics); computer science and engineering (cybersecurity, information and communications technology, operational technology, automation); electrical engineering (smart grids, power electronics, embedded systems); civil engineering (geotechnical and hydraulic engineering, risk management); health technology and biomaterials; environmental engineering. In 2022-2023, the graduate school enrolled 8,350 students, a quarter of whom came from abroad. Research is also carried out, as the school is home to thirty-eight particularly active research laboratories, with an average of 230 theses defended each year. Moreover, the vast majority (80%) of the teaching staff is involved in scientific research³³.

All the lecturers participating in this study work in the same department and teach in degree programs in energy, civil and environmental engineering. In particular, these lecturers teach in one of the following international MEng courses: Hydraulic and Civil Engineering, Electrical Engineering, Fluid Mechanics, and Energy Systems. As all the lectures are delivered

³³ The data in this paragraph were retrieved from the website of the graduate school of engineering.

in English as an additional language, to international student audiences, the context can be categorized as EMI. A detailed description of the participants is given in the next section.

3.2 Participants

There are two groups of participants in this study: the lecturers, and their students. A detailed description of the lecturers is indispensable to better understand the CoFEL findings. The student audiences are also described because they are inseparable from the pedagogical situation and the *raison d'être* of IME-MEng courses. For the record, lecturer-student and student-student interaction, and the resulting “communities of practice” (Wenger, 1998), were to be at the heart of the present research. However, online teaching during COVID-19 cancelled a year’s work in this direction, at the start of 2020. The object of this thesis thus became the lecturer’s monologic discourse.

3.2.1 Participating lecturers

Fourteen MEng lecturers from the department of energy, civil and environmental engineering of the graduate school volunteered to participate in this study. Two of them did not take part in the end, as no opportunity or possibility of recording their courses in English arose. Table 2 lists the twelve participating lecturers, together with some personal and professional details. To respect confidentiality and anonymity, a randomly-chosen three-letter code is used systematically to designate the lecturers, notably in data file names. They appear in the table in chronological order of recording, with the exception of the first one to be observed (MIL), but who was not recorded. This was the earliest observation of an engineering lecture in this study, where, as an observer, I sought to get a first impression of the pedagogical context of an engineering lecture, before installing all the audio and video recording equipment in the lecture hall. Thus, this left eleven speakers to observe and record.

Table 2

Graduate school of engineering: list of the 12 participating lecturers

Rec. #	Lecturer code	Gender	Status	Age (2022)	Nationality & L1	Broad engineering discipline
None	MIL	M	Associate professor	46	French	Environmental engineering
1	FAC	M	Associate professor	35	French	Fluid mechanics
2	CAM	M	Professor	51	French	Civil engineering
3	ELO	F	PRAG*	51	French	Civil engineering
4	AMI	F	Postdoctoral fellow	32	Italian	Civil engineering
5	OFA	F	Postdoctoral fellow	30	Italian	Civil engineering
6	PAC	M	Associate professor	45	French	Energy systems and markets
7	SAJ	M	Professor	54	French	Electrical engineering
8	UPA	M	Associate professor	53	Romanian	Automation and intelligent systems
9	LUY	M	Professor	54	French	Mechanical engineering and energy
10	DIM	M	Associate professor	38	French	Mechanical engineering and energy
11	TAY	M	PRAG*	55	French	Electrical engineering

*PRAG: French term (*professeur agrégé*) for a secondary education teacher employed in higher education.

The lecturer group includes both women and men (three women and eleven men), and they range in age and experience from post-doctoral researchers in their early thirties, through junior and experienced associate professors, up to senior professors in their fifties. The majority (nine) are L1 speakers of French, complemented by two Italians and one Romanian. All the participants lecture in English as an additional language. In the field of public tertiary engineering education in France a majority of lecturers are *enseignants-chercheurs*, that is to say, faculty with a PhD, and with an equal share of lecturing and research responsibilities. Out of the twelve faculty participants in this study, only two of them are *professeur agrégé de l'enseignement du second degré exerçant dans l'enseignement supérieur*, or “PRAG”: tenured secondary education teachers with the highest qualification called *agrégation*, employed in tertiary education. Some PRAGs hold a PhD, and although it can be an asset in a recruitment, it is not a formal requirement for them. Secondary education tenured teachers account for around fourteen per cent of all higher education faculty members nationwide (Adedokun & Tourbeaux, 2022, p. 1). Two other participants were postdoctoral research fellows on a temporary contract. Although the panel of participants is small, it is nonetheless representative of the distribution of lecturers' profiles in the French public higher education sector, in terms of qualifications and experience. The same can be said of the gender, nationality and age distribution, with three female lecturers out of twelve (39% nationwide—Adedokun & Tourbeaux, 2022, p. 4) and with nine French nationals, two Italians and one Romanian, and covering an age range from 30 to 55 years old in 2022.

Teaching and managing highly heterogeneous groups of multinational, multilingual, multicultural students may represent a significant linguistic challenge for lecturers, as speakers and listeners. While they are certified and recognized experts in their fields of engineering, in the EMI context they must deliver expert scientific or engineering content in an additional language. The level of English of the 11 lecturers on our panel was assessed for two skills (listening and reading) using Cambridge University's Linguaskill General test (Cambridge English, 2022). This test was chosen because its results are aligned with the EU's Common European Framework of Reference for Languages (CEFR, Council of Europe, 2020), evaluating levels on a scale from A1 (breakthrough) to C2 (mastery). The results ranged from 'waystage level' A2 to C1+ (or 'C1 or above', the maximum possible with this test: advanced to mastery levels): one lecturer tested at level A2, one at B1, two at B2 and seven at C1+. While these lecturers' test results are just one indicator, notably as only two skills were tested, one third of the lecturers on the panel do not have C1 level proficiency. Unlike some other countries (O'Dowd, 2018), in France lecturers are not systematically tested for their level of English, and no certificate of proficiency or accreditation is required before institutions entrust them with responsibility for EMI courses. The graduate school of engineering where this study took place did not have such requirements when CoFEL was completed.

3.2.2 MEng student audiences

The typical MEng student audience in this study is a mix of two distinct groups, in varying proportions: international (i.e., non-French) students, and French students. International students are the main target audience for the research-oriented MEng courses, often leading to a doctorate. These students generally hold a bachelor's or 1st-year Master's (M1) degree from their home university, enabling them to enroll in French MEng courses in related disciplines, at M1 or M2 level, having thus two years or 1 year of study remaining, respectively. In order to be attractive and keep up with the global EMI trend in higher education, in the early 2010s the institution decided that these international MEng programs would be taught in English from the outset. In contrast, a majority of French students do the classic 5-year engineering curriculum of the *Grandes Écoles* system. The first two years consist of the elite preparatory course called "*classes préparatoires*", which are done outside the graduate school of engineering for the majority of students; the school hosts two *classes préparatoires*, which account for just 10% of the total number of students admitted each year. This is followed by three years of specialization in engineering, leading to the engineering degree (*diplôme*

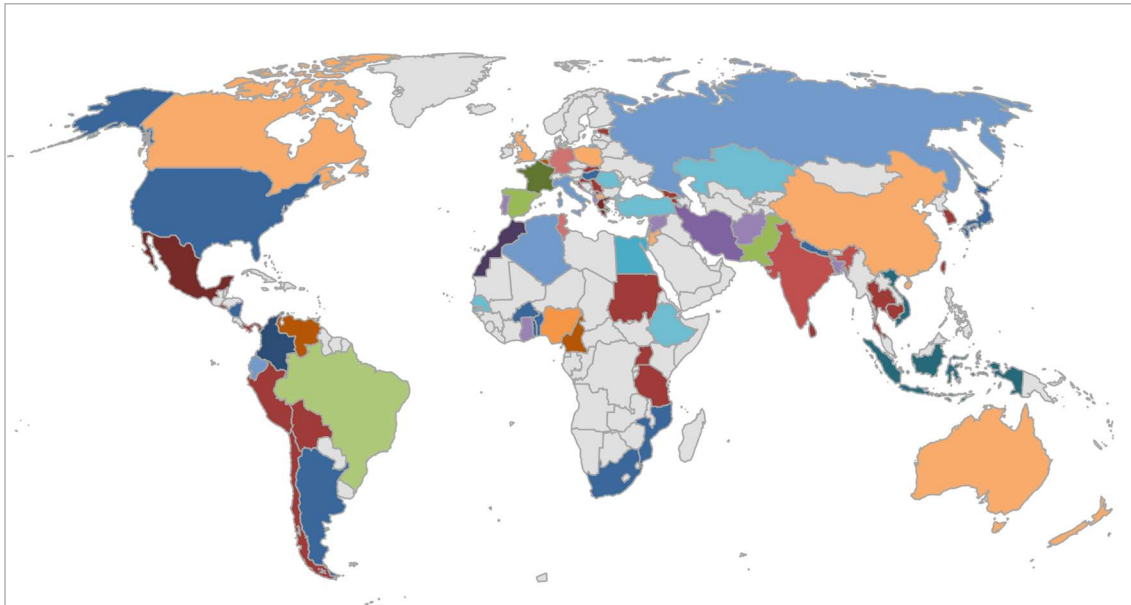
d'ingénieur-e), in one of the school's eight departments of engineering. Although in practice the first of these three years is equivalent to a bachelor's degree, and the next two years to a Master's degree, the *diplôme d'ingénieur-e* term is used to designate the degree that is delivered at the end of the training. For the majority of students, this engineering degree leads directly to employment and a career. Only fourteen per cent of the junior French engineers continue with three more years of study and research to prepare a doctorate. Thus, on the whole, the final objective of the two groups of students can differ: some go on to do doctoral-level research while others go directly into salaried employment outside of academia.

Despite these differences, for practical reasons and economies of scale, French and international students at MEng levels are most often grouped together in the same course, when subjects converge. This convergence is the main factor leading to the use of the English medium in the MEng engineering courses, and thus to the emergence of EMI. Student groups frequently differ in size and composition from one session to the next, depending on students' study paths, as a host of options and elective courses are offered to students, and the type of classes vary (lectures, tutorials, design sessions, practical or laboratory work).

By way of illustration, details will be given for the department of energy, civil and environmental engineering, where our observations were made. In this department, as the map in Figure 3 shows, international students come from all over the world. The seventy-eight countries that contributed at least one student in the first ten years of the MEng programs at the department are shown in color. Enrolment spans all five continents. A number of these foreign countries have French as their national, official, or vehicular language, particularly on the African continent, for example, the Democratic Republic of the Congo, Cameroon, the Republic of Côte d'Ivoire, Benin, the Togolese Republic, or the Central African Republic.

Figure 3

Department of energy, civil and environmental engineering: origin of MEng students (2013-2022)



Note. The 78 colored countries contributed a minimum of 1 student.

The graph in Figure 4 shows the general upward trend in international student cohorts. Although the COVID-19 pandemic dealt a severe blow to the student enrolment dynamic in 2020, headcount progress is constant, on average. The 2021 peak is due to pandemic adaptation and compensatory measures, as enrolment quotas were increased to enable the left-out students of 2020 to enroll with those of 2021. The peak in 2021 is due to adaptation to the pandemic and compensatory measures, as enrolment quotas were increased to allow students left out in 2020 to enroll with those in 2021. Over the 2013-2022 ten-year period, the number of MEng programs³⁴ rose from two to four, which naturally led to an increase in the headcount.

³⁴ As of 2022, the four programs were: Fluid Mechanics and Energy (FME), Hydraulic and Civil Engineering (HCE), Electrical Engineering for Smart Grids and Buildings (SGB), and Mobile, Autonomous and Robotic Systems (MARS).

Figure 4

Department of energy, civil and environmental engineering: student headcount trend (2013-2022)

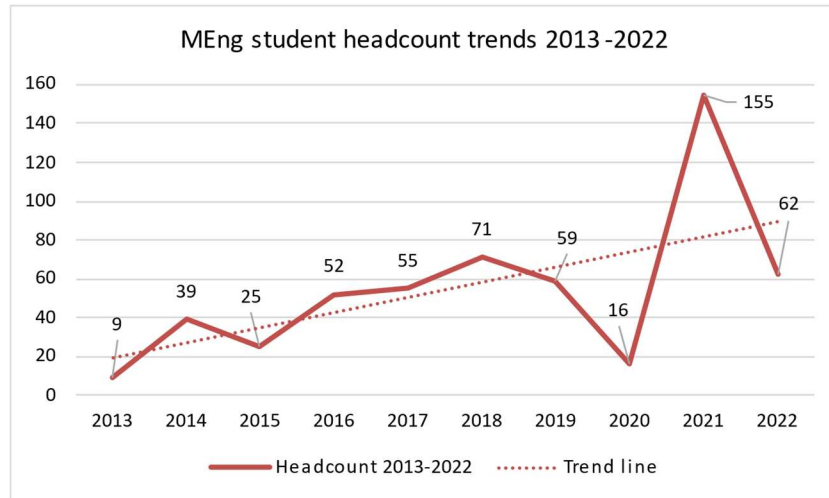


Table 3 presents the twenty-eight top contributing countries (out of 78, in decreasing order) to the MEng cohorts, over the first ten years of existence of the programs, highlighting the great diversity in the geographical origins. The higher figures in the table are the result of the graduate school of engineering's active policies of international agreements and partnerships with foreign universities (over 350), and the implementation of the EU's Erasmus exchange program. As the table shows, a proportion of French students ($n = 15$) from traditional university backgrounds (i.e. not from the *classes préparatoires*) also joined the programs at Master's levels. Today, despite the increase in numbers, MEng students represent only 7% of the total yearly population of about 1,000 students in the department. Thus, to date, the vast majority of students (93%) still follow the standard three-year curriculum of specialization leading to an engineering degree. The presence of this minority of international students has, paradoxically, contributed to the majority of courses in the last two years of the classic French curriculum being taught in English to all students. This is how English has become the lingua franca (ELF) used in the pedagogical situations today at this institution, which is representative of others in France and across Europe.

Table 3

Department of energy, civil and environmental engineering: countries of origin of MEng students

Country	n	Country	n	Country	n	Country	n
Lebanon	57	Colombia	19	Belgium	9	Germany	7
India	53	Greece	16	Cameroon	9	Tunisia	7
Pakistan	30	Mexico	16	Venezuela	9	Brazil	6
Spain	30	France	15	Algeria	8	Afghanistan	5
Iran	27	Morocco	12	Ecuador	8	Albania	5
Egypt	22	Indonesia	11	Italy	8	Bangladesh	5
Nigeria	21	Vietnam	11	Russia	8	Ghana	5

Note. 28 top countries for the 2013-2022 period

Being a student in an EMI context means facing certain challenges. On the one hand, there is the multi-lingual, multicultural class group made up of students from all over the world. On the other hand, English levels can range from the official minimum admission requirement, the EU's 'threshold level' B1 (Council of Europe, 2020), to students from English-speaking countries for whom it is the native language. There may be strong internal differences in psychology, attitudes, behaviors, or beliefs, making group dynamics and relationships potentially complex. Adapting to a new local language (French), a new language of learning (English), new academic and pedagogical approaches, immersion in a new country and a new culture represent considerable challenges for many of the students, both international and French nationals. For the latter, they are still in their home country, yet immersion in the new mix of international students and the discovery of EMI after a lifetime of studying in French can be different from any of their previous educational experiences.

3.3 Data collection

The Corpus of French Engineering Lectures (CoFEL) comprises videos, audio recordings and annotated transcripts from ELF-EMI engineering lectures delivered to MEng-level students, between fall 2019 and fall 2021. Studying what goes on in the lecture hall during an EMI engineering lecture means going there as an observer and recording the observed sessions for further analysis. To keep observer impact to a minimum, the recording equipment was installed before the session, and the participants were invited to fill out the consent forms online, when possible. Data were collected in two ways, depending on two distinct and quite different recording periods: pre-COVID, and during COVID. Initially, the lectures were recorded using high-definition video cameras and audio recorders in the lecture hall, and both lecturers and students were recorded, as the focus was to be on lecturer-student and student-

student interaction. However, from March 2020 on, due to the outbreak of the COVID-19 pandemic in France, most courses were conducted remotely, via online platforms. The project to study live interaction on the site was no longer viable, and the choice was made to focus the research on the lecturer's monologic discourse. From then on, the recording of lectures was done online, via the recording function available on those platforms, and only the lecturers were recorded. The consent of participants was always collected, both from lecturers and students. The lecturers' recording consent form can be found in Appendix A. For the on-site sessions, the students' recording consent form was available in French, or in English for international students (cf. Appendix B, English version). Over the two-year period (2019-2021), twenty-eight lectures were recorded in twelve different engineering subjects, for a total of sixty hours. As the lectures were recorded, the next step was to compile and rigorously organize them into a functional, usable database: a corpus of lectures. For the purposes of the exploratory study reported in the present thesis, only about half (27 hours) of the raw corpus has been analyzed. Thus, in the remainder of this thesis, the acronym "CoFEL" refers only to this part of the raw corpus that was annotated, and transcribed.

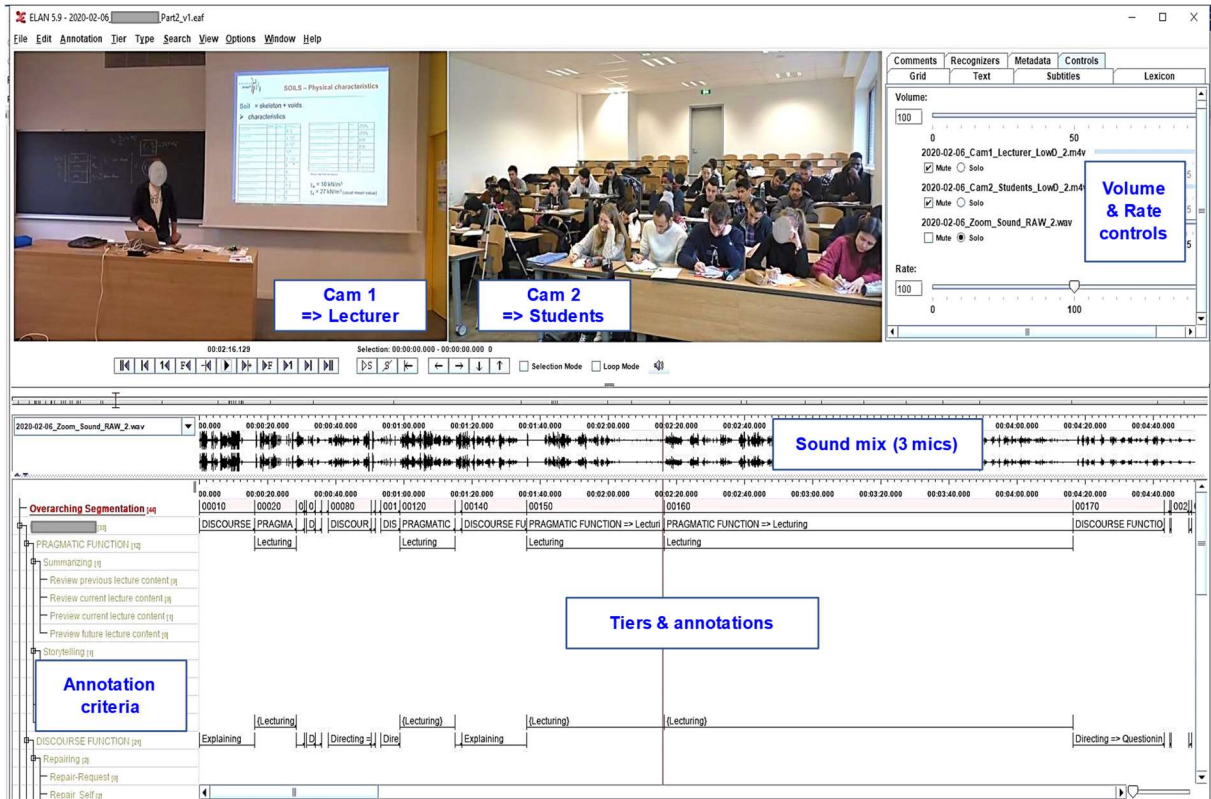
3.4 Corpus analysis tools

3.4.1 *The ELAN software*

The ELAN annotation software provided by the Max Planck Institute for Psycholinguistics (The Language Archive: 2019 to 2023 versions) was used to annotate and transcribe the synchronized audio and video recordings. Figure 5 shows the typical ELAN interface. The software was chosen for its ease of use, flexibility, and accurate multimedia synchronization capabilities. It supports an unlimited number of textual annotations to audio and video recordings. These annotations can be performed on multiple layers, called 'tiers', which can be organized and interconnected in a hierarchical fashion. One asset of ELAN is that the annotated text is time-aligned, making it easy to view, retrieve or export annotations, transcripts, and time codes synchronized with the multimedia timeline, with up to millisecond accuracy. The annotation-transcription work was done manually. This task requires a great deal of attention and time: one hour of lecture recording generally represents fifteen to twenty hours of work. Processing recordings in English as a lingua franca (ELF) makes the task even more challenging because of the presence of norm-divergent language items.

Figure 5

ELAN: typical workspace window (pre-COVID recordings)



Note. Typical ELAN window display: 2 video camera windows, multimedia controls, and annotation criteria and space.

The transcription of the utterances into written text was done using the Vienna-Oxford International Corpus of English conventions (VOICE Project, 2007; 2021) as guidelines. These conventions were chosen for two reasons: the VOICE corpus is one of the main internationally-recognized corpora studying EMI, reflecting years of prominent researchers' experience in the field, and its annotation conventions are freely accessible to other researchers. The main VOICE-inspired codes used in the CoFEL excerpts and transcripts in this study are listed in Table 4. During the transcription process, the video data was particularly useful for disambiguating certain unclear audio passages, because one can see the lecturers' gestures, facial expressions, chalk-board text, drawings, or course material in slideshows. Where available, course handouts and slideshows in paper or digital form were collated and often proved valuable in understanding specialist terminology.

Table 4

Key to abbreviations and codes used in CoFEL excerpts and transcripts

Mark-up convention	Meaning
er	Hesitation
the: / the::	Colon: sustained syllable (1 second / 2 seconds or more)
many MANY years ago	Capitals show insistence / higher voice volume
(ok?)	Text in brackets: uncertain transcription
on roads- railways	Hyphen + space: fragment; change of focus in text; enumeration
(4)	Number in brackets: pause or elapsing silence, in seconds
@@@	Indicates laughter: one "@" corresponds to moderate laughter (approx. 1 syllable); "@@@@" corresponds to big laughter
{erases chalkboard}	Added contextual information / Transcriber's comments
xxx	Unclear speech / one "x" corresponds approximately to 1 syllable (when identifiable)
<FRENCH> / </FRENCH>	Participant speaks another language than English, here French (code-switching) / End of code-switching passage
<SOFT> / </SOFT>	Participant speaks in a soft voice / End of soft passage
SS	Students (e.g., when reacting noticeably to the lecturer's speech)

Note. Adapted from the VOICE Project (2007)

Annotation designates the identification stage, where pragmatic and discursive categories are used. An analysis grid was developed, encompassing these categories inspired by the relevant EMI literature, as shown in Figure 6. In ELAN, this analysis grid appears under the form of a dendrogram of tier dependencies, displaying the various annotation layers, or tiers, and their hierarchy. Then, a systematic serial reference number was attributed to every annotation made in each lecture. The numbering system uses four digits, from 0000, to the last annotation in the lecture, for example, 0840, in steps of 10. The annotations were made in tens as a precaution, so that they could be divided into nine additional annotations if further detail was added later. In this thesis, the in-depth analysis is limited to the two top tiers in the grid, namely summarizing and storytelling, covered in Chapters 4 and 5 respectively. The rest of the pragmatic and discursive functions that make up the entire corpus have nevertheless been transcribed. This makes it possible to carry out measurements on annotated text, such as length (number of tokens), duration, or frequency, and facilitates future exploitation of the corpus.

In corpus analysis, the length of a string of text like a summary can be measured by counting the number of tokens that make up that string (Alsop, 2015). A “word” is defined as a sequence of characters with a word separator at each end, and “tokens” refer to running words, also known as the total number of units that make up a text (Scott, 2020). A token is the smallest

unit that a corpus consists of, or to put it differently, it is anything that lies between two spaces, notably including non-words like hesitation marks (*er* in CoFEL; see subsection 3.4.2 below), or abbreviations (Kilgarriff et al., 2014). Word “type” are distinct as type refers to a class, to different words, whereas token refers to individual instances of that class, each type generally being illustrated by more than one token (Scott, 2020). Thus, in corpus analysis, tokens make distinct units from words and types as they are counted linearly regardless of their frequency of repetition. As such, tokens are units which can be counted to determine the length of a text, while the duration measures the elapsed time of an utterance, or a set of utterances, and includes pauses.

Figure 6

Screenshot of CoFEL tier dependencies (annotation grid) in ELAN.



The ELAN software was used in the annotation and transcription processes of CoFEL multimedia data files. However, due to the limitations of our local computer system or those of ELAN's data processing capacity and functions, I had to use other software to manage and exploit our files at corpus level. These limitations became glaringly apparent when combining and synchronizing several high-definition (HD) two-hour videos and audio soundtracks in a single ELAN file: I experienced frequent software crashes, making the whole-lecture work difficult at best, and often impossible. A twofold solution was found: divide the two-hour multimedia files into four lighter 30-minute files for work with ELAN, and use specific manager software at corpus level: the Sketch Engine online platform (Kilgarriff et al., 2014). Thus, the two software packages were used in a complementary way, with Sketch Engine effortlessly bringing together the 28 hours of transcribed and annotated CoFEL data on a single platform, albeit without the multimedia files.

3.4.2 *Hesitation markers*

Hesitation is a common, ordinary feature of natural speech (Chafe, 1985; in Alsop, 2015, p. 38). Hesitation marking issues appeared very early in the transcription-annotation process of CoFEL lectures with ELAN. In terms of the corpus as a whole, the omnipresence of hesitation markers is a recurrent salient feature in all the lecture transcripts. In CoFEL, the choice was made from the start to unify the systematic transcription of hesitation with only two parameters: the marker itself and its duration. A choice had to be made between *uh* and *er*, and the latter was retained. The presence of the consonant “r” made it a more likely candidate for vowel sound sustaining or lengthening³⁵, a very common phenomenon in CoFEL hesitations (see sections 4.5 & 5.4 for transcript samples). Looking at Coventry University's ELC figures for comparison, and retaining only non-words, the amount of hesitation markers is 3.5% of the corpus tokens³⁶. In CoFEL, there are 7,505 occurrences of “er” out of a total of 154,279 tokens, giving almost 5% of all tokens, which is considerable. This leads me to posit that, as 77% of the ELC tokens come from lectures delivered in L1, from the UK and New Zealand (Alsop,

³⁵ N.b., the feature may concern any vowel sound lengthening in CoFEL transcripts, not just hesitations (see also Table 4). The transcription of hesitation sound lengthening in CoFEL is as follows: “er” for a brief sound (inferior to one second); “er:” between one and two seconds; “er::” above two seconds. The symbol used is inspired from the broad IPA (International Phonetic Alphabet) phonetic transcription of a word like *blur*, giving /blɜː/, without having recourse to special phonetic symbols in the corpus transcripts.

<https://www.internationalphoneticassociation.org>

³⁶ The ELC annotates “ah”, “er” and “um” as discrete elements. Estimates (Alsop, 2015, pp. 56 & 242) give 18,400 occurrences for the three markers, against 528,157 tokens for the whole corpus. The marker “well”, being a word, was not included in my count.

2015, p. 56), and 100% of the CoFEL lectures are given in EAL, lecturing in an additional language increases the frequency of hesitations significantly by putting a heavier load on the speakers' working memory. This outcome converges with that of Fehringer and Fry (2007) who showed that proficient speakers produced a higher rate of hesitations in additional languages, indicating a heavier cognitive load imposed on their working memory when using these additional languages. In some cases in CoFEL, hesitation is so widespread that it is likely to disturb the consistency and coherence of the lecture, as in the Soil & Rock Mechanics excerpt below, especially when combined with false starts and other disfluencies (all emphasized in bold), or divergent ELF forms (in italics):

so- it's more the job of **er** the geologist **to: er** investigate sites and determine the different **orient-** orientation of **er** discontinuities, but **er** from **er** *these information* **er**, we can have **er** graphical *representation* of **er** this **er** plane **er** *which are* **er** *the: stereographic representation.*

[20200213_ELO_SRM2_0280]

3.4.3 Sketch Engine

Sketch Engine is a corpus manager and text analyzer designed to search large collections of texts. Complex linguistically-motivated queries enable researchers to extract essential information from a corpus of up to one million words (or more, in advanced versions), in over 90 languages. It is a user-friendly corpus linguistics tool, widely used by language practitioners and researchers, including lexicographers, translators, and language instructors. Its name derives from one of its main features, 'word sketches': automatic one-page summaries of the grammatical and collocational behavior of a word. In the collocation search, a word co-occurrence analysis is performed to display the most frequent words in relation to a search word, which can be considered collocation candidates. Word sketches group collocations according to particular grammatical relations (e.g., subject, object, modifier, etc.). Figure 7 shows an example of a word sketch from CoFEL, using "power" as a search entry.

Figure 7

Screenshot of a Sketch Engine 'word sketch' example from CoFEL, with the lemma "power" as a search entry

modifiers of "power"	nouns modified by "power"	verbs with "power" as object	verbs with "power" as subject	"power" and/or ...
output ... the output power	plant ... power plants	transmit ... transmit the power	function ... power function	route ... power and signal routes
heat ... the - the heat power	converter ... a power electronic converter	regroup ... regrouping the power	provide ... power provide	rule ... power and signal rules
total ... the total heat power	system ... the power system	supply ... supplying power	generate ... generated by a power	heat ... heat and power
transmitted ... transmitted power	supply ... the power supply	model ... model a power	flow ... power flows	voltage ... voltage and power

Among the features available on the platform, some proved particularly useful in this study. The fundamental one was the corpus building and management function, which enables users to create their own corpus from uploaded texts. Once a corpus like CoFEL has been compiled, the platform provides practical functionalities from its dashboard. For instance, the researcher can: browse the corpus and view documents, folders and edit its metadata; make it bigger, by adding texts; download it or share it with other users; create and manage subcorpora, like the Summarizing or Storytelling subcorpora in CoFEL. Other Sketch Engine functions were used in the data mining phase, like frequency word lists, part-of-speech (PoS) tagging³⁷, or n-grams³⁸, notably in the identification and analysis of the summarizing function and its strings of text (Chapter 4),

The features above were useful in this study in combination with the concordance search feature, which finds examples of a word form, phrase, or complex structure in the corpus and displays the query in context. Search results can be sorted, filtered, and counted. As shown in Figure 8, the results are displayed with the search word or search expression (n-gram) in the middle, with their left and right contexts on each side. In this example, a 3-gram search was conducted with engineering terms (displayed in red) in the whole corpus, and 5 occurrences were found. The platform displays the left and right contexts of the utterance, enabling the researcher to quickly locate the query in the corpus. The query displayed in Figure 8 is an

³⁷ PoS tagging: the process of marking up a word grammatically in a corpus as corresponding to a particular part of speech. The platform uses a reference 'tagset', based on both definition and context. The PoS tag is the label assigned to a token (word) in the text corpus to indicate the part of speech and other grammatical categories such as tense, number, case etc.

³⁸ The 'n-gram' tool produces frequency lists of sequences of tokens: e.g., a 4-gram search will produce a list of the most common combinations of four tokens in decreasing order. The user has a choice of filtering options to specify in detail which n-grams should have their frequency generated.

example concordance line for an aeroelasticity engineering term, ‘the flutter phenomenon’, where KWIC refers to Luhn’s Key Word in Context extraction (Luhn, undated, in Manning & Schütze, 1999, p. 35):

Figure 8
Sketch Engine: screenshot of an example 3-gram concordance search result in CoFEL

<input type="checkbox"/> Details	Left context	KWIC	Right context
1 <input type="checkbox"/> ⓘ doc#0	are going to start the interesting part, which is	the flutter phenomenon	. all right? and today we are going to start the in
2 <input type="checkbox"/> ⓘ doc#0	e are going to start the interesting part which is	the flutter phenomenon	. all right? that gonna be f subsection.</s><s>(x
3 <input type="checkbox"/> ⓘ doc#0	lear probably the way i’ve– i’ve explained it but	the flutter phenomenon	is hidden here. so– it basically says that we hav
4 <input type="checkbox"/> ⓘ doc#0	plification of the vibration which corresponds to	the flutter phenomenon	. and here what we have said is that delta is eq
5 <input type="checkbox"/> ⓘ doc#0	ere will be only one frequency corresponding to	the flutter phenomenon	. if you remember the explanation that we have

In addition to the features presented above, the platform can also indicate what is typical of the selected corpus compared to a large reference corpus, such as the BNC³⁹ or OANC⁴⁰. To do so, Sketch Engine provides the extraction of key words and multi-word terms from texts, based on frequency count and linguistic criteria. Results can take several forms, for example, individual key words, key terms (multi-word expressions in a format typical of terminology in the selected corpus), or n-grams (any sequences of tokens for which only items which appear more frequently than in the reference corpus are included). These last features, in addition to the other ones above, made Sketch Engine an essential tool in this study.

3.5 CoFEL: corpus description

CoFEL includes fourteen engineering lectures from eleven different lecturers. The main characteristics of the lectures are listed in Table 5, in chronological order of recording (L01 to L14), with their generic file names, subjects, duration and length. The lectures were observed and recorded over a period of two years, between January 2020 and December 2021. They range across six broad disciplines: fluid mechanics, civil engineering, electrical engineering, mechanical engineering, automation, and energy systems. The corpus covers twelve specific lecture subjects, ranging from aeroelasticity to power systems, or from rock

³⁹ British National Corpus (BNC): collects written and spoken language of British English from the later part of the 20th century and comprises 100 million words. <https://www.sketchengine.eu/british-national-corpus-bnc>

⁴⁰ Open American National Corpus (OANC): includes all genres and transcriptions of oral American English data produced from 1990 onwards and comprises 11 million words. https://www.sketchengine.eu/oanc_masc-corpus

mechanics to cybersecurity. Some specific features stand out: lecture pairs L02-L03 and L09-L10 are consecutive and are delivered by the same lecturers, on the same subject, one week apart; L07 and L08 are delivered by the same lecturer, on distant dates, and cover different subjects. L05 and L06 are extended practical laboratory sessions lasting almost three hours and four and a half hours respectively, while the rest of the lectures last about two hours or less.

Table 5

Fourteen engineering lectures of CoFEL, in chronological order of recording

Lecture code	Generic file name (Date_Lecturer_Subject)	Lecture subject	Duration (hh:mm)	Length (tokens)
L01	20200117_FAC_AERO	Aeroelasticity	02:04	6023
L02	20200206_CAM_SRM	Soil and rock mechanics	00:52	7385
L03	20200213_CAM_SRM	Soil and rock mechanics	01:22	10172
L04	20200213_ELO_SRM	Soil and rock mechanics	01:35	7803
L05	20200407_AMI_SSD	Shear strength of sand	02:51	13905
L06	20200408_OFA_CSP	Characterization of soil properties	04:33	23731
L07	20200820_PAC_SPI	Smart power industry	01:31	11006
L08	20201008_PAC_EMS	Energy market structures	02:05	12896
L09	20201016_SAJ_EMG	Electromagnetic compatibility	02:00	15797
L10	20201023_SAJ_EMG	Electromagnetic compatibility	01:58	15946
L11	20210111_UPA_CYBER	Cybersecurity of industrial systems	02:00	16961
L12	20210318_LUY_THERM	Thermodynamics	02:03	14772
L13	20210506_DIM_RICE	Reciprocating internal combustion engines	00:38	4644
L14	20211208_TAY_DCS	Design and control of power systems	01:37	12380
Total			27:09	173421

Note. Key to generic file names: lecture dates YYYYMMDD; lecturers' names anonymized; subject codes according to course subjects.

The names of CoFEL files include metadata. Thus, a rich file name was preferred to a simple serial number, so that any section, annotation, or transcript from the corpus could be easily traced back to its original lecture source. When the longer lectures had to be divided into several parts for work in ELAN, a part number was added to the lecture subject code.

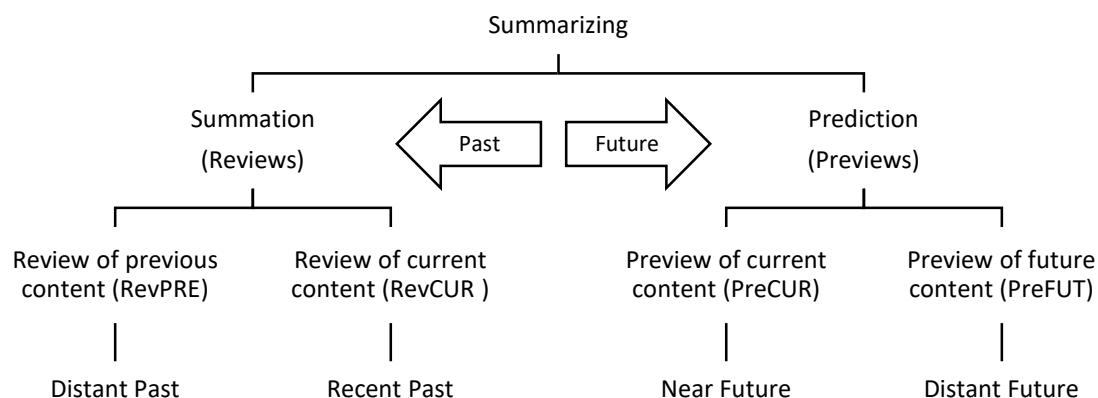
4 Summarizing

Communicating complex and rigorous engineering content to future engineers can be challenging, considering the heavy responsibility entailed by the profession, when designing a suspension bridge or building a nuclear power plant, for instance. Similarly, absorbing engineering content, often laden with intricate mathematical formulas, during sessions of two to four hours in a row, can be demanding for students. Lecturers are inclined to organize their discourse in such a way as to maintain focus as the lecture unfolds, while giving a clear, supportive structure to the knowledge imparted. Therefore, scientific or academic lecture discourse is likely to contain instances of summarizing: lecturers classically use previewing and reviewing to organize the delivery of their lecture content, conveying a roadmap and milestones to their students. Alsop (2015) defines a review as containing reformulated information or reminders of information already given either in the past or current lecture; a preview contains upcoming content either in the current or in a future lecture.

This chapter examines the pragmatic function of summarizing as it appears in monologic engineering lectures taught in English-Medium Instruction (EMI) environments. In this study, a summary is defined as an occurrence in a lecture when topical content is reviewed or previewed by the lecturer. Summarizing can happen at the beginning, at the end, and at more or less regular intervals throughout the lecture. Nesi et al. (Nesi et al., 2014) identify two main summarizing functions, summation and prediction. The former serves the purpose of reminding the audience of the ground that has been covered; the latter is destined to herald what lies ahead in the lecture. Summation is divided into two types of reviewing: the review of previous lecture content (RevPRE) and the review of current lecture content (RevCUR). For prediction, two types of previewing are also defined: the preview of current lecture content (PreCUR) and the preview of future lecture content (PreFUT). Figure 9 illustrates this organization.

Figure 9

Breakdown of summarizing into two main functions and four types



Note. Adapted from Nesi et al. (2014) and Alsop (2015).

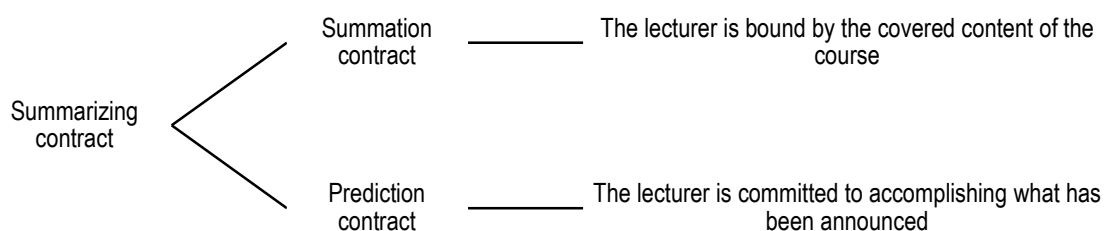
4.1 Conceptual and research background

Summarizing in oral academic discourse has drawn less attention than its written counterpart, and while parallels can be drawn, the natural language of spoken discourse is usually less neat and well-organized. It contains hesitation, interruption, rephrasing, repetition, pauses, and formulaic language. Speakers monitor what they say in real time, and this monitoring will often become part of the discourse. An abundance of sounds, words and phrases with no apparent meaning (*er, so, well, actually, okay, etc.*), and linguistic devices (repetitions, pauses, etc.) signal how the speaker is interpreting the ongoing discourse (Sinclair, 2004). This may be true for lecturing in one's native language, and may be even more likely in EMI, where lecturers teach complex engineering content in an additional language.

Regardless of the language used, lecturers are committed by what they announce (Tadros, 1994). With summation, or reviewing, the lecturer is bound by the actual content of the course that has been delivered so far, and with prediction, or previewing, the lecturer is committed to accomplishing what is announced. In both cases, a contract is established between the lecturer and the student audience (Figure 10), and the lecturer undertakes to fulfill this contract. Sinclair (2004) refers to the partial or total fulfilment of the retrospective or prospective contracts drawn. He also introduces the concept of a directionality of discourse and its associated previous and subsequent language.

Figure 10

The contract established between the lecturer and the students when summarizing



Note. Adapted from Sinclair (2004) and Tadros (1994)

In this section, I begin by describing some of the main features of the summarizing function in general, then some of the more specific features of summation and prediction. I will continue with the scaffolding role of summaries in the overall lecture structure, before addressing our research questions.

4.1.1 Deixis

A first feature common to all summaries is deixis. Deixis is a technical term for one of the most basic roles that utterances have, which is pointing via language: identifying the people, objects, events, processes or activities referred to (Lyons, 1977; Yule, 1996). As such, deixis plays a key pragmatic role in summaries of oral discourse. Reviews correspond to a retrospective or backward temporal deixis, pointing towards preceding lecture content depending upon some antecedent oral text. Previews correspond to a prospective or forward temporal deixis, pointing towards proceeding lecture content depending upon some postcedent oral text. Deixis occurs when a word or a phrase refers to a place in the discourse (discourse deixis), a time frame (temporal deixis), or a person (person deixis): a vital connection between time, discourse and the participants involved is established (Huddleston & Pullum, 2002). It creates points of intersection, a reference between language and the social setting (Hanks, 2011; Lyons, 1995). It describes the way the meaning of words is bound to the situation of the utterance: if one is not aware of that situation, one cannot decode the meaning of the utterance (Crystal, 2008; Peters, 2004).

Discourse deixis has been defined as the matrix of linguistic material in which the utterance plays the role of preceding and following other parts of the discourse (Fillmore, 1997), as a co-reference or reference to an earlier discourse (Hanks, 2011), or as deixis in text that

unfolds in real time, for both speaker and addressee (Marmaridou, 2000). As the lecturer's discourse unfolds in time, temporal deictics are commonly used to refer to portions of this discourse, creating a connection between discourse and time deixis, a possibility of referring from one utterance to others, backwards or forwards (Levinson, 1983, 2001).

With temporal deixis, the lecturers' *now*, and their location, form the deictic center, also termed 'origo' by some authors: the zero-point of the spatiotemporal co-ordinates of the deictic context (Lyons, 1977; Yule, 1996). In the EMI lecture hall, everything is related to the lecturers' viewpoint in an egocentric situation: proximal (near the lecturer: *this, here, now*), and distal (away from the lecturer: *that, there, then*). Temporal deixis refers to the events mentioned in utterances, relatively to the lecturing time. It appears in the form of conjugation tenses, temporal adverbial deictics (*yesterday, tomorrow, tonight, next time, last year*, etc.), or prepositional phrases (*at noon, in the morning, on Tuesday*, etc.). The temporal deictic function relies heavily on conjugation tenses, which anchor utterances in their context deictically (Levinson, 1983). The lecturers' origo (here and now), controls tense, which falls into three categories: the moment of lecturing, the Present; earlier than the moment of lecturing, the Past; later than the moment of lecturing, the Future (Lyons, 1977; Matthews, 2014). English has only two basic forms, the Present and Past tenses, which are morphologically differentiated, with the Present tense as the proximal form and the Past tense as the distal form (Matthews, 2014; Yule, 1996). The temporality of the Future tense is imbued with modality: it is never a purely temporal concept, necessarily including an element of prediction or a degree of uncertainty (Lyons, 1977; Matthews, 2014).

Person deixis, in the EMI context, refers to the properties of utterances determined by the identity of the interlocutors, namely the lecturers and their students (Fillmore, 1997). Above all, it concerns the use of pronouns, which are familiar ways of encoding the participants' roles in utterances. The use of personal pronouns, especially interpersonal ones, is a key feature of summaries, and pronouns may be used as an indicator of student engagement (Yeo & Ting, 2014). First person encodes the lecturers' reference to themselves (means lecturer inclusion); second person encodes one or more student addressees (means addressee inclusion), and third person, refers to persons who are neither lecturers nor addressees, thus lecturer and addressee exclusion (Levinson, 1983). First-person plural *we* has a special status, and *we* is not the plural of *I*: when *we* includes the addressee, it is said to be inclusive; if not, it is exclusive of the addressee (Lyons, 1968). Particularly relevant to the lecturing situation is the lecturers' use of pronouns *we* and *you* in the presentation of information in summaries, with a dual role: first, to

highlight and reinforce the student audience's awareness of a shared context; second, to make up for the lack of interaction between the lecturers and their students (Hansen & Jensen, 1995).

4.1.2 Discourse markers

Discourse markers (henceforth, DMs) are a second feature common to all summaries. In EMI lecturing contexts, it is widely recognized that learners can have difficulties in following the structure of lectures. This has inspired studies on the role of DMs and macro-discourse organization in lecture comprehension (e.g., Allison & Tauroza, 1995; Thompson, 2003, for an overview). DMs in oral summaries can play a relational role, a signaling role, and a discursive deictic role. As their name suggests, their roles are better described at the discourse level rather than the sentence level, and their meanings is better analyzed for what they indicate or mark rather than for what they describe (Blakemore, 2002). Furthermore, there is disagreement on their precise definition, categorization, meaning and function, and they have been given various labels, such as discourse connectives, discourse operators, pragmatic connectives, sentence connectives, or cue phrases (Fraser, 1999; Schourup, 1999). Certain forms catalogued as DMs are found mainly in oral rather than written discourse, such as lectures (e.g., *so, now, by the way, well, okay, right, after all, I mean, you know, mind you, etc.*).

DMs can influence foreign students' understanding of lectures, because for them the structuring and organization of information may be even more fundamental than for native speakers. Studies have showed that students understood the lectures better when DMs were included than when they were deleted (Eslami & Eslami-Rasekh, 2007; Flowerdew & Tauroza, 1995; Pérez & Macià, 2002). They can also be used by lecturers to compensate for the lack of interaction with the audience, and alongside pauses to signal changes in topics or viewpoint, to evaluate or give directions (Hansen & Jensen, 1995).

Being able to recognize connections, relationships, and hierarchies in the lecture content are some key aspects of the active listening skill, provided these DM signals are given in a timely and appropriate manner by the lecturer. It has been shown that, in an English as an additional language (EAL) context, the high- and low-level information in lectures was much better recalled when DM cues oriented students to the overall structure of the lecture, as the content was explicitly previewed, enumerated, or summarized by the lecturer (Jung, 2003). A distinction can be made between micro- and macro-markers: micro-markers (*so, ok, right, etc.*) have a low semantic value and mainly allow the speaker time to plan the next utterance; macro-makers are explicit, elaborate expressions of lecture information planning (Chaudron &

Richards, 1986). While micro-markers do little more than allow the lecturer to reflect and plan the next utterances, and sometimes blur the message, macro-markers require phrasing attention to serve as signposts, signals, or milestones that will enhance students' comprehension in effective listening.

4.1.3 Information packaging and enumeration

Information packaging and enumeration are a third feature common to all summaries. Information packaging refers to the way elements of lecture content are packaged and structured within clauses (Biber et al., 1999; Huddleston & Pullum, 2002). Information packaging is one of the hallmarks of summaries, as can be found in EMI lectures, notably through the grammatical structure called *wh*-clefts or pseudo-clefts (Deroey, 2012). Pseudo-clefts are rather common in oral discourse (Collins, 1987) and tend to occur at strategic points in lectures, namely in lecture introductions, at the beginning of sections or in explanations (Deroey, 2012). Deroey also emphasizes that the structural features of the clauses show a systematic preference for the active voice and for the inclusive pronominal subject *we*. These clauses usually signal a turning point or introduce a new point of interest, thus playing an essential pedagogical and structuring role.

Students can benefit from a lecturer's summative or predictive signals, and identifying and decoding these signals is a key listening skill in all types of academic discourse (Alsop, 2015). Tadros (1989, 1994) proposed a framework of the predictive aspect of text, emphasizing its pedagogical value, and one of her predictive categories is enumeration. Enumeration is a type of prediction in which the predictive item carries a signal which prompts the lecturer to name a list of several predicted points one by one, that usually belong to the same functional category (Dubois & Sankoff, 2005; Tadros, 1989). It can outline the day's session or the breakdown of upcoming course content. In EMI lectures, enumeration is not limited to prediction and can equally be found in summative summaries when lecturers recapitulate course content delivered previously. When enumeration is used together with the inclusive *we*, lecturers strengthen interpersonal relationships and improve comprehension by giving clear, structuring signals.

4.1.4 Salient features of summative summarizing

Summative summarizing breaks down into two types of reviews: reviews of previous and current lecture content. Reviews correspond to a retrospective or backward deixis, pointing

towards preceding lecture content depending upon some preceding oral text. Anaphora is the process where a word or phrase, the anaphor, refers back semantically to another word or phrase, the antecedent, which was used earlier in a text or conversation, and with which an anaphoric link, or relationship, is created (Richards & Schmidt, 2010). Anaphora and backward deixis have much in common, as certain word forms or markers can be both deictic and anaphoric, but whereas deixis generally refers to the extralinguistic situation, anaphora refers to the text itself (Huddleston & Pullum, 2002). In EMI lectures, summative markers can be identified at the level of single, or groups of lexical items and help to signal how content is organized in stretches of discourse (Alsop, 2015). For example, the use of the definite determiner *the* implies a de facto anaphoric reference. Anaphora often corresponds to an articulation in academic lectures, for while it summarizes a complex concept or process, it also heralds the next related content (Alsop, 2015). It is a key element in making discourse cohesive, and communication coherent: it is the most common way to hold a text together, and most word categories can be found in the anaphoric chain of reference (Hands et al., 2017; Peters, 2004).

Retrospective labelling is one case of anaphora. Summative summarizing requires information to be recalled from elsewhere earlier in the text, and is the post-positing closure of that already-given information (Francis, 1994). By assigning labels to noun phrases, lecturers can then use them to link and organize their discourse. The retrospective label serves to encapsulate, or package a stretch of discourse. As such, it is not a repetition or a synonym, but an equivalent of the replaced text, with a name (the label) that ensures the intelligibility and cohesion of the discourse. EMI lecturers may use labels to signal that they are about to move on to the next stage of their argument. The lecturer often disposes of the preceding stage by encapsulating it in a single label, a shell noun (Hunston & Francis, 2000; Nesi & Moreton, 2012), whose meaning is supplied through an anaphoric relation to the past act of doing an exercise, or of watching an instructional video. In academic contexts, scientific or engineering disciplines may require more precise anaphoric devices than other registers, which can be provided by a relatively high frequency of noun phrases with demonstrative determiners (Biber et al., 1999).

4.1.5 Salient features of predictive summarizing

Symmetrically to summative summarizing, predictive summarizing breaks down into two types of previews: previews of current and of future lecture content. Previews correspond to a prospective or forward deixis: they point to postcedent lecture content. However, the role

of verbs in the utterances differ: summative utterances stem from the past and are built on facts, while predictive utterances, which deal with the future and unfulfilled actions, are tinged with modality. Tadros (1989, 1994) proposed a framework of the predictive aspect of text, emphasizing its pedagogical value. Following previous research in the EMI engineering lecturing context (Alsop, 2015; Nesi et al., 2014), four of Tadros's functions are expected to be particularly pertinent: enumeration (addressed in subsection 4.1.3), advance labelling, recapitulation, and question.

Research has long shown that the prior introduction of new concepts can facilitate learning and retention (Ausubel, 1960), by contributing to bridge the gap between what students already know and what they need to know. Ausubel's advance organizers work on the basis of two factors: the first factor increases the familiarity and meaningfulness of the pedagogical task; the second one provides optimal anchorage for the learning material to come (*I will start my lecture with . . . ; you can build different relationships, and this is what we are going to do now*). More recent research has favored the term advance labelling (Francis, 1994; Tadros, 1994). When lecturers use the pragmatic function of advance labelling, they tell their students what to expect, thus aiding the learning and retention of meaningful verbal material. The function is distinctive of predictive summarizing.

The function of recapitulation is to begin a section of the lecture by predicting new information from what has already been said, in a predictive-predicted structure (Tadros, 1989). Such an utterance is often articulated around an adverb like *then*, separating the predictive part, for example an if-condition, from the concluding predicted part expressed by the adverb. A Present Perfect tense at the opening of the utterance will refer to the past actions, while in the predicted part the Future tense will be found, with the modal *will* (*we have seen that, if the two variables . . . , then we will have an occurrence of the phenomenon*). The mathematical demonstration will be complete provided the predictive if-condition is satisfied.

The fourth and last of Tadros's function is (rhetorical) question. Rhetorical questions do not work through interaction (*So, where do we go from here? Of course, we need to . . .*). When lecturers utter them in their monologic lecture discourse, they intend to provide the answer themselves (Peters, 2004). The rhetorical question function concerns both the detachment and involvement of the lecturers. Although the lecturers detach themselves from the question they are asking, this detachment suggests that they will intervene at a later date to give their view on the matter (Tadros, 1994). Khan and Aguilar-Pérez (2023), examining the

metadiscursive role of rhetorical questions, found two categories: macro questions marking a topic shift, and micro questions marking the lecturer's navigation within the topic. Khan and Aguilar-Pérez argue that, on the one hand, they engage students and maintain their attention; on the other hand, rhetorical questions perform both textual and interactional metadiscursive functions, organize discourse, and deconstruct complex knowledge.

Inchoation (or inception) refers to an action that is about to begin, or that is beginning (Peters, 2004). Inchoative structures complete the lecturer's flow of course content delivery by giving such metadiscursive information. In EMI, it is an instant call to action from the lecturer, which implies immediate audience involvement (Crawford Camiciottoli, 2004). The student audience is involved by means of the inclusive pronoun *us*, comprising both lecturer and students (*let us/let's resume our topic; let's start from the beginning*). Inchoation can have a positive effect on lecture comprehension in EAL settings and student engagement (Crawford Camiciottoli, 2004).

Utterances expressing volition concern what the lecturer wants to do, or wants to be done by the students (*you will calculate one point corresponding to one location*). They usually contain the modal auxiliary *will* and use proximal forward time deixis, with implicit or explicit temporal references (Olsen & Huckin, 1990). The auxiliary constructions with *will* can express more than predictions and convey the determination that something should happen, in an expression of modality (Peters, 2004). Volition in EMI lectures is also found in distal forward deixis utterances, characteristic of references to future lectures, where more challenging topics are left for a later date (*this is very complex, and we'll see that together next year*).

Structures conveying intentionality give information about the current upcoming lecture content and convey the purpose of the lecturer, with the *be going to* structure (or the more casual *(be) gonna*). The quasimodal verb *be going to* is the most straightforward means English has to express futurity, with no particular implications that restrict its use (Peters, 2004). This means that a rather high frequency of occurrence of such forms can be expected in the EMI lecturing genre, when expressing intentions (*calculating the value of . . . is what we're gonna/going to do now*).

4.1.6 *Scaffolding the lecture: from micro to macro structures*

Lecturers use structural cues, signposts, and summaries which, when timely and appropriate, clearly signal the articulation of sections, scaffold the lecture, shape and facilitate

their audience's comprehension and retention of information. Some researchers have attempted to identify a macro-structure model for lectures (e.g., Connor & Mauranen, 1999; Connor & Upton, 2004; Flowerdew, 1994a; Kanoksilapatham, 2005; Straker Cook, 1975; Swales, 1990). However, results in this area have been rather unsatisfying for our research purposes: a problem-free, predictable structural model, comprising a set of categories for a functional description of the lecture discourse, remains to be found (Alsop, 2015).

A number of studies have found that macro-markers lead to the students' better recall of lecture content than micro-markers and that, conversely, an inability to recognize the macro structure of a lecture impairs the students' understanding, or claim that rhetorical form is more important than content (Carrell, 1987; Chaudron & Richards, 1986; Flowerdew, 1994b; Lebauer, 1984; Malavska, 2016; Wijasuriya, 1971; Young, 1994). Young did comprehensive work on oral lecture structures, examining the micro-structural features in relation to lecture macrostructure, and her oft-cited work on lecture structure divides lectures into six phases. Her results show that each of these phases is repeated discontinuously and intertwined with others throughout the sessions, forming a complex, variable macro-structure (Young, 1994). In a study on lecturer strategies, Miller (2009) found that lecture staging (the division of the lecture into identifiable parts, or stages) and mapping (the stages organized into a structure, handed out or projected) facilitated student comprehension. Malavska (2016) explored students' preferences in EMI lecturing styles, and the lecturers' own style choices. She used a holistic approach to analyze lectures: opening, main text, and closing. She proposed a four-move model of lecture macro structure: warming up, setting up the lecture framework, putting the topic in context, concluding the lecture. Although Malavska's model may be appealing, her results are based on a limited corpus of seven lectures and show some fluctuations and repetitions in the fourteen steps that she identified within the moves.

4.2 Research questions

With regard to summarizing in EMI engineering lectures, I propose to address the research questions below.

- How can summaries and their boundaries be identified in EMI pedagogical situations?
- How can summaries, once found, be categorized?

Then, in order to investigate the quantitative summarizing data in CoFEL, I propose to address the next research questions.

- What is the amount of summarizing activity in lectures, and what is the number and duration of summarizing episodes?
- Does the duration and token length of summarizing episodes vary according to their type?
- For consecutive lectures given by the same lecturer on the same subject, are there any differences in summarizing between the first and the subsequent lectures?
- Do summary types depend on the location of the summarizing episodes in lectures?

Finally, in the light of our findings, I propose to examine our non-quantitative data and address the last research question.

- What is the pragmatic and linguistic variation of summaries in French EMI engineering lectures?

The next section presents the method used to locate, delineate, and categorize the summarizing instances that I hypothesize exist in the engineering lectures of CoFEL.

4.3 Method

In order to explore the role of summaries in CoFEL, I propose the following method. First, the summaries and their boundaries need to be identified. For this purpose, I have built and implemented an analysis grid, as explained in subsections 4.3.1 and 4.3.2. Next, an example of how the grid is used is given (subsection 4.3.3), with the identification of reviews of previous content. Then, the categorization process of summaries into one of the four types is explained, (subsection 4.3.4). Finally, I present the visualization solution that I developed to further explore, understand, and share our data (subsection 4.3.5).

4.3.1 Building an analysis grid

Building an analysis grid involves the identification of subsets of criteria that should be mutually exclusive and collectively exhaustive (MECE—Minto, 2021). The design of the analysis grid for CoFEL consisted of two steps. The first step involved a close examination of the work carried out by Nesi et al. (Nesi et al., 2014) and Alsop (2015), who analyzed the summarizing pragmatic function in the Engineering Lecture Corpus (ELC, 2014), the corpus most akin to CoFEL. Salient lexicogrammatical terms and phrases from their work were extracted to form the basis of our analysis grid for the identification and analysis of summarizing data⁴¹. The second step consisted in supplementing these items with converging ones from the CoFEL data-collecting and corpus-building pilot phase (the 2019-2020 academic year), using the ELAN software. Then, all the collected lexicogrammatical items were classified into n-grams, from 1-grams to 4-grams and above, under the four summary types: the review of previous lecture content (RevPRE), the review of current lecture content (RevCUR), the preview of current lecture content (PreCUR), and the preview of future lecture content (PreFUT). Next, they were grouped according to their grammatical categories (nouns, adjectives, verbs, etc.), while excluding all the duplicates. Finally, using these items, the analysis grid was designed (Table 6) and broken down into one grid per summary type. These 4 sub-grids were used to retrieve all the ‘candidate summaries’ from CoFEL (strings of text that are likely to be summaries), and were located with Sketch Engine (Kilgarriff et al., 2014). Later, these candidate summaries underwent an identification-assessment stage performed by three independent raters, as described in the next subsection.

⁴¹ The detail of the original ELC extracted data in this early phase is available in Appendix D, Appendix E, and Appendix F

Table 6

Analysis grid for retrieving candidate summaries from CoFEL, according to the 4 summary types

Review of previous lecture content (RevPRE)	
1-grams	summary
	former, latest, preceding, previous
	ago, already, before, briefly
	recap, remember, review, sum, summarize
2-grams	considered, demonstrated, described, did, discussed, explained, had, reviewed, said, saw, seen, spoke, started, stopped, studied, talked
	last Monday, last Tuesday (etc.), last time(s), last class(es), last session, last week(s), last month(s), last semester(s), last year(s)
3- to 5-grams	i'll say this once more, i always find it hard to explain, i was showing you (last week), i showed you (last time)
	what we did/we've done (the last two weeks), what i/we said, what you/we learned/have learned, what we've covered/been covering
	(a bit) more difficult, (it's/it was) the hardest (part/module), (it's going to get/it got) get (quite) complicated, (it's/it was) not (that) easy (to understand), we had a/this problem
	we looked at, we have seen that, we saw that, we have defined, we defined, (the work) we've done so far, let's make sure that
Review of current lecture content (RevCUR)	
1-grams	complex, difficult, earlier, major
	compare, highlight, remind
	mentioned
2-grams	just now, so- again,
	it's complicated, note this
3-grams	i repeat, i've explained, we already
	the most important, not too much
	going back to, please remember that
	that is how, that's for the, that was just, this ends the
4-grams or above	we have considered, we have done, we have given, we have said, we have seen, we have studied a few moments ago
	(it is) very important
	that was maybe a bit hard, i hope it was not too (difficult), i hope you understood
	what we (have) said, what we discussed today
Preview of current lecture content (PreCUR)	
1-grams	example
	look
2-grams	carefully, today, together
	how to, it's just, let's see, now let's, (let's) move onto
3-grams	a little bit, a little problem, just a little
	the interesting part, the outline of
	to move on, will be explained
	i propose to, i will explain, i will focus on, i will show, i'm going to, now i'm starting
4-grams or above	we will discuss, we will examine, we will see, we'll try to, we're going to
	the aim is to, the first part of the lecture, the outline of this lecture is about
	look a little more closely at, that will be the main, the next part will be
	i will ask you, i will finish with, i will just present, i will start with, i would (just) like to
	now we will start the, we will figure out, we will have to
	what comes next is, what is actually being, what we are going to, what we have to, what we will do

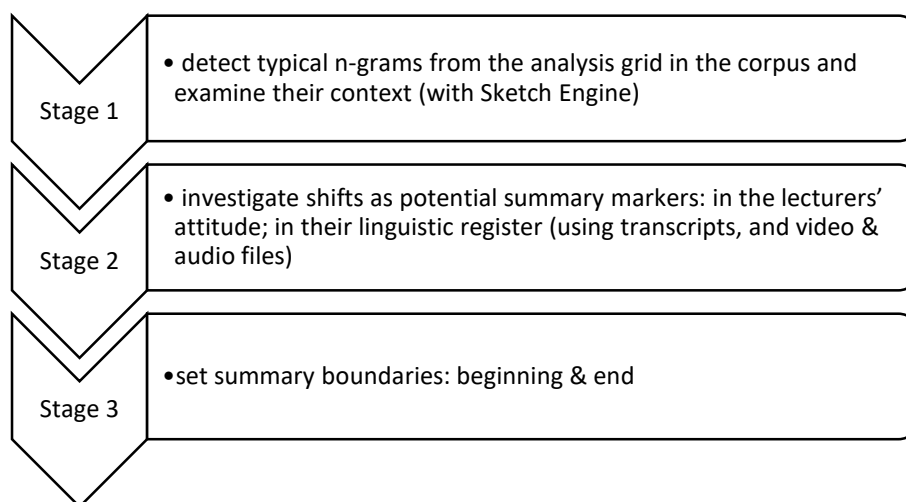
Preview of future lecture content (PreFUT)	
1-grams	later
2-grams	next week, next month, next semester, next year
3-grams	after the holidays, how we can, it'll be in (January, February, etc.), will be used
	we will apply, we will have, we will see, we'll be using, we'll focus on you will do, you will understand
4-grams or above	in the next lecture, in your 2nd/3rd year
	it will depend on, remember that you have, where are we going?
	we still have one session, we will (not) deal with, we will be doing, we'll be able to, we'll see that together
	i'm going to do, it is going to concern, we're going to determine, we're going to figure out, we're going to study, we're going to continue with you're going to examine, you're going to start dealing with

4.3.2 Identifying candidate summaries

Identifying summaries in CoFEL involves both broad and narrow corpus data mining. The identifiers of summaries vary with each of the four attributes identified in our analysis grid, as shown in Table 6. Firstly, searching the corpus broadly for summaries with the ELAN software (ELAN, 2023) was done during the early transcription-annotation process of the corpus. Secondly, the narrower, more systematic summary identification-confirmation process involved three stages and was undertaken after the completion of the corpus (December 2022): (1) locating the key elements in the corpus, with Sketch Engine (Kilgarriff et al., 2014), using the n-grams in the analysis grid as search words and phrases, and examining their context; (2) locating shifts in the lecturers' attitude, in their linguistic register or pragmatic function used, with recourse to the audiovisual recordings, as necessary; (3) determining the precise boundaries of a summary: where a summary begins and where it ends. Figure 11 recapitulates this three-stage process used in the identification of summaries.

Figure 11

The 3-stage process to identify and delineate the four summary types in CoFEL



The temporal anchoring of the summarizing function makes it possible to start pre-categorizing examples just after their retrieval from the corpus. However, the conjugation tenses of the verbs used by the lecturers may not always be reliable: in ELF, inflectional suffixes or auxiliaries are often divergent, and omitted or not pronounced by EAL speakers. Verbs are often rendered in their bare infinitive forms instead, or mimic the lecturer's native language patterns divergently (e.g., a French *Passé composé* systematically rendered as a Present Perfect, whereas its main English equivalent translation is the Simple Past). Thus, a higher degree of accuracy in the summary categorization process can be achieved by combining verbal information with the more dependable temporal markers (*last week, next session, etc.*).

Subsection 4.3.3 will cover the RevPRE category in detail, as an example of the process. The other three categories were identified and categorized in the same way, but with their own search data sets from the analysis grid. The process being repetitive, it has not been detailed in this section.

4.3.3 An example: identifying RevPREs

The first stage consists in detecting typical 1-gram lexical markers. Table 6 (top section) displays the first lists of these terms that were used to identify passages in the corpus that might potentially function as RevPREs. Such summaries deal with past events by definition. However, although they are conjugated in the Present tense, 1-gram verbs like *recap(-itulate)* or *remember* also make an anaphoric reference to previously delivered content, as does the noun *summary* and adverbs like *ago* or *already*.

Using the key word in context (KWIC) concordance findings helped considerably in the identification, disambiguation, and categorizing processes of the summarizing instances. Sketch Engine lemmatizes the search terms automatically, subsuming all the formal lexical variations of the lemma (*talk*, for example, subsumes *talking*, *talks* and *talked*). Three KWIC extraction results with some of the terms of Table 6 are provided as examples below. Table 7 illustrates how 1-grams reveal the co-occurrence of other temporal deictic markers in their close context, like *remember*, *recall*, *last time*, or *the last class*, confirming the classification of the returned strings of text as belonging to the RevPRE category.

Table 7

Examples of a 1-gram KWIC concordance extraction from three CoFEL engineering lectures

Left context	KWIC	Right context
remember we have er two class left. [...] and today I just	summarize	the most important result that we obtained. so- er last time [...] er we started to study er the intact rock <small>[20200213_ELO_SRM]</small>
ok. so- er well- the last class er we had was er more than one month	ago	i will er start again- er er the beginning of the- the last class <small>[20210111_UPA_CYBER]</small>
so- just to recall you what i	explained	to you last time concerning [...] rock mechanics <small>[20200206_CAM_SRM]</small>

To increase the accuracy of the KWIC 1-gram extraction results, it is useful to complete it with a 2-gram search. Table 8 displays three representative KWIC extraction results with some of the 2-grams as examples. The 2-gram searches returned many more accurate RevPRE findings than the 1-grams. These anaphoric deictic markers often contain the adjective *last* in RevPREs, indicating various degrees of remoteness from the present. Also revealed are instances of divergent ELF verbal conjugation, like the Present Perfect *have seen* or *have done* (left context column) instead of the Simple Past *saw* or *did* in this context (n.b., the last line contains the French abbreviation *t-p* (*travaux pratiques*), for laboratory classes, which is an instance of code mixing).

Table 8

Examples of a 2-gram KWIC concordance extraction from three CoFEL engineering lectures

Left context	KWIC	Right context
you make the measurement of the water content and we have seen	last time	that gamma d is equal to gamma a multiplied by one [20200213_CAM_SRM]
all clear? do you remember what we have done	last week	? more or less? we have started to study the dynamic aeroelasticity [20200117_FAC_AERO]
we can do the test with water and we can do the test without the water.	last year	the students that did er this t-p* were doing the test without water [20200407_AMI_SSD]

A final round of even more targeted searches of 3- to 5-grams potentially found in RevPREs was carried out successfully. Table 9 provides a few examples illustrating a KWIC concordance search using some of these terms.

Table 9

Examples of a KWIC concordance extraction (3- to 4-grams) from four CoFEL engineering lectures

Left context	KWIC	Right context
it's exactly the situation where i never remember the formulas.	i always need to	make a demonstration. ok? [20200213_CAM_SRM]
so- it's already complicated to solve. and er it's	what we did	propose is one way to solve risk analysis for complex and interdependent infrastructures [20200820_PAC_SPI]
but it's very quickly become er a- er something that is complicated and the trustability is	not that easy	. so- er i think that most probably we need to couple tools- because they are tools that are very strong to do one thing [20200820_PAC_SPI]
this er ratio here- depends a lot on the initial cracks that we have inside the sample. er in the intact rock	we saw that	depending on the minerals- when the minerals behavior are quite homogeneous and uniform, we don't have many initial cracks in the material [20200213_ELO_SRM]

The cumulative findings of the above searches should represent a majority of the RevPRE candidate summaries in the corpus. However, in order to be as exhaustive as possible, or to confirm certain equivocal findings, a second stage was implemented: shifts were identified in the lecturers' attitudes, or linguistic register, so as to investigate their potential use of summaries. For example, audiovisual recordings were used for identifying instances where a lecturer moves from direct address to the audience, with eye contact, to launching a new slideshow from their laptop computer. Such moves are often accompanied by discursive changes: from a rundown on previously learned concepts (*last time, we did...*), to the synopsis of the new section (*today, we will...*). Sometimes, short bursts of RevPREs are found embedded in the main delivery of course content (e.g., *of course, you will remember this from the January lecture*).

Finally, in a third and last stage, the boundaries of the candidate summaries were established. Starter-markers like *right*, *now*, or *ok* seem to be frequently associated with summaries, although not exclusive to them. Common end-markers of RevPREs are *ok?* or *all right?*, their rising intonation being an invitation for students to ask questions.

4.3.4 *Categorization of summaries*

Temporal anchoring is a key feature of the summarizing pragmatic function. Thus the temporal references and conjugation tenses used in the lecturers' summaries often made it possible to allow an initial categorization of the summaries simultaneously or rapidly after their retrieval from the corpus. Individual raters reviewed the candidate summaries, then compared their results and decided together: (a) whether to include or exclude them; (b) to categorize them into one of the four types; (c) to decide their beginning and end. The first round consisted in using the analysis grid (Table 6) and the KWIC concordance searches with Sketch Engine to exhaustively identify all candidate summaries in the corpus. Ninety-nine candidate summaries were found.

In the second round, the candidate summaries were presented in a random order to three raters. These raters validated the cases where they agreed, and listed the uncertain cases to be discussed further, in a third round. Out of the ninety-nine candidate summaries, sixty-eight posed little or no problem and were quickly categorized into one of the four categories. For a third of the summaries (31 of them) one rater or more disagreed.

In the third and last round, the three raters decided on the remaining summaries. Whenever fine-grain disambiguation was required, an in-depth examination of the front or back context of the summary was carried out, or the position of the summary in the lecture was examined. The five types of uncertain summaries identified were: (1) "Not a summary" (5 instances); (2) potential "ELF issues" (5 instances); (3) "Uncertain types", with context to be checked (11 instances); (4) "Chaining instances: standard" (14 instances); (5) "Chaining instances: embedding" (1 instance). The last two types concerned candidate summaries which contained several summaries in one. This fine-grain examination meant that the final number of uncertain cases increased to a total of thirty-six, raising the agreed figure to 104 summaries in the end.

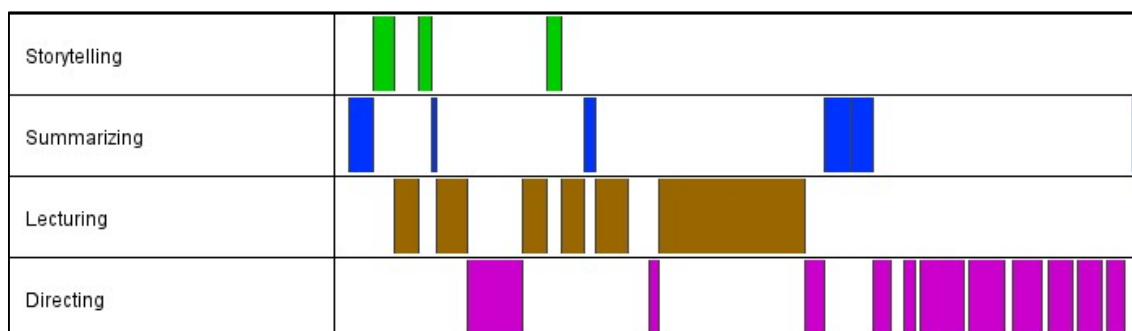
4.3.5 *Summary data visualization*

This subsection presents SumVIZ. It is a basic summary visualization solution which was instrumental in exploring CoFEL summaries, that I developed. The visualization of textual corpus data may help to explore and analyze textual data, as well as share or communicate on results (Allen, 2017; Card et al., 1999; Siirtola et al., 2014). SumVIZ has its origin in the Annotation Density Plot (ADP) of the ELAN software (ELAN, 2023), and in Alsop's data visualization tool (2015), ELVis. However, unlike these two tools, SumVIZ does not depend on a computer program or language, other than a suite of standard office software with chart design capability. Before presenting SumVIZ in detail, brief overviews of ADP and ELVis are given.

ELAN's ADP functionality schematizes the distribution of annotations resulting from the segmentation-annotation processing of an audiovisual file. The distribution of the annotations can be displayed at various adjustable levels of granularity. The tool displays the annotation density of one or several "annotation tiers" (in ELAN terminology), in the form of superimposed horizontal column graphs, with one graph per tier (see example in Figure 12). The horizontal axis represents the elapsing lecturing time, from left to right. The width of each column represents the duration of the annotation. The column height is arbitrary. ELAN offers a limited range of options to adjust the visualization size and aspect beyond the default settings. While useful for showing annotation densities, the tool has significant limitations and is not self-contained: annotation values are absent, and column graph units or axis labels cannot be displayed. Another limitation lies in the difficulty ELAN has in managing large multimedia files at once, which often makes it difficult for the ADP to display the whole lecture in one graph. For example, the longest CoFEL lectures had to be divide into 30-minute parts, but with all the quantitative data always remaining available from the ELAN software.

Figure 12

Example of an Annotation Density Plot (ADP) in the ELAN software



Note. Selection of annotation tiers from a Smart Power Industry lecture. Display range: whole lecture. Granularity level: Pragmatic Function (low-level of detail).

[L07 – 20200820_PAC_SPI]

Alsop (2015, pp. 95–106) developed and implemented the ELVis data visualization software tool to investigate the macro-level patterns of the Engineering Lecture Corpus (ELC - Nesi et al., 2014). The decision to develop ELVis was taken when a review of existing software failed to provide satisfactory visualization solutions. However, the software appears not to have been used in other corpus studies and is not publicly available. The relatively small scale of CoFEL (14 lectures), compared to ELC (76 lectures), made the investment in dedicated software development questionable.

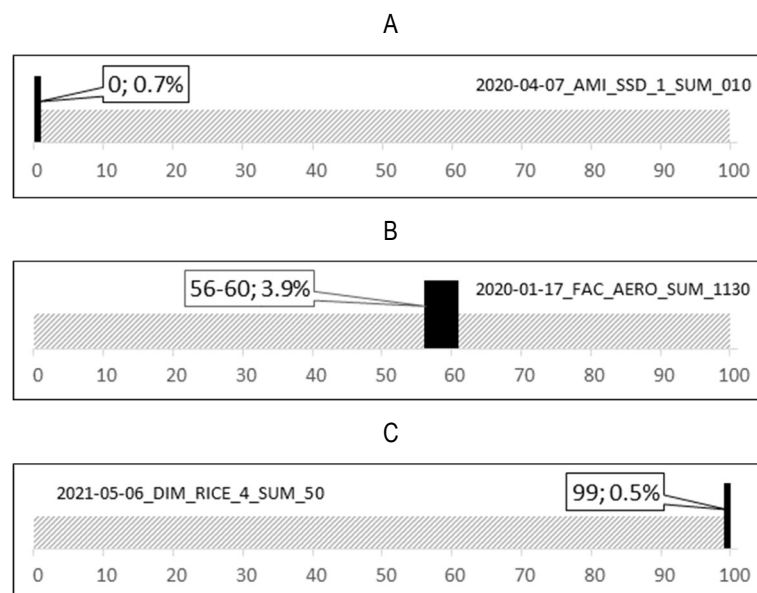
Thus, SumVIZ, the tool that I developed, is a basic visualization solution based on a standard software suite with chart design capability. It is fed manually with tabulated quantitative corpus data extracted from ELAN. SumVIZ graphs facilitate: (1) instant identification of the distribution of summaries within a particular lecture; (2) easy comparison of the distribution of summaries among various lectures; (3) clear, concise, and convenient communication on summarizing in lectures.

The basic SumVIZ graph is a horizontal line graph, or timeline, the x-axis representing either the normalized length of an episode as a percentage of the whole lecture, or its duration in seconds. SumVIZ graphs come in two main types: single-summary (one column) display (Figure 13), or multiple-summary displays (Figure 14). On these graphs, the x-axis, unit, and size of the graphs are normalized, and therefore look identical: the x-axis unit is 1% of the lecture duration, for a total of 100% in length (0% to 100% scale, similar to both ADP and ELVis). For manageability and legibility, the finest accuracy in this graph display is 1%, thus, any summary lasting between 0.1% and 1% of the lecture will appear in a column of the same width. In addition, a label pointing to the column provides: (a) the position of the summary in

the lecture, with a first number indicating its start point (tag) in hundredths of the lecture; (b) the precise summary duration, with a percentage number, with a decimal precision of 1/100th of the lecture, which is 7.2 seconds in a two-hour lecture. For example, for Lecture B in Figure 13, this summary ranges from 55/100th to 60/100th of the lecture, and its duration is 3.9% of the whole lecture. This precision is sufficient for a macro-level approach and communication purposes. Accurate data figures, like beginning or end tags, or durations, always remain available from ELAN if needed, with millisecond precision. The single-summary graphs below can be stacked to reveal the summarizing behavior of a lecturer throughout the whole lecture⁴².

Figure 13

Three examples of SumVIZ single-summary graphs



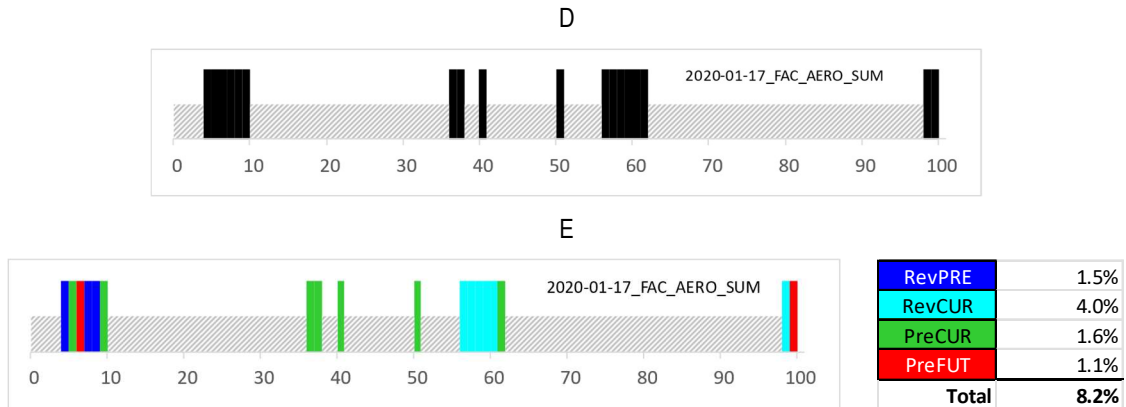
Note. Graphs from different lectures illustrating the full range of summary distributions in lectures: beginning, midsection and ending. First number in label indicates position of summary in lecture: 0 (in A) very beginning; 99 (in C) very end. Second number indicates duration of summary in % of total lecture duration. In lecture B, this summary starts at 55/100th of the lecture and its duration is 3.9% of the total. Minimum graph accuracy in the display: 1% of the lecture.

Another way of displaying multiple summaries is to compile all the summary instances of one lecture into a single graph. The result is similar to ELAN's ADPs, but with the option of displaying axis labels or other information as required, as in the two examples of Figure 14. Graph D shows the same lecture as graph E, but with the summaries differentiated by type, along with their relative duration.

⁴² Although very legible and potentially useful, this 'stacking of summaries' was finally not used in this thesis. An example of this visualization possibility is shown in Appendix C.

Figure 14

Examples of two multiple-summary display possibilities for the same lecture

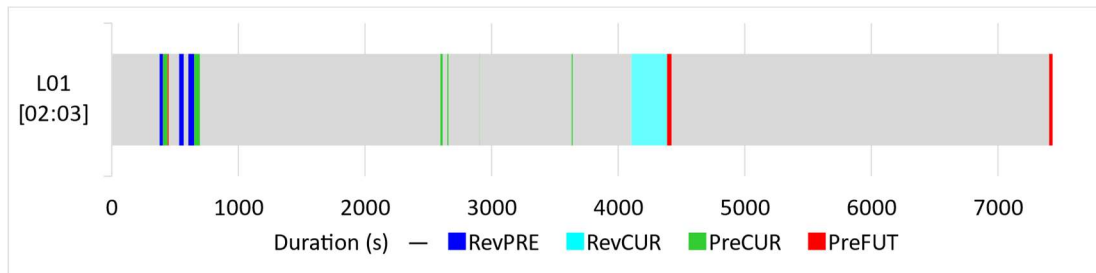


Note. This lecture contains 14 summaries. Graph D compiles all summarizing instances (undifferentiated summary types). Graph E displays the same information with differentiated summary types. The total portion of summarizing amounts to 8.2% in this aeroelasticity lecture. [L01 – 20200117_FAC_AERO]

If necessary, a higher granularity display can be used in SumVIZ. In that case, the x-axis is graduated in seconds, so that the granularity limit becomes that of the resolution of the computer system. For summaries lasting a few seconds legibility can be problematic, with hairbreadth displays. An example of such display is given in Figure 15.

Figure 15

SumVIZ: example of a fine-grain summary distribution display per type

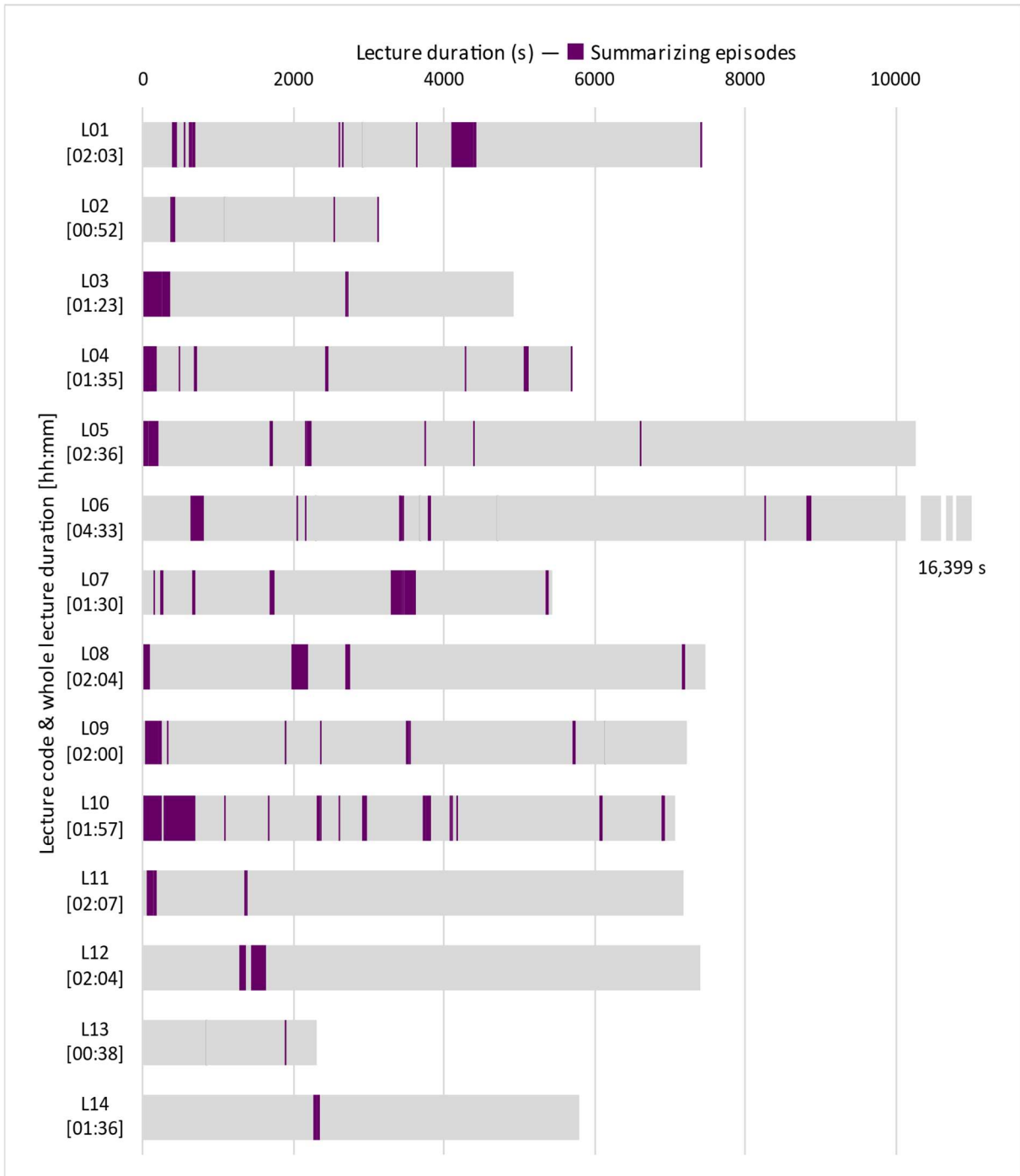


Note. Aeroelasticity lecture. Total lecture duration: 2h03 (7,380 s). [L01 – 20200117_FAC_AERO]

The duration of CoFEL lectures varies from one to 4 hours, and lecture comparisons can be carried out, related to our research questions. As with ELViz, normalized graphs can be combined to compare elements of the corpus, or the whole corpus. Figure 16 displays a fine-grain overview of all the combined summarizing instances in CoFEL, in chronological order of recording, highlighting the wide variation in summary duration and distribution, which will be presented in the following sections.

Figure 16

SumVIZ: overview of the summarizing activity in CoFEL



Note. 14 lectures: L01 = lecture #1 in chronological order of recording. Lecture duration appears in seconds (s) on the x axis and in square brackets in the [hh:mm] format on the y axis.

N.b.: L06 is partially represented here: its total duration is 16,399 seconds; its hidden part contains no summary.

To conclude, SumVIZ visualization options will be used in the following section dealing with the analysis of the various summary types, with a granularity level adjusted to

needs. Although SumVIZ has limitations and its implementation remains manual, it is effective for a corpus the size of CoFEL.

4.4 Results: quantitative summarizing data in CoFEL

This section presents a quantitative description of the summarizing activity in CoFEL. The 14 lectures in the corpus are labelled in order of recording, from L01 to L14. Three pairs of lectures are delivered by the same lecturers, identified by letters A, B and C. At the graduate school of engineering, a standard lecture lasts two hours. The two longest “lectures” (L06 and L05) are practical laboratory sessions (n.b., for practical reasons, the term “lecture” encompasses any type of class format in this thesis). The shorter lectures are non-standard formats, mainly due to COVID-19 degraded situations, with the lectures being delivered online.

4.4.1 *Summarizing activity*

Figure 16 above illustrates the summarizing activity during lectures in CoFEL, and Table 10 below shows its corresponding numerical data. The overall lecture duration ranges from 38 minutes to 4½ hours, with a mean of almost two hours (01:56), coming close to the standard institutional lecture format of two hours. All lectures contained some summarizing activity. The table shows the mean duration of summarizing activity in a two-hour lecture (6.4 minutes, or 5.6 percent). The overall summarizing activity varies greatly amongst lectures, ranging from half a minute and 1% of the lecture (L13), to nearly 19 minutes and almost 16% of the lecture (L10_C).

Table 10

Summarizing activity and summarizing episodes in CoFEL lectures

Lecture		Summarizing activity			Summarizing episodes	
ID	Duration (h:mm)	Duration (minutes)	%	n	Mean duration (seconds)	Frequency (episodes/hour)
L01	2:04	9.8	7.9	14	104	6.8
L02_A	0:52	1.9	3.6	5	23	5.8
L03_A	1:22	6.8	8.3	4	102	2.9
L04	1:35	6.6	6.9	10	40	6.3
L05	2:51	6.7	3.9	8	51	2.8
L06	4:33	6.9	2.5	13	32	2.9
L07_B	1:31	8.9	9.9	9	60	6.0
L08_B	2:05	7.0	5.6	4	41	1.9
L09_C	2:00	6.6	5.5	9	42	4.5
L10_C	1:58	18.8	15.9	18	63	9.2
L11	2:00	2.7	2.3	4	44	2.0
L12	2:03	5.0	4.1	2	150	1.0
L13	0:38	0.5	1.2	2	14	3.1
L14	1:37	1.3	1.3	2	38	1.2
Mean	1:56	6.4	5.6	7.4	57	4.0

As for the summarizing episodes, on average, a CoFEL lecture contains 7.4 summaries, lasting almost one minute (57 s) each, but again, a lot of variation can be observed. The range is rather wide, as some lectures contain only two episodes, while one lecture contained a total of 18 episodes. Moreover, in case of very few summaries, these can be either very long (L10 has the longest: 7 minutes), or extremely short (L01 has the shortest: 5 seconds). In terms of frequency (episodes per hour), on average, lecturers use summaries four times per hour. But we observe important differences between lecturers with frequencies as low as one summary per hour and as high as nine summaries per hour. Thus, some lecturers seem to use a few short summaries, while others use longer or more frequent ones.

Two pairs of lectures were delivered by the same lecturer and on the same subject: lecturer A delivered lectures L02 and L03 and lecturer C delivered lectures L09 and L10, both a week apart. In both cases, the amount of summarizing increased notably from the first to the second lecture (from 3.6 to 8.3 percent for lecturer A and from 5.5 to 15.9 percent). An interpretation could be that lecturers tend to summarize more in the second lecture, as previously-delivered knowledge has to be activated and summed up before delivering new content. Lecturer B delivered two lectures on a different subject showing the reverse summarizing pattern with a decrease in percentage (from 9.9 to 5.6 percent in the second lecture).

4.4.2 Type of summarizing activity in lectures

Table 11 shows the summarizing episodes in lectures by types in two categories of data, per lecture: the frequency of episodes and the percentage of summarizing duration.

Table 11

Type of summarizing activity in CoFEL lectures

ID	Number of episodes per lecture				Amount of summary types (% of total)			
	Review Previous	Review Current	Preview Current	Preview Future	Review Previous	Review Current	Preview Current	Preview Future
L01	3	2	6	3	18	49	19	13
L02_A	1		2	2	20		48	31
L03_A	1	1	2		63	5	32	
L04	4		4	2	53		27	20
L05		1	7			8	92	
L06	1	1	9	2	5	1	76	18
L07_B		4	3	2		48	18	34
L08_B		1	2	1		53	38	8
L09_C		3	5	1		28	63	9
L10_C	3	5	4	6	61	15	9	15
L11	1	1	1	1	62	3	18	16
L12	1	1			33	67		
L13		2				100		
L14		1	1			48	52	
Mean	1.1	1.6	3.3	1.4	23	30	35	12

Although all lectures contain summarizing activity, they do not contain all types of summarizing. Only four lectures (L01, L06, L10 and L10_C) contain all four types of summarizing. All the other lectures miss at least one type, and one lecture even misses three types (L13). Thus, some lecturers do not feel the need for, or omit, to give signals and signposts, or to review or preview their course content. We can see that reviews of previous lecture content and previews of future content were the summaries the most often left out. All but two of the lectures contain reviews and previews of the current content. Reviews and previews of current content also dominate in terms of episode frequency (a mean of 1.6 and 3.3 episodes for reviews and previews, respectively), compared to reviews of previous content and previews of future content (a mean of 1.1 and 1.4 episodes respectively).

Reviews and previews of current lecture content also surpass the other types in terms of duration (30% and 35% of all summarizing for reviews and previews, respectively). Thus, the preference of lecturers for the summarizing of current content is confirmed for duration. However, large individual differences can be observed. Four lecturers devoted more than 50%

of their summarizing activity to reviewing the previous lecture content. Two of these four lectures actually correspond to the second lecture on the same subject by the same lecturer (L03_A and L10_C). Thus, in consecutive lectures on the same subject, summarizing seems to be used to reactivate previously taught content. In L10, a large amount of summarizing is dedicated to reviewing the previous lecture (three episodes, adding up to 61% of all summarizing in the lecture). Likewise, 6 episodes are dedicated to previewing the future content of the next (unrecorded) lecture (15% of all summarizing in the lecture). Lecturer B taught two lectures on different subjects, but no particular summarizing pattern could be found in that case. This lecturer does not use reviews of previous lecture content as these lectures bear on independent single subjects and are not part of a series or module, unlike the pairs of sessions of lecturers A or C.

4.4.3 *Token length and duration of episodes as a function of summarizing type*

In this subsection, we want to know whether the token length and the duration of summarizing episodes vary according to their type. Table 12 shows the median, mean, and standard deviation of length and duration of the summarizing episodes, expressed in tokens and in seconds. The table shows high standard deviations which differed depending on summarizing type. The Kruskal-Wallis H test, a rank-based nonparametric test, was used to determine whether summarizing types differed in the number of tokens or in their duration in seconds. The Kruskal-Wallis H test showed that there were no statistically significant differences neither in number of tokens, $\chi^2(3) = 5.576$, $p = 0.134$, nor in number of seconds, $\chi^2(3) = 5.843$, $p = 0.120$, between episodes of different summary types.

Table 12

Length of the summarizing episodes expressed in tokens and in seconds

Summarizing type	Tokens				Seconds			Tokens/second	
	n	Median	Mean	SD	Median	Mean	SD	Mean	SD
Review previous	15	91	198.6	235.1	42.6	100.3	118.0	2.04	.48
Review current	23	70	107.7	120.3	36.7	60.0	76.2	2.16	.64
Preview current	46	50.5	78.3	81.2	26.3	38.3	38.4	2.19	.60
Preview future	20	61	73.0	55.8	26.5	35.8	32.3	2.21	.52
Total	104	57	101.1	125.8	30	51.6	66.8	2.17	.58

Pairwise comparisons were made between the summary types for token length, with a significance value of .05 (5%). Outcomes indicate that there is a statistically significant

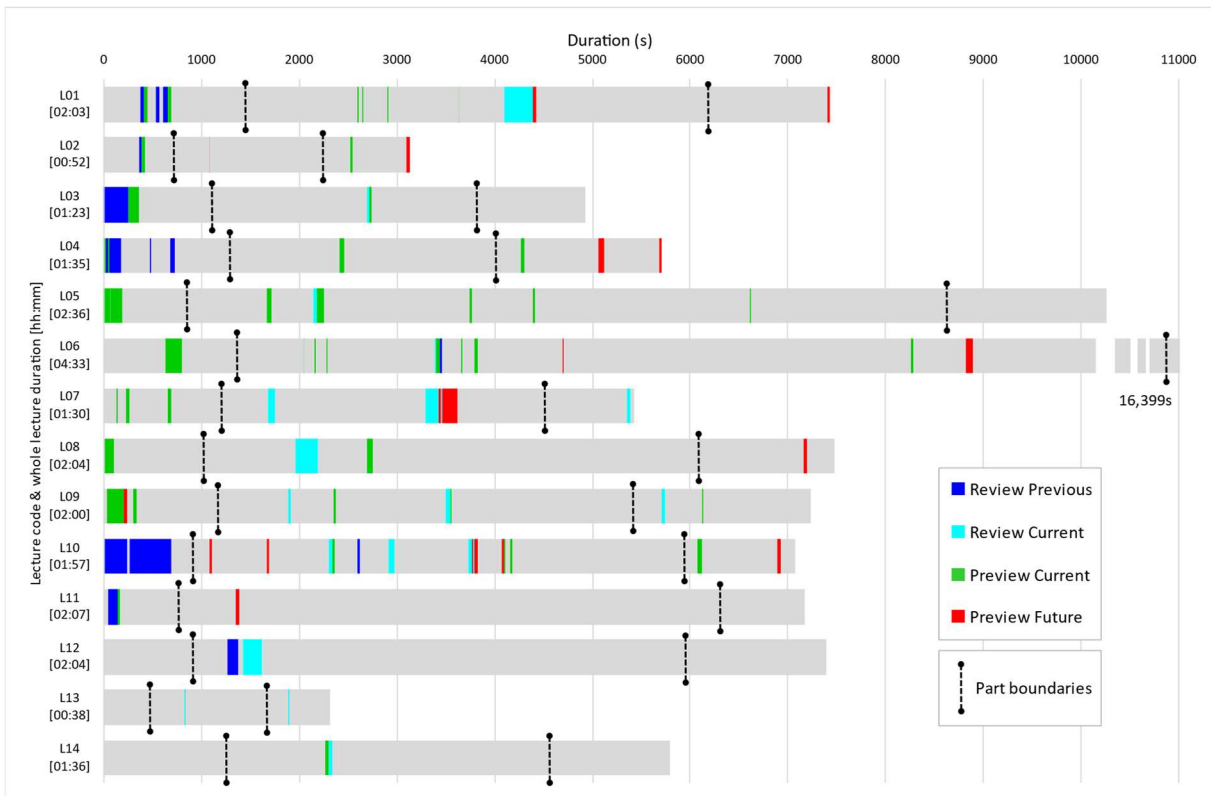
difference only between the Review previous episodes and the Preview current episodes in both number of tokens and in duration ($\chi^2(1) = 2.339, p = 0.019$). Moreover, the difference between Review previous and Preview future episodes also attained significance for duration ($\chi^2(1) = 1.965, p = .049$). Thus, although the four types of summarizing episodes did not significantly differ in length overall, Review previous episodes were somewhat longer than Preview current (both in length and duration) and the Preview future episodes (in duration only).

4.4.4 Summarizing type as a function of location in the lecture

This section examines summary episodes according to their location in the lecture. The initial premise was that there might be differences in the distribution of summaries depending on the different phases of the lecture. For example, reviewing previous content may take place at the beginning of lectures, before new subjects are introduced and developed, or previewing future content may take place at the end of lectures, to lead into the future module sessions. We considered that every lecture has a macro-structure: a beginning, a midsection, and an ending. Moreover, we considered that the first and last long continuous stretches of lecturing without summaries were the markers of the boundaries. Four working principles were observed: (1) a boundary cannot fall within a summary; (2) the longest stretch of non-summary after an initial series of summaries marks the beginning boundary; (3) the longest stretch of non-summary before a final series of summaries marks the ending boundary; (4) in case of long initial or final stretches of non-summary, they mark a boundary. Two raters worked intersubjectively towards an agreement on the boundaries for each of the lectures. Figure 17 below displays the result of this process.

Figure 17

CoFEL - Summarizing overview by type: whole lectures with the three-part boundaries



Note. Vertical black dashed lines indicate part boundaries: beginning, midsection, ending.

N.b., the ending part of L06 has been shortened; the missing part contains no summaries (total duration: 16,399 s).

The summarizing types as a function of their location in the lecture are shown in Table 13. The figures in the bottom (total) row show that the largest number of summarizing episodes by far are found in the midsection of the lectures (59 out of 104 episodes). The next one is the beginning of the lectures, with 30 episodes. The end part of the lectures contains the smallest number, with only 15 episodes.

Table 13

Summarizing type as a function of the location of summaries in lectures (number of episodes)

Summary type	Beginning	Midsection	Ending	Total
Review previous	12	3	0	15
Review current	0	18	5	23
Preview current	16	26	4	46
Preview future	2	12	6	20
Total	30	59	15	104

Pearson's chi-squared test (χ^2) was used to examine whether the two categorical variables, summary type and location in the lecture, are independent. The relation between these variables was significant, $\chi^2 (6, N = 104) = 36.2, p < .001$, thus confirming that the type and location of summaries are dependent to a large extent. The beginning of the lectures contains relatively more reviews of previous content. Indeed, eighty per cent (12 out of 15) of the summarizing activity in the beginning is reviewing content which has been dealt with in past lectures. This happens to the expense of reviewing current content and previewing current and future content. Concerning the midsection of lectures, Table 13 shows that fifty-seven per cent (59 out of 104) of all summarizing activity takes place here. The reviewing of current content is clearly favored at seventy-eight per cent (18 out of 23). Finally, Table 13 shows that summarizing mainly consists of reviewing current and previewing future content towards the end of lectures.

4.5 Pragmatic and linguistic variation in CoFEL summaries

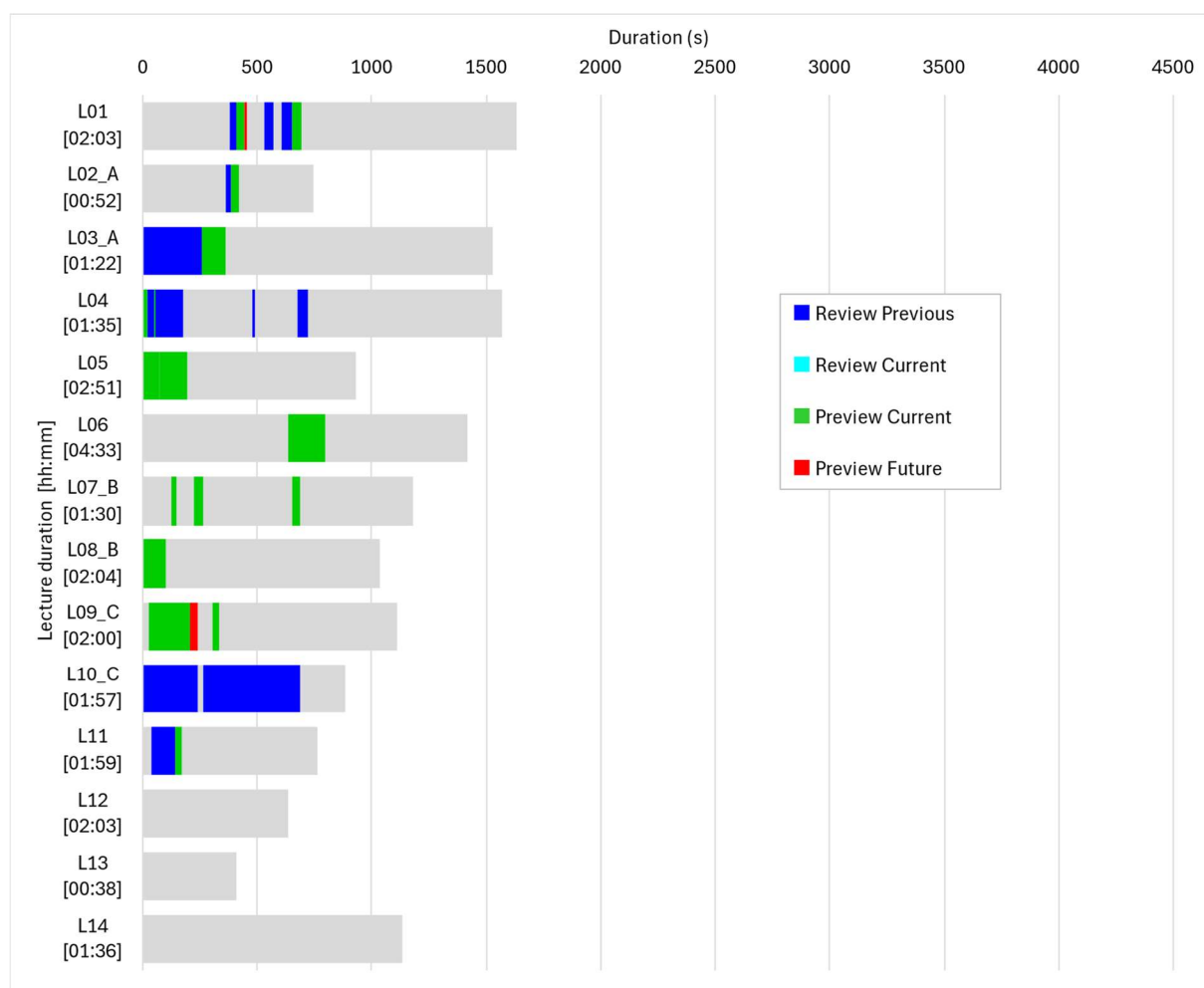
The following four subsections examine linguistic and pragmatic variation in summaries, starting with the beginning part of lectures, with RevPREs (subsection 4.5.1), and then PreCURs (4.5.2). Variation in lecture midsections and endings is covered in subsections 4.5.3 and 4.5.4. A final subsection will explore chaining of summaries (4.5.5).

4.5.1 *A trait of lecture beginnings: RevPRE*

This subsection studies the review of previous content (RevPRE) summary type as a trait of the beginning part of CoFEL lectures. The purpose of RevPREs is to restate lecture content that was given earlier in the course or module, thus functioning anaphorically. In engineering lectures, this type seems to play two main roles: (1) to ensure that all key concepts and fundamental knowledge are in place as the basis for future learning (Alsop, 2015); (2) to

mark the relevance and underline the importance of certain vital points in the module or course (Deroey & Taverniers, 2012). The pragmatic and linguistic variability of RevPRE summaries will be illustrated using three examples taken from the beginning part of lectures: one RevPRE in the initial position of the lecture (analyzed in depth), and then the shortest and longest RevPREs (more briefly commented). Figure 18 below shows an enlargement of the beginning part of lectures.

Figure 18
Summarizing in CoFEL: beginning part of the 14 lectures

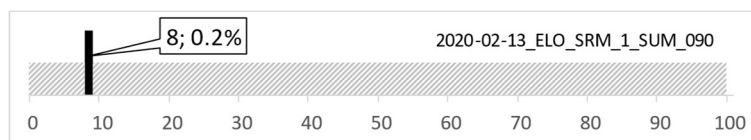


In CoFEL, restating lecture content can be done briefly, in a few words and seconds, or lecturers can summarize at length, accounting for a large part of the lecture. The shortest RevPRE summary lasts twelve seconds, while the longest, at seven minutes long, is the longest of all summaries in the corpus. Three summaries will now be examined for linguistic and pragmatic features, from shortest to longest.

The shortest of all RevPREs comes from a Soil & Rock Mechanics on-site lecture (L04). This summary takes place shortly after the beginning of the lecture and represents 0.2% (34 tokens, 12 seconds) of the whole lecture duration (Figure 19).

Figure 19

Position and duration of the shortest RevPRE summary in CoFEL



Note. Position: 8/100 of the lecture. Duration: 0.2% of the lecture. Soil & Rock Mechanics lecture.

[L04 – 20200213_ELO_SRM1_SUM_0090]

Using the Simple Past tense (*it was not; you analyzed*), the lecturer anaphorically recalls an “exercise” done in a previous lecture that involved one or several “curves” (Transcript 1). The episode is introduced by the most common DM in CoFEL: *so*. The past temporal backward deictic reference *last time* does not appear in this episode, but it does a little earlier and this passage is still under its influence. The singular demonstrative *this* and pronoun *it* do not match the plural *curves*, and are instances of ELF divergence.

er: so- this curves- *it was not* exactly the same but *you analyzed* this on *the exercise*.
remember? to have er different er couples of value er sigma three and sigma one-
the maximum.

Transcript 1. Example of a short RevPRE summary (the shortest in CoFEL: 12 s, 35 tokens). Typical markers are *italicized*.

[L04 – 20200213_ELO_SRM1_SUM_0090]

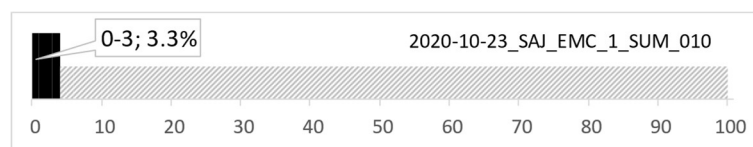
The lecturer uses the definite nominal group *the exercise* anaphorically, in an act of retrospective labelling and lexical cohesion (Francis, 1994) to connect to previously delivered content and organize her discourse. She signals to her audience that she is moving on to the next stage of her argument by packaging the information in a single nominalization (Francis, 1994), and provides a link between the backward and forward contexts of the lecture. The distal deictic mental verb *remember* appears in the elliptic interrogative form “(Do you) remember?”. This implicit pronoun *you*, together with *you analyzed*, directly addresses the students, in a move to draw their attention and involve them. Simultaneously, the lecturer points a finger at the image displayed on the slideshow, in a deictic gesture. Thus, in this summary the lecturer’s pragmatic intention seems to be threefold: first, to link two parts of the lecture; second, to help

students to recall information from previous lectures; finally, to build rapport with students, via person deixis with the personal pronoun *you*.

The next summary is one of the two RevPREs found in initial position in lecture L10, on Electro-Magnetic Compatibility (Figure 20, full transcript in Appendix G). When in initial position, the summary's primary function is to introduce the current lecture with an anaphoric recapitulation of the previous lecture. It is rather long, lasting almost four minutes (460 tokens, 3.9 minutes). As the lecturer is delivering his summary, he quickly goes through the slideshow from the previous session, making reminder comments along the way, accompanied by mouse or laser pointing gestures. These gestural deictic movements are also verbalized by brief descriptions of the slides accompanied by four occurrences of the deictic adverb *here* (ll. 6 to 8 & 27, Appendix G).

Figure 20

Position and duration of a RevPRE summary in initial position in CoFEL



Note. Position: 0/100 of the lecture. Duration: 3.3% of the lecture. Electro-Magnetic Compatibility lecture.

[L10 – 20201023_SAJ_EMC1_SUM_0010]

The summary opens with the two most common DMs in CoFEL: *ok* and *so*. They are used consecutively, with a short pause in between: the lecturer signals the start of the lecture, in an inchoative move, and calls for the participants' attention. He is also giving himself and his students a few seconds' concentration time before starting the lecture proper. The DM *so* appears a total of ten times in the transcript, including once more in combination with *ok* (l. 29) and signals the end of summary, as it opens the coda (ll. 29-30: *that was the first point where ... i will not detail more than this*). This statement is also framed by two four-second pauses, during which the lecturer gives himself time to reflect, probably to make sure that he has not forgotten to say anything essential. The common adverb *basically* (l. 1), can also be considered a DM and a time-filler.

A concluding statement comes just after the second pause and a long *er* of hesitation, as an uncertain afterthought (*i think we also started to define the standards*). Backward temporal deictic markers, like *last week* (l. 1) co-occur with convergent markers, like verbs in the Simple Past tense: the inclusive inchoative *we started* (l. 1) and the inclusive reminder *we reviewed*

(l. 3). The mental verbal expression (*as you may*) *remember* (l. 11) is also typical of these reviews. The recurring use of direct-address, personal pronoun *you* (10 occurrences, with possessive *your* included, l. 14) may reveal the interpersonal consideration of the lecturer for his students.

Retrospective labelling appears in the first line of the summary, with the shell noun (*we started with some*) *introduction* (l. 1). It encapsulates the entire previous lecture and re-introduces the main topic seen then: how electrical equipment can be disturbed by other devices (ll. 1-2). Some relevance marking follows, with the highlighting of *the definition of the electromagnetic compatibility* mentioned as being *the core* of the subject (ll. 2-3). Thus, the first three lines serve as an act of retrospective encapsulation (Francis, 1994). Then, the majority of the text (ll. 2-29) is devoted to recapitulating the previous session in detail, before the coda. Relatively little ELF divergence occurs in the intended message of the lecturer, who achieved the highest possible score on the language test⁴³. Nonetheless, the derivational suffix in the third person singular verbs is missing, for example, *divide* and *split* (l. 12), or *induce* (l. 28), which the lecturer self-corrects (l. 29). However, divergence is not systematic, as in the standard English (SE) form *generates* (l. 28). A few other mainly conjugational divergences also appear: *shielding must either be avoiding* (ll. 18-19; SE: *must also be avoided*), or *to allow the current flowing* (ll. 24-25; SE: *to let the current flow*). In terms of lexical-semantics, there is the repeated use of the term *diaphony* (four occurrences, ll. 9 & 16), which appears to be a key term in the lecture. *Diaphony*, although plausibly an English word like *telephony*, seems to be the lecturer's coinage from French *la diaphonie*; the English word cannot be found in dictionaries or specialized Internet searches. The standard translation of the French term is *crossstalk*, referring to an undesired effect between two close circuits in signal transmission. The risk is that students will not understand the term and its context, or worse, that they will retain an incorrect term and reuse it later, including in examination papers or in their future professional communication.

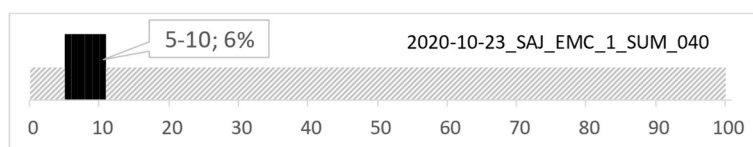
As the longest RevPRE episode in the corpus happens to be from the same lecturer as above, its in-depth analysis is not included here (Figure 21, full transcript in Appendix H). It lasts over seven minutes (873 tokens, 7 minutes, 6% of the lecture). The longest summary of all in CoFEL is quite naturally of this type, as RevPREs are by far the longest, on average. It comes from an on-line, two-hour Electro-Magnetic Compatibility lecture (L10). The topic is

⁴³ Level C1+/C2, on the Council of Europe's CEFRL scale (Council of Europe, 2024).

the compliance of electronic devices with European standards. In terms of token length, it represents thirty-seven per cent (over a third) of the whole lecture (2,346 tokens). This percentage is explained by the fact that a large part of the lecture time is devoted to application exercises of the reviewed preceding content, during which the lecturer is silent or does not interact orally with his students until the correction. As he speaks, the lecturer quickly scrolls through the slideshow used in the previous session and summarizes its content. This summary takes place shortly after the beginning of the lecture, and the lecturer takes great care to recapitulate at length and in detail the content of the previous lecture anaphorically.

Figure 21

Position and duration of the longest summary in CoFEL (RevPRE type)



Note. Position: 5/100 of the lecture. Duration: 6% of the lecture. Review of previous content (RevPRE) type. Electro-Magnetic Compatibility lecture. [L10 – 20201023_SAJ_EMC1_SUM_0040]

These three examples of the fifteen RevPREs are of contrasting position and duration and illustrate the considerable variation that can exist for these parameters in CoFEL. By examining some of the salient features of these summaries, we have tried to highlight their main pragmatic function, namely the reiteration of lecture content given earlier in the module, to reveal how this was done linguistically.

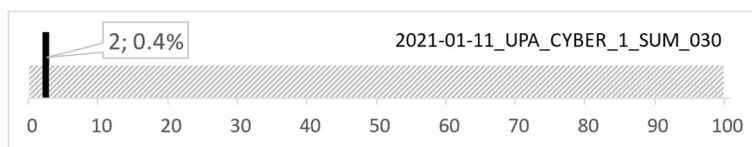
4.5.2 *Another trait of lecture beginnings: PreCUR*

The preview of current content (PreCUR) summary type can be found in initial and delayed positions, as found in lectures L09 (initial) and L11 (delayed). The explicit metadiscursive signaling function and strong discourse-structuring function of these summaries have been highlighted in previous research (Alsop, 2015; Crawford Camiciottoli, 2004; Young, 1994): they present the outline and announce the topic of the upcoming class.

A short PreCUR is found in L11 (Figure 22, Transcript 2) and comes from a cybersecurity lecture. It lasts only 30 seconds (63 tokens), and comes shortly after the beginning of the lecture, but not in initial position.

Figure 22

Position and duration of a short PreCUR summary in CoFEL



Note. Position: 2/100 of the lecture. Duration: 0.4% of the lecture. Preview of current content (PreCUR) type. Cybersecurity lecture. [L11 – 20210111_UPA_CYBER1_SUM_0030]

The lecturer only begins to deliver his topic content after some regulating and lecture outlining. The summary opens with the typical PreCUR temporal reference *now*, which is also a micro-DM here, signaling the beginning of the topic delivery to the audience. It is followed by a rhetorical question (*What means to secure something?*) of the micro-type (Khan & Aguilar-Pérez, 2023), supporting the lecturer’s navigation within the topic. As a rhetorical question does not work through interaction, the lecturer plans to answer it himself (Peters, 2004; Tadros, 1994), which he does with the rest of the episode. The three micro-DMs (boxed, in Transcript 2), the inclusive pronoun *we*, and the question form are ways of securing the audience’s attention and engaging them. PreCURs do not rely much on straightforward temporal references as the other types, however *now* and the question are both PreCUR type markers here, anchoring the summary in immediacy:

now- *what means er to secure something?* *well*- to secure something means to respect some preliminary predefined er security policy. *so*- *we'll* see in the second part of the- of today class what er security policy is- who has to define the security policy and what's the place of the security policy in the: overall er: (2) picture of er: industrial systems cyber security.

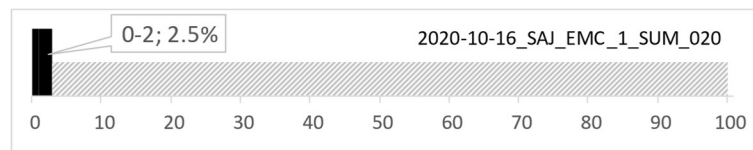
Transcript 2. Example of a short PreCUR summary. Type markers are *italicized*. DMs are **boxed**. The personal pronoun is **highlighted**. [L11 – 20210111_UPA_CYBER1_SUM_0030]

The episode contains some ELF divergence, such as the absence of: an auxiliary in the question on line 1 (SE: *What does it mean to secure something?*); a genitive on line 3 (SE: *the second part of today’s class*), which could be either a grammatical issue, or a pronunciation issue, with the speaker simplifying the inter-lexical consonant cluster /zkl/.

The longest summary in the PreCUR category in CoFEL and comes from Electro-Magnetic Compatibility lecture L09 (full transcript in Appendix I). It lasts almost three minutes (178 seconds, 380 tokens) and opens the lecture (Figure 23). Occupying 2.5% of the lecture, this episode reveals the care that the lecturer takes in scaffolding the content to come.

Figure 23

Position and duration of the longest PreCUR summary in CoFEL



Note. Position: 0/100 of the lecture. Duration: 2.5% of lecture. Electro-Magnetic Compatibility lecture.

[L09 – 20201016_SAJ_EMC1_SUM_0020]

The summary is rich in elements of interest, albeit with some repetitions, and has a complex structure. In order to facilitate comments on the contents of this summary through its structure, we have inserted reference numbers in square brackets for the nine parts identified in Appendix I. In part [1], the lecturer announces the topic of the day: Electro-Magnetic Compatibility, a part of the Power Electronics module. The overall course organization is revealed in part [7]: four lectures and two laboratory sessions. The summary, and thus the lecture, starts with the classic discourse marker (DM) *so*, discussed earlier (see 4.5.2); this DM occurs five times with the same role of pause filler throughout the summary (boxed in the text). The occurrence of *so* on line 2 has a different role, signaling causality. Part [1] at once begins with a PreCUR marker (italicized) with the temporal deictic *today*, as part of the string of text announcing the day’s topic: “the lecture of today is focused on” (l. 1). The type marker also appears in lines 10 and 22. Another such PreCUR marker can be found farther down, as part [6], with the use of the modal *will* and the Future tense (“that will be the main points that will be addressed”). Part [6] functions as a concluding statement to parts [1] to [5], the introduction to the lecture (parts [7] to [9] will be dealt with further below).

Part [2] creates a break in the summary, as the lecturer switches from the Present tense in part [1], to the Past tense, with the repeated *you started*. Strictly speaking, part [2] can be seen as a brief example of embedding, in which one type of summary (here, RevPRE) lies within one of another type (PreCUR). In part [2], the lecturer briefly recapitulates what ground was covered in previous lectures, delivered by a colleague, on complementary topics. Two even shorter cases of fleeting embedding, similar in principle, can be found in parts [4] and [7]. Both are of the PreFUT type, a type examined in greater detail in the following subsection: “we will go on next week” (l. 11, part [4]), and “we have four lectures and also two different labworks” (all of part [7]).

The articulation with part [3], in which the lecturer returns to the topic of the day, is clearly marked with the PreCUR temporal deictic marker *now*, and the return to the Present

tense *we are working* (l. 5). This part contains one of the instances of relevance marking in this summary, in bold type (l. 6), as Electro-Magnetic Compatibility is presented as “something which is quite important” in Power Electronics. Research has shown that marking relevance can help students process the lecture and form a coherent mental picture of its content (Deroey, 2013; Deroey & Taverniers, 2012). The lecturer seems aware of that role, as he uses relevance marking in three other instances (also in bold type) in part [5], plus one in part [6] (*the main points*). Part [5] announces what is *the main goal* of the lecture, that knowledge of the Electro-Magnetic Compatibility standards is *mandatory*, and what students and professionals *need to know*. This last expression is introduced by the inclusive pronoun *we*. There are eight occurrences of this pronoun, and the same number of 2nd person pronoun *you*.

As seen earlier in this section, the personal pronoun *we* plays a key role in establishing lecturer-student rapport when used in its inclusive form, which is the case for all of these eight occurrences (see 4.5.1 for a discussion). The pronoun *you* is a direct address to the students, in a move to attract their attention and involve them. In this summary, the lecturer’s pragmatic intention seems to be to help students recall information from previous lectures (exposed in part [2]), in an interpersonal move towards his students, via person deixis, with the frequent use of personal pronouns *we* and *you*. But this lecturer goes further in this direction, showing strong empathy with his students for the workload that is required of them, given the complexity of the subject and the long hours of a full day’s study. This is done in two passages in parts [4] and [5]: “because you are at the end of the day and you may be tired” (ll. 10-11), and “the main goal of the lecture is to help you in understanding how to design an e-m-c filter” (ll. 13-14). The lecturer empathizes with his students about their potential fatigue at the end of a long day of content-rich online lectures⁴⁴. Intentionality is also present in part [3], with the modal *will* in the phrase “I will try to help you in understanding” (l. 8). In this way, the lecturer insists on how he wants to help his students understand the complexities of the day's subject.

To finish, this summary constitutes an enumeration of course points (Dubois & Sankoff, 2005; Tadros, 1989). Starting with *now* in part [3], the enumeration progresses with *after the basis* (part [4]), *and finally* repeated twice (part [5]), to finish with *and then* (part [9]). After concluding his summary with part [6], the lecturer adds a complex coda, made of the PreFUT embedding (part [7]), the repeated comment on the hypothetical completion of the day’s lecture (part [8]), and a new and more detailed conclusion to the summary (part [9]). In

⁴⁴ 2020 COVID period lecture, delivered online.

this final part, the phrase *how to* is repeated four times, as the lecturer commits himself to meeting the practical needs of his students in their knowledge of Electro-Magnetic Compatibility. Finally, the episode contains some repetitions (notably parts [4] and [8]) and hesitations, and only one divergent ELF form in the mismatch between a plural noun and its singular determiner, in line 25, with *this disturbances*. This could be either a pronunciation or a grammatical issue (the lecturer achieved the highest possible score on the language test⁴⁵).

To conclude on these two PreCUR summaries (Transcript 2, a short and delayed summary; Appendix I, an initial and long summary), they were both from online lectures, and they differed greatly in length. They had in common some basic functions that are characteristic of the type, although not identical ones: the first outlines and structures the lecture; the second begins the delivery of the topic content of the day. Thus, the fact that the first one was initial to the lecture and the other delayed by a few minutes did make a functional difference. Both summaries also contain hesitations and ELF divergences, with passages that could disturb comprehension in student audiences, especially the second summary. With two summaries illustrating the great variability in duration of the PreCUR type, we have exemplified the most common of all four summary types in CoFEL (46 occurrences out of 104). We examined the linguistic realization of the main pragmatic feature of the type, a reference to information to come in the current lecture.

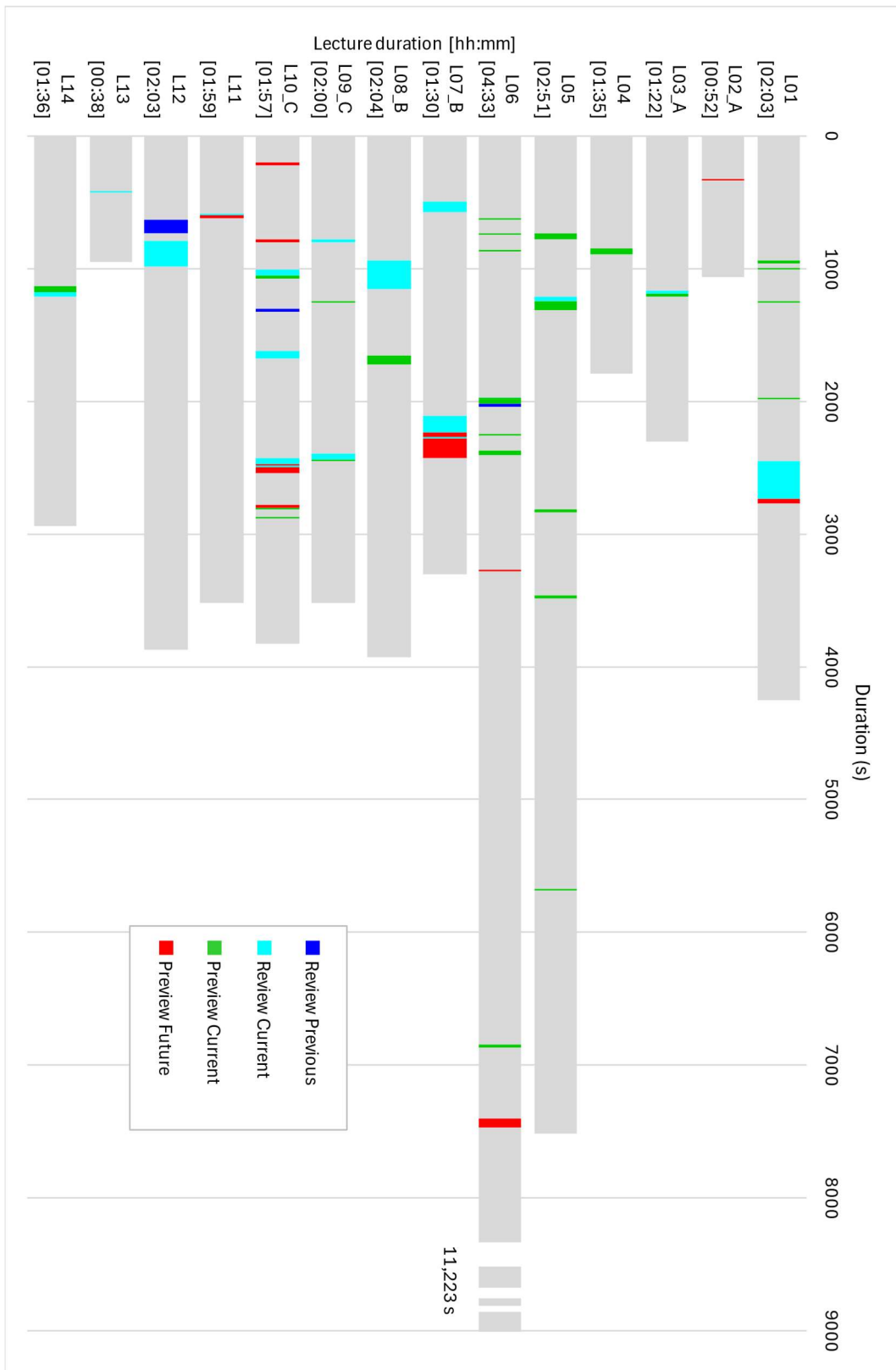
4.5.3 *A trait of lecture midsections: RevCUR*

Reviews of current lecture content (RevCURs) are completely absent from the beginning part of lectures but do occur in the midsection of CoFEL's 14 lectures. Figure 24 shows that, although all four types of summary are present in this part of the lectures, they are not evenly distributed. Together with PreCURs (26 occurrences), RevCURs (18 occurrences) represent the dominant type in this part by far (only 3 RevPREVs & 12 PreFUTs). The graph also shows that RevCUR episodes, on the whole, appear to be longer than the other types, and notably longer than PreCURs. Yet, four lectures (L02, L04, L06 & L11) contain no RevCURs.

⁴⁵ Level C1+/C2, on the Council of Europe's CEFRL scale (Council of Europe, 2024).

Figure 24

Summarizing in CoFEL: midsection of the 14 lectures

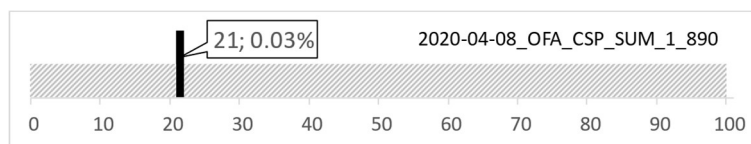


Note. For L06 (11,223 s), this part without summary has been slightly shortened to improve overall visualization.

The first example of the RevCUR type concerns the shortest summary (Figure 25 & Transcript 3), which comes from a civil engineering laboratory lecture (L06). Concluding the first quarter of the lecture, it links the end of the first topic to the beginning of the next.

Figure 25

Position and duration of the shortest summary in CoFEL (RevCUR type)



Note. Position: 21/100 of the lecture. Duration: 0.03% of lecture. Characterization of Soil Properties lecture.

[L06 – 20200408_OFA_CSP1_SUM_0890]

Structurally, this summary consists of only two elements: a reversed pseudo-cleft packaging the information (Biber et al., 1999; Deroey, 2012), and the topic title of what has just been dealt with, *primary consolidation*, which refers to the consolidation of a sand sample in the laboratory, in the process of characterizing sand properties. The first element is introduced by the most common discourse marker (DM) in CoFEL: *so*. Concerning the reversed pseudo-cleft *this is for what concern*, the divergent ELF verbal form *concern* lacks the “s” ending of the SE conjugation. Omitting this “s” is a salient feature of EMI-ELF speech in general (Seidlhofer, 2004), and throughout CoFEL in particular. This reversed pseudo-cleft is of a proximal backward deixis nature. The whole RevCUR summary acts as a conclusion encapsulating the topic that has just been dealt with, in a single act of retrospective labelling (Francis, 1994).

so- this is for what concern the primary consolidation.

Transcript 3. Example of a short RevCUR summary. The reversed pseudo-cleft is *italicized*.

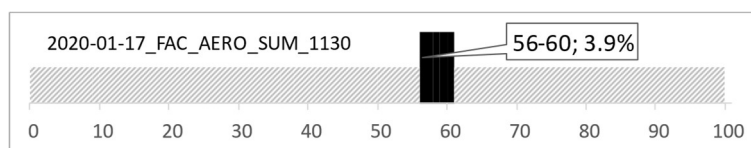
[L06 – 20200408_OFA_CSP1_SUM_0890]

The next example of a RevCUR in a lecture midsection, lasts 4.75 minutes (503 tokens; 3.9% of the lecture), and is the longest in the type (Figure 26, full transcript in Appendix J). It is part of a math-heavy Aeroelasticity lecture (L01). In this summary, the lecturer recapitulates all the points that have just been seen in the course anaphorically, while referring in turn to the equations that have been written on the board. The subject of the lecture is the ‘flutter phenomenon’, which can lead to the destruction of an aircraft when vibrations are uncontained. The phenomenon had first been studied statically (cf. l. 18) and is now being studied dynamically (cf. l. 28). The references to the flutter equations on the chalkboard are apparent

in the text through the repeated use of the adverb *here*, which appears 7 times (highlighted in light gray). The lecturer uses the board as a visual aid as he writes on it: he explains what he is writing, and then he refers back to it and re-explains things. This might be an attempt to help students organize, understand, and memorize the heavy mathematical course content.

Figure 26

Position and duration of the 2nd longest summary in CoFEL (RevCUR type)



Note. Position: 56/100 of the lecture. Duration: 3.9% of lecture. Review of current content (RevCUR) type. Aeroelasticity lecture. [L01 – 20200117_FAC_AERO_SUM_1130]

In lines 1 and 35, both DMs (boxed) in this affirmative *all right* form are utterance-initial and play the same role of pause filler, giving the lecturer some time to gather his thoughts before speaking, and giving the students some processing time by getting their attention before delivering the content; this potentially improves comprehension (Flowerdew & Tauroza, 1995). Moreover, DMs can be combined, as in line 1 where *all right* is used with two other pause fillers: *so* and *actually*. More specifically, the first *all right* marks the boundary of the transition (Richards & Schmidt, 2010) in the pragmatic function, between plain lecturing and the long recapitulation summary that is just beginning. The second one marks the end of the 35-second silence during which the lecturer writes a complement to the formulas already on the board. Finally, a new call to attention is made before concluding the summary with a final utterance. Variants of *all right* include the two occurrences of the DM in lines 13 and 17, in the interrogative form, as “All right?”, or simply “Right?”. In both cases, the lecturer uses the DMs to check student comprehension, in a similar fashion to the DM “ok?” in line 20. These three DMs, used at regular intervals in the summary, show how the lecturer wants to check that his complex content is understood. The DM *so* is a pause filler here and appears seven times in this summary. Other elements of interest include RevCUR type markers (italicized in Appendix J), such as: *we have seen that* (l. 4), *then we have considered* (ll. 7-8), or *what we have said here* (l. 14).

To conclude on this summary, the passage comprising lines 20 to 27 (from “and if we look...”, to “...the flutter phenomenon.”) lasts 39 seconds and is almost 100 tokens long, and is representative of a phenomenon that can occur in spontaneous situations: (almost) endless

utterances. As they deliver their absorbing content, sometimes lecturers go on and on. They can do this without realizing it, forgetting that they, like their audiences, need breathing spaces, pauses, and links in their discourse, to rest and digest the heavy content. Too many words in one utterance can thwart the transmission of content and the pedagogical quality of the lecture, and lead to pragmatic failure (Blum-Kulka & Olshtain, 1986).

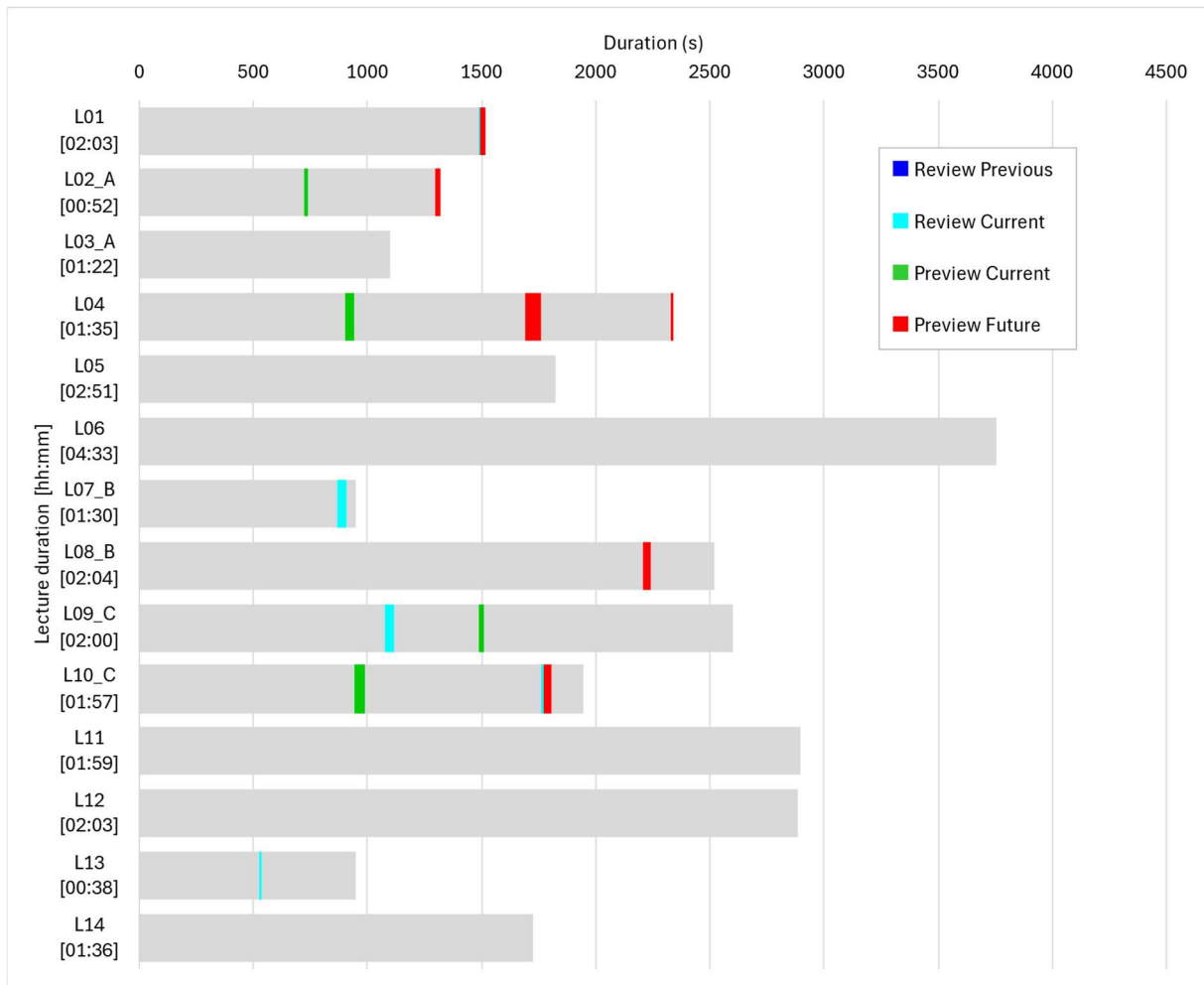
The two RevCUR examples of this subsection, out of a total of twenty-three in CoFEL, illustrate the even greater variation that can exist in duration than in the first RevPRE type (subsection 4.5.1). With this pragmatic sub-function of summarizing, lecturers make a retrospective reference to information already given in the current lecture. Depending on the lecturer and the situation, this can be done in just a few words, or in very long stretches of text. Here again, we have examined the salient features of the RevCUR summaries in detail and have tried to highlight their main pragmatic function, and to reveal how this was done linguistically. Their discursive and linguistic realization is consistent with their location in the midsection of the lecture

4.5.4 A trait of lecture endings: PreFUT

Variation is also found in lecture endings. Figure 27 shows an enlargement of this ending part, and the summaries it contains. Some of the features are: a total absence of summaries at all in some lectures (L03, L06, L11, L12 & L14); few occurrences of summarizing compared to the beginning and midsection parts; absence of RevPRE type; the six PreFUTs are the final summaries of their lectures, and three of them terminate the lectures (L01, L02 & L04). This last observation led us to compare two PreFUT summary samples from this ending part, one terminal, and one not terminal. While its pragmatic function remains the same, the PreFUT type is thus not confined to the ending of lectures, since it is found in all three parts, in varying numbers: very few in the beginning part (Figure 18); highest number of occurrences in the midsection (Figure 24); few, but dominant in occurrences, in the ending part (Figure 27).

Figure 27

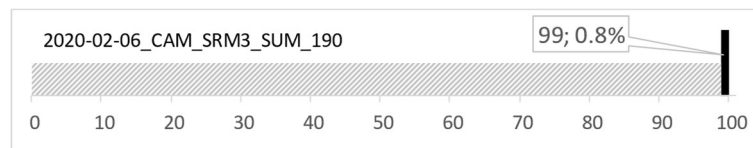
Summarizing in CoFEL: ending part of the 14 lectures



In our selection of two PreFUT examples, we have favored episodes from lectures that had not been selected earlier in this section (L02 & L04). Both chosen PreFUTs come from the same Soil & Rock Mechanics course, a branch of the Civil Engineering specialization, but they are delivered by different lecturers, in different modules, and on different narrow subjects. The first PreFUT comes from lecture L02, is terminal, and its subject is the classification of soils (Figure 28).

Figure 28

Position and duration of a PreFUT summary: ending of SRM lecture L02



Note. Position: 99/100 of lecture. Duration: 0.8% of lecture.

[L02 – 20200206_CAM_SRM3_SUM_0190]

According to research (Alsop, 2015; Nesi et al., 2014), PreFUTs function as a reference to information to come in the next lessons of the module or course. That is exactly what this short, end-of-lecture summary does (72 tokens, 26 seconds): the lecturer announces that the students will carry out fieldwork in April (the lecture is dated February), when spring arrives, and the snow has melted (Transcript 4). The future lecture content, measuring groundwater levels⁴⁶ (l. 2), is revealed, as one of the things soil engineering students and professionals do *classically* (three occurrences, ll. 1-3). This summarizing episode contains several type markers, temporal and conjugational (italicized). For the temporal, we find the distal (here, 2 months) forward deictic *in April* (l. 2), repeated twice and already mentioned, announcing a future session. This temporal marker is relayed by the series of verbs in the Future tense that dominate throughout the summary: *we will play* (ll. 1 & 4), *we will measure* (l. 1-2), *it will be* and *it will depend* (ll. 2-3).

classically *when we will play* er outside to make site invest- investigation *we will measure* the groundwater level. *so- it will be in april.* classically *in april it will depend* on the snow on these days- but classically the snow melt. *so-* the river is high. *ok?* the level of the river is (h)igh and the ground water level when *we will play-* one meter- two meters- it depend(s) (on it). *ok?*

Transcript 4. Example of a terminal PreFUT summary in CoFEL (SRM). PreFUT type markers are *italicized*. Discourse markers are *boxed*. Personal pronouns are *highlighted*.

[20200206_CAM_SRM3_SUM_0190]

The classic discourse markers *so* (pause-filler) and *ok?* (interrogative, comprehension check) are used at regular intervals (two separate occurrences of each). The inclusive personal pronoun *we* appears three times, as the lecturer engages his students in the future field activities. This engagement is relayed by the repeated use of the colloquial phrase *we will play* (ll. 1 & 4). As seen earlier in this section, the idea of *play* and *playing*, as in a game, seems to be commonly used by engineering lecturers (38 occurrences in CoFEL). Finally, the summary contains some

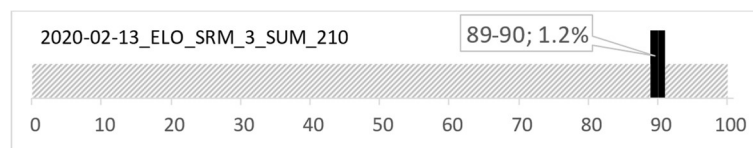
⁴⁶ As the measurement site is located in a mountainous area prone to snowfall, and close to a river, groundwater levels depend on seasonal precipitation and river levels.

ELF divergence. For example, in SE: the Present tense would follow the conjunction *when* in “when we play” (l. 1 & 4-5: “when we *will* play”); “the snow on these days” would be “snow conditions at the time”, and “the snow melt” as “the snow melts” (both on l. 3); the ellipsis of the verb would be filled in “the ground water level [...], one meter, two meters” (l. 4-5: “the ground water level [...], *could be* one meter, two meters *high*”). Despite this, the PreFUT summary plays its role: bringing the lecture to a close by providing information on upcoming sessions, and bridging the two-month gap between the current indoor lecture and future outdoor field sessions.

The second PreFUT summary examined comes from lecture L04, on rock mechanics and tunnel design. As shown on Figure 27, this lecture holds two PreFUTs in its ending part: we chose the longer, non-terminal one, to broaden the scope of the study (Figure 29; 81 tokens, 47 seconds).

Figure 29

Position and duration of a PreFUT summary: ending of lecture L04, Soil & Rock Mechanics



Note. Position: 89/100 of lecture. Duration: 1.2% of lecture.

[L04 – 20200213_ELO_SRM3_SUM_0210]

This PreFUT summary (Transcript 5) is quite similar to the previous one, in that it announces the nature of the three future lectures on the module, which include an “in situ” field session (l. 2). Here, the future topic is the measurement, not of groundwater levels, but of the resistivity and seismic velocity of soils⁴⁷.

⁴⁷ In geomechanics, soil resistivity measurement data is used to make sub-surface geophysical surveys (ore locations, depth to bedrock, etc.). Seismic velocity measurements using wave propagation help to understand the internal structure, material composition, and physical state of soils.

er remember that **you** have three laboratory- laboratory session(s) for this er er er this er course unit -two in laboratory and one in situ. and the one in situ **you** will- er i think there is one resistivity measurement of the soils from the surface and this electrical er method is (x) the general geophysics method. for the geophysic method **you** can use the electrical method- electromagnetical method- er er: seismic velocity. **so**- this is one of the principal one.

Transcript 5. Example of an ending (non-terminal) PreFUT summary in CoFEL. PreFUT type markers are italicized. Discourse marker is boxed. Personal pronouns are highlighted.

[20200213_ELO_SRM3_SUM_0210]

The organization of the three future lectures is declared from the start (ll. 1-2: “remember that you have two laboratory sessions and one in situ for this course unit”), which is a PreFUT type marker in itself. The presence of the mental verb *remember* is rare in a PreFUT summary. However, with *remember* as the first word, the role of this first utterance is to create a distal deictic connection between the previously delivered course information and the future program that awaits the students, in an immediate past-future link. The content of the *in situ* (field) session is announced on the next lines (ll. 3-5), with short explanations on the two soil measurement methods. The Future tense appears once only, in the false-start verbal phrase “*you will-*”, but the introductory statement (l. 1) unambiguously conveys the lecturer’s pragmatic intention about future pedagogical activities. Finally, the summary contains little ELF divergence (the adjective *geophysic*, for *geophysical* in SE), but it contains a good number of false starts, repetitions and hesitations. While spoken language typically contains such features, they could combine with other ELF divergences to impair students’ understanding of key content.

To conclude on these two PreFUT summaries (Transcript 4, terminal & Transcript 5, not terminal), neither contains clear position markers making it a terminal or non-terminal summary. They could be found anywhere else in this ending part, or in the other parts of the lecture: there is no solid anchorage, even if their presence towards the end of lectures seems more likely. Both summaries fulfill their PreFUT role: to announce the future lectures. This is done concisely, and the future subjects or lecture contents are not developed at length, which is consistent with earlier research (Alsop, 2015). However, both summaries also contain hesitations or divergent ELF, and passages that could be ambiguous or misinterpreted by students. Thus, these two PreFUT summaries illustrate the great variability in the type, although lecturers tend to develop this type of summary to a lesser extent and with less detail than the other types. We have highlighted the salient features and commented upon the linguistic

realization of the PreFUT pragmatic function: to provide a reference to information to come in the next lectures of the module or course.

4.5.5 Chaining of summaries

This section investigates the chaining of summaries in CoFEL lectures. Chaining is defined by Alsop (2015) as a phenomenon in which summaries of different types occur contiguously and seamlessly. Chaining can occur in any part of the lectures, but the phenomenon concentrates most cases in the beginning and midsection part of the lectures: lecturers seem more inclined to link reviews to previews in the first half of their lectures before delivering the main course content. For Alsop, the occurrence of such chains may reveal a relationship between various summarizing types that spans previous, current, or future lecture content. However, she found that there were no identifiable patterns of co-occurrence of summaries in the ELC corpus. Table 14 displays all the summary chains found in CoFEL, organized by number of summaries per chain: six pairs, two chains of three (including one embedding), and three chains of four.

Table 14

CoFEL: list of summary chains

Nb. of summaries per chain	Chaining type	Nb of chains
2	PreCUR + PreFUT	1
2	PreCUR + RevCUR	1
2	PreFUT + PreCUR	1
2	RevCUR + PreFUT	1
2	RevCUR + PreCUR	4
2	RevPRE + PreCUR	2
3	PreFUT + RevCUR + PreFUT (embedding)	1
3	RevPRE + PreCUR + PreFUT	1
4	PreCUR + RevPRE + PreCUR + RevPRE	1
4	RevCUR + PreCUR + PreCUR + RevPRE	1
4	RevCUR + PreFUT + RevCUR + PreFUT	1
Total nb. of chains		15

Only one instance of PreCUR + PreFUT chains was found in CoFEL: Transcript 6 from an Electro-Magnetic Compatibility course, presented below. The lecturer (L09) introduces a series of lectures and the topic of the day (“today we’ll finish the introduction” – PreCUR) that will end with “two labworks” (laboratory classes – PreFUT) that are also briefly introduced. The relationship between the two summaries is established, since what is announced in the

PreFUT summary continues what was presented in the PreCUR one, within the same engineering course module.

so- normally *today we'll finish the introduction* but as i told you it may be a bit too long. so- maybe we'll manage to (xxx) in another way and then we will go with disturbance sources and how to er- and how to er to er simulate er in an efficient way this disturbances and how to design the filter and how to address the high-frequency behavior of the components. || *the two labworks*. one will be to validate er the e-m-c model of a boost converter. er a comparison with time (domain) simulation and the other labwork which will be six hours is supposed to help you in er the validation of the e-m-c* filter design.

Transcript 6. Single instance of a summary chain combining PreCUR and PreFUT in CoFEL. Italics emphasize the PreCUR marker. The PreFUT marker is boxed. One "x" conventionally represents an unintelligible syllable. A double vertical bar indicates the PreCUR-PreFUT boundary. *e-m-c: Electro-Magnetic Compatibility. [L09 – 20201016_SAJ_EMC1_SUM_0020-0030]

A single instance of PreCUR + RevCUR was found in the corpus, in an Electrical Engineering lecture (L14). The lecturer introduces the students to various electrical devices, some of which are located in the laboratory room. In the PreCUR summary, after presenting some high-voltage equipment, the lecturer announces the content of the remaining part of the lecture to come. The only occurrence of PreFUT + PreCUR found in CoFEL was in a Power Electronics lecture (Transcript 7). In this passage, the lecturer introduces a chapter ("Chapter 2") that will be dealt with in a future lecture, as clearly indicated by the transition phrase "but it will not be today" (emphasized in the excerpt). Then the subject of the current lecture is introduced by "so- we will start with" that follows. In this case, apart from the fact that the two lectures (or "chapters") evoked in the two summaries belong to the same module, it is more a case of juxtaposition of summaries than a relationship: the summaries could be far apart and still function in the same manner.

er chapter 2 is er really related to er our design models. it is really intend to show you how we can model a power electronic converter regarding (3) e-m-c disturbance modeling. (so- we can emission.) so- we really want to be able to model the power electronic converter and i will show you how to er: make this in an efficient way er *but it will not be today*. || *so- we will start with the-* the main origin of the electromagnetic disturbances which is the switching aspects of a power electronic converter.

Transcript 7. Single instance of a summary chain combining PreFUT and PreCUR in CoFEL. Italics emphasize the boundary between the two summary types. The number in brackets indicates a pause in seconds. A double vertical bar indicates the PreFUT-PreCUR boundary. [L10 – 20201023_SAJ_EMC3_SUM_0030-0035]

A single instance of RevCUR + PreFUT chain was found in the corpus, at the very end of an Aeroelasticity lecture (L01). This combination of summaries begins with the conclusion

of the current lecture and continues with the introduction of the program of the new lecture to come. In the same vein, two instances of the RevRPE + PreCUR type were found, combining previous lecture content and forthcoming current lecture content. Finally, four occurrences of RevCUR + PreCUR were found, making it the most common chain type in the corpus. This chain summarizes the back and front of the current lecture, presenting a straightforward articulation between preceding and proceeding content: a relationship is clearly established between the two summaries. Transcript 8 is an example of this chain type. Here, the lecturer concludes a theoretical part of this Cyber-Security session by briefly summarizing what has just been done (*italics*), and introduces an application of the theory through an exercise based on a concrete “industrial system example”.

so- those are the MAIN step in the risk assessment and risk analysis || and er: we'll try to apply those er steps- those main- er: (1) er: main- er: risk assessment er: er: steps- especially i- identification analysis and evaluation on a fictitious- on a very simple industrial system example.

Transcript 8. One example of a summary chain combining RevCUR and PreCUR. Italics emphasize the type markers. A double vertical bar indicates the RevCUR-PreCUR boundary. The number in brackets indicates a pause in seconds. Capitals show the lecturer's original oral emphasis. [L11 – 20210111_UPA_CYBER4_SUM_0100-0105]

The next example is a peculiar case of embedding, with one summary appearing in the middle of another, resulting in a PreFUT + RevCUR + PreFUT chain (Transcript 9). The overall idea in this chain is for the lecturer to preview the content of several future lectures (PreFUT). This is confirmed by the plural “lectures” in the first line and recourse to the Future tense (*we will* —) in the first and third summaries. However, in the middle of this introduction, the lecturer embedded a brief summary concerning the subject covered in the current lecture (RevCUR), in the Simple Past tense, brought in with “but I wanted you to be familiar with”. In the last string of text, the lecturer resumes the initial intention of presenting the full module to come.

||PreFUT|| so- in the- er in the: lectures that we are doing now- and especially in the network- *we will focus* on emission here (2) because *we will er build* a power electronic converter and *this power electronics converter will need* er: to be able to cope with the emission standard. *we will not deal* with the susceptibility standards ||RevCUR|| but *i wanted* you to be familiar with the fact that there are some er: tests which are normally applied to your equipment. ||PreFUT|| so- the standards *we will be using* er in the: in the lab work *which will be* er i think the: industrial equipment or an equivalent standard which is defined in aircrafts er: application- which is the d-o one sixty*- (2) which is in the same spirit.

Transcript 9. Single instance of a summary chain combining three types: PreFUT + RevCUR + PreFUT in CoFEL. The RevCUR summary is embedded in the PreFUT one. *Italics* emphasize type markers. Double vertical bars indicate the type boundaries. The numbers in brackets indicate pauses in seconds. *DO-160: *Environmental Conditions and Test Procedures for Airborne Equipment* is a standard for the environmental testing of avionics hardware.

[L10 – 20201023_SAJ_EMC3_SUM_0010-0012-0014]

To finish, here is an example of a complex summary chain, which presents a RevCUR + PreFUT + RevCUR + PreFUT formula, from lecture L07 (full transcript in Appendix K). With such chains, lecturers can draw up a complete picture of their course modules. This is a session from the middle of a module of Smart Power Industry lectures. In the first RevCUR summary, the lecturer reviews what has been covered in the current session, using the Simple Past tense and typical markers (*the take-back-home information; remember; let me summarize*). Then, in the first short PreFUT string he announces the ground to cover in future sessions (*you will see in further work*). In the second, brief RevCUR summary, the lecturer recalls again what has been done in the current session (*i was pushing toward you*), with more precision about the content. In the second, long and last PreFUT summary, the lecturer announces “further work” to come in the future lectures (*So, further work. Where are we going?*) and concludes the summary chain, and his lecture, with “So, my presentation is over”. Despite its complexity, there is a consistent relationship in the topic throughout the chain.

In conclusion, this section has investigated the chaining of summaries in CoFEL lectures. With fifteen chains in which summaries of different types are linked in a contiguous sequence, the phenomenon concerns about a third of the total summaries in the corpus⁴⁸. Although chaining can occur in any part of the lectures, the phenomenon is mostly found at the beginning and midsection parts of the lectures, which is consistent with Alsop’s (2015) findings. We found that in all but one of the examined chain samples covering all the chain types, the occurrence of such chains revealed a relationship between the various summarizing types, spanning previous, current, or future lecture content. When the summaries in the chain

⁴⁸ 30 summaries are chained out of a total of 104 in CoFEL.

were both of the current lecture content types (RevCUR + PreCUR or PreCUR + RevCUR), the relationships concerned the continuity of the topic and the articulation of two parts of the lecture. When the summaries differed in time frames (e.g., RevPRE + PreCUR or PreCUR + PreFUT), the relationship concerned the continuity and integrity of the overall module or course by linking different course sessions together. In both cases, chains of summaries play an essential role in supporting coherence for both the lecturers and their students, either within the ongoing lecture or within the overall course. In the only case where no relationship was found (Transcript 7), the two summaries just happened to follow each other and could have been separated without changing their function. However, like Alsop (2015), no systematic pattern of summary co-occurrence could be found in CoFEL.

4.6 Summarizing in CoFEL: conclusion and perspectives

In this chapter on summarizing in CoFEL, we identified the salient features of summaries in their summative or predictive roles, and their pragmatic or linguistic realizations. We explained our method: how we used a mixed-methods approach, first quantitative, then qualitative guided by the quantitative, to answer our research questions; how we built an analysis grid to investigate summarizing in the corpus; how we identified and categorized the summaries into four types: reviews of previous (RevPRE) and current (RevCUR) lecture content, previews of current (PreCUR) and future (PreFUT) content; how we developed SumVIZ, a data visualization solution to meet our needs. We presented our quantitative findings: the amount of summarizing activity in the corpus; the share of each of the four types of summarizing activity in the lectures; the token length and duration of summaries; the relationship between summary types and their location in the lectures. Some qualitative findings followed, complementing the quantitative data: the pragmatic and linguistic variation in the four types of summaries; the influence of location and duration, in the beginning, midsection, or ending part of the lectures; the phenomenon of the chaining of summaries.

Our main quantitative findings were, first and foremost, that the treatment of summarizing varies considerably from lecturer to lecturer, in most of the variables: summary frequency, token length or duration (e.g., for duration: 5 seconds for the shortest summary; 7 minutes for the longest). This wide variability means that no particular pattern concerning these variables emerges at corpus level, in the lecturers' implementation of the pragmatic function, except that lecturers favor summaries of the current lecture content, RevCURs and PreCURs.

We found that there were no statistically significant differences either in number of tokens or in number of seconds, between episodes of different summary types. Secondly, concerning the location of summaries, one significant outcome was revealed, thanks to an intersubjective decision to divide the lectures into three congruent parts (beginning, midsection, and ending): the location of summaries and their type are largely dependent. Thirdly, we found that some types of summaries were completely absent from certain parts of the lectures: this is the case of RevCURs in lecture beginnings, and RevPREs in lecture endings. Finally, that there are almost three times as many occurrences of the RevPRE type in the beginning part of lectures than of the other types, and twice more of the PreFUT type in the ending part.

With regard to our qualitative findings, we chose to examine the pragmatic and linguistic variation in the shortest and longest summaries in the corpus, for each of the four types in turn. We revealed how the reiteration (reviews), or anticipation (previews) of the lecture or course content was realized pragmatically and linguistically, by highlighting the salient features of these summaries in detail, whether they were brief or extended. The main observation is that summaries of any length can play their expected scaffolding role. While short summaries constitute simple links between two parts of the lecture, long summaries usefully enumerate partial or full lecture outlines, albeit with the increased risk of including long-winded passages, hesitations, or ELF divergence. As far as this ELF divergence is concerned, the odd disfluency, mispronunciation, or non-standard form is unlikely to be problematic. Rather, it is when these features combine that they can have more influence on the quality of communication: for example, a shift in word stress, combined with unreduced vowel sounds and elided final consonants in verbs can lead to more serious consequences.

We also examined archetypal summaries for the three parts of the lectures: beginning, midsection and ending. We found that there was a certain degree of dependence between the type of summaries and their location at the meso level of the lecture part, but not at the micro level of their position within the part (e.g., initial or delayed, or terminal or non-terminal made no difference). All the episodes examined play their role as summaries, and their location in the lecture is generally adequate for this role, even if, once again, there are repetitions, hesitations, or ELF divergence which may hinder this role.

The last qualitative aspect examined was the chaining of summaries into summarizing sequences of various types, and of various numbers of episodes. With a third of the summaries involved in chains, this phenomenon was worth investigating, as it is a common practice. Our

findings show that chains of summaries contribute to establishing coherence for both the lecturers and their students, either within the ongoing lecture or within the overall course. The presence of such chains reveals a relationship between the various summarizing types, spanning previous, current, or future lecture content. This relationship may either concern the continuity of the topic and the articulation of two parts of the lecture (summaries in the chain are both of the current lecture content types), or may concern the continuity and integrity of the overall module or course by linking different course sessions together (summaries differ in time frames). However, like Alsop (2015), we found that no recurring pattern of summary chaining could be found in CoFEL.

Despite its wide variability at all levels, summarizing is a rather common pragmatic function in CoFEL, being present in all of the lectures, whether in minimal or considerable amounts. We have highlighted its lecture scaffolding role by analyzing its pragmatic and linguistic realizations in detail, through a range of examples representative of the four types. Overall, our results are fairly consistent with previous research, but sometimes to different degrees. One of these differences lies in the apparent higher presence of ELF divergence in CoFEL, when compared to the Engineering Lecture Corpus (ELC; Nesi et al., 2014; Alsop, 2015), for instance. This is mainly due to the fact that all CoFEL participants speak English as an additional language. To varying degrees, the linguistic challenge of teaching in EAL could contribute to the lecturers' repetitions and disfluencies mentioned throughout this chapter. In addition, the high cognitive load imposed by the combination of EAL and online teaching could lead to increased fatigue for both lecturer and students (Roussel et al., 2017; Roussel & Tricot, 2017).

5 Storytelling

In graduate schools of engineering, lecturers of content courses (i.e., all courses other than foreign language courses) commonly have professional experience in industry or research. To unpack what engineers truly do and think, many lecturers use narratives based on personal background and professional experience to complement or illustrate more traditional, bookish, or formalistic content. Thus, stories are a sign of professional competence (Dyer & Keller-Cohen, 2000).⁴⁹

5.1 EMI, engineering lectures and storytelling

When communicating science to nonexperts, Dahlstrom (2014) recommends the use of storytelling. A good story contextualizes content and encourages emotional connections, which may favor learning (McNett, 2016). Landrum et al. (2019) also emphasize ‘the pedagogical power of storytelling’ and advise implementing storytelling as an educational approach, praising its virtue in achieving desired learning outcomes. Finally, Nesi and Alsop (2021, p. 25) found more frequent storytelling in experienced lecturers’ discourse, in both EMI and L1 contexts, when content was more complex and abstract. Further, they argue that stories help lecturers to make sense of their experience, and that narration meets psychological and cognitive needs, such as breaking from routine or building rapport.

Research addresses aspects of storytelling which are pertinent for EMI. For example, Alterio and McDrury (2003) promote it as a teaching and learning tool and encourage a ‘storytelling culture’ in higher education. Andrews et al. (2009) go into more detail, identifying four story-based instructional methods: case-based (learner as an outside observer), narrative-based (learner within the narrator’s context), scenario-based (learner in an interactive, real-time experience) and problem-based (learner as the director of learning activities). Considering the learner’s perspective is also important for Abrahamson (1998), who stresses the role of stories in reducing depersonalization and promoting critical thinking, and refers to them as the ‘foundation of the teaching profession’. This may partly be due, as argued by Gibbs (2014), to the way narratives give force to evidence, anchor experience in time, enable inferences, and favor appropriation of knowledge. Moreover, Gibbs highlights the way narratives contribute to

⁴⁹ Most of this Chapter 5 was published as an article (Picavet et al., 2023): “Storytelling in L2 English-medium engineering lectures: A typology”.

“the social transmission of experience” (2014, p. 4). Finally, Martin observes the social dimension, noting that “both discourses are social ones. They form community.” Martin (2008, p. 53) and explains that storytelling complements traditional scientific lecture discourse.

5.2 Defining the terms ‘story’ and ‘storytelling’

What is a story? According to Genette (1983), a story is an act or an event reported in text; it relates a transformation through time, or the passage from a previous state to a resulting subsequent state. Labov and Waletzky’s seminal work (1967), presents an analytical framework for the analysis of oral narratives of personal experience, defines the basic units of a narrative, and outlines its ‘normal structure’. What do stories do? For Bruner (1990), a story organizes experience and is rooted in our ancient heritage. It appeals to the human organizational predisposition for the narrative form. We will use ‘story’ as an umbrella term to designate story types and retain the following basic definition of ‘story’: “a description of how something happened, [will happen, or could happen], that is intended to entertain people, and may be true or imaginary”.⁵⁰ The term ‘storytelling’ will designate the act of telling a story. Finally, a ‘narrative’ is one of the six sub-categories of story types. All the stories referred to belong only to spoken discourse.

5.3 Research questions

In response to calls for more EMI corpus research (e.g., Jablonkai, 2021), we will use corpus linguistics methods to analyze academic discourse in the EMI context. We hypothesize that lecturers tell stories to their international student audiences, when delivering their engineering lectures in English as an additional language, in graduate schools of engineering. The research objective is to find, delineate and categorize these stories. Our context of study is pedagogical situations specific to a Master’s degree in engineering (MEng), in France. Therefore, we address the following research questions:

- How can stories and their boundaries be identified in EMI pedagogical situations?
- How can the stories, once found, be categorized?

⁵⁰ Longman Dictionary of Contemporary English (online) [accessed March 2024]: <https://www.ldoceonline.com/dictionary/story>

The next section describes how an analytical framework for studying storytelling was built to investigate storytelling in our corpus.

5.4 Method

5.4.1 Building an analytical framework for studying storytelling

In this section, we build an analytical framework enabling the identification and categorization of potential stories. Alsop and Nesi explored storytelling extensively in their multinational Engineering Lecture Corpus (ELC) (Alsop et al., 2013; Alsop & Nesi, 2014; Nesi et al., 2014; Alsop, 2015), highlighting the fact that several classifications exist. These range from ‘classical’ Labovian Narratives (Labov & Waletzky, 1967), to Recounts, Exempla, Anecdotes (Martin, 2008), Story-Likes (Alsop, 2015), or Scenarios (Nesi & Alsop, 2021). Labov and Waletzky’s definition of a narrative requires it to include a strict temporal sequence that can “recapitulate past experience in the same order as the original events”. However, some researchers have found Labov’s approach too narrow or somewhat selective, and “an extremely problematic one” (Martin, 2008, p. 44). Martin proposes an extension of the notion of a single narrative structure by identifying four types of stories: 1. Recount, 2. Anecdote, 3. Exemplum, and 4. Narrative. Alsop (2015) concurs that Labov’s model is not fully satisfactory and readily adopts Martin’s approach (Alsop et al., 2013, pp. 8–9). Nonetheless, to accommodate field data, Alsop added ‘Story-likes’, and Nesi and Alsop (2021) further added ‘Scenarios’ (see Figure 30). These story types will now be defined.

Narratives, in the typical Labovian sequence, involve disturbed and restored equilibrium, and are broken down into six phases: an optional abstract, an orientation, a complication, a floating evaluation, a resolution and an optional coda. The Labovian analysis highlights how Narratives are constructed and the functions they perform. It also informs how narrators perceive, react to, or give shape to events, how they make a point; how they perceive themselves, or whether they recognize the conventions of storytelling. The evaluation element plays a major role: it tells the audience what the narrator feels about the events. It can add more evidence to the overall story and provide moral elements.

Three other types of stories (Anecdote, Exemplum and Recount) were proposed by Martin as an improvement to Labov’s classification. Martin argues that Labov standardized his corpus, narrowing the range of narrative functions and formats, by orienting participants to

relate stories on a single theme and sometimes asking them at the end about what happened next (2008, p. 44). In Martin's classification, Recounts narrate unproblematic events and take the form of running commentary. Anecdotes and Exempla narrate problematic events which are not resolved. Anecdotes tell of unexpected disruption generating emotional empathy, whereas Exempla deal with a noteworthy incident leading to moral judgement.

Alsop's 'Story-likes' deal with events that do not systematically belong to the past, often having a present-continuous, or permanent character, yet are presented by the lecturer in a similar fashion to a Narrative (Alsop, 2015).

Finally, 'Scenarios' are defined as hypothetical events, when lecturers invite their audiences "to imagine a fantastic hypothetical situation", in sequences in which they stray away from the delivery of facts "into the realms of fantasy" (Nesi & Alsop, 2021, p. 25). These authors report that "scenario-telling is part of the experienced lecturer's repertoire and is an engaging way of elaborating abstract technical terms, concepts, or arguments central to a disciplinary topic". Our acceptance of the term 'scenario' is broader than that of Nesi and Alsop, as it includes hypothetical events that are not fantastic.

5.4.2 Identifying stories in CoFEL

Identifying storytelling instances in CoFEL involved three stages: (1) detecting typical morphosyntactic story markers in lexis, tenses, or personal pronouns; (2) locating shifts in stance, in the linguistic register or the pragmatic function; (3) determining the boundaries of a story, that is to say, where a story begins and where it ends. Based on examples of stories in the literature, we selected search terms to identify passages in CoFEL that could potentially function as stories (Alsop et al., 2013; Alsop, 2015; Nesi & Alsop, 2021).

Firstly, a word search was carried out of sixteen terms likely to be found in the six story types (Table 15).

Table 15

Selected search terms to identify potential story passages in CoFEL

analogy	anecdote	example	happen
if you	imagine	let me	let's say
picture	say you	scenario	story
suppose	tell	what if	you know if

In addition, we searched for words relating to incidents or accidents in the field of engineering (Table 16). Items from this lexical field are frequently found in engineering stories (Alsop, 2015). The canonical form of words (lemma) was used in the search, with a tolerance for all inflections of a given word (e.g., ‘catastrophic’, from ‘catastrophe’).

Table 16

Selected search terms relating to incidents or accidents in engineering.

accident	awful	casualties	catastrophe
cause	consequence	damage	death
disaster	hurt	incident	injury
lives	problem	terrible	tragic

Tenses other than the various present forms are strong indicators of the presence of a story in a lecture (Alsop et al., 2013; Nesi & Alsop, 2021). Thus, CoFEL was searched for the nine past, future and conditional auxiliaries listed below (Table 17), including two contracted forms frequently found in oral texts.

Table 17

Selected search terms relating to tenses other than the Present

did	had	should
was	were	will
would	—'d	—'ll

Moreover, by using personal pronouns lecturers address their student audiences directly, and therefore, may indicate the presence of stories. For Luukka (1992, pp. 79–80), the ‘I’ pronoun signals the presence of the lecturer; the collective ‘you’ indicates the presence of the audience; ‘we’ marks the presence of both; all of these are “signals of interactional attitudes”. Thus, CoFEL was searched for the three pronouns ‘I’, ‘you’ and ‘we’. When the searches for lexis, tenses and pronouns pointed to the same passage, it was considered to be a candidate story for further analysis.

Secondly, we used Martin’s model (2008, pp. 53–54) to locate shifts in stance, register or pragmatic function. Building on Bernstein’s (1999) work on vertical and horizontal discourse, the model lists types of shifts within discourse: from scientific to artistic, technical to accessible, logical to rhetorical, rational to humane, objective to subjective or institutional to local. A typical shift from exposition to narration is when lecturers move from presenting rational scientific data to an experienced real-life situation. Such a shift can be characterized by changes in physical, attitudinal, or moral positioning (e.g., moving closer to the audience, looking for eye contact) and in type of discourse (e.g., from lecturing to building rapport, entertaining, telling historical facts or anecdotes, or attempting humor). Likewise, lecturers may interrupt the exposition of scientific content in order to invite students to imagine themselves in a real-life engineering situation. We searched through CoFEL (video, sound, and transcripts) and identified such shifts from scientific discourse to discourse directly addressing the students.

Thirdly, we established the boundaries of the stories. We decided to include the lead of a story when it was essential to understanding the overall story. In case of rater disagreement, we decided to set the boundaries so as to include the longer version of the story. Moreover, some stories contained one or several sub-plots. When a sub-plot was self-contained, we considered it a separate story. On the contrary, when several sub-plots were interrelated, they were taken together as one complex story. Some candidate stories were rejected outright, such as Transcript 10 which presents a condition (‘if you have’), but lacks even the rudiments of a plot:

so that's the correct idea. it doesn't really make sense to study the cases corresponding to: velocity which are beyond the: divergent velocity- right- because if you have already broken your: structure then you don't really matter- you don't really focus on the vibration of it.

Transcript 10. A candidate story (with a condition) that was rejected.

[20200117_FAC_SCEN_10]

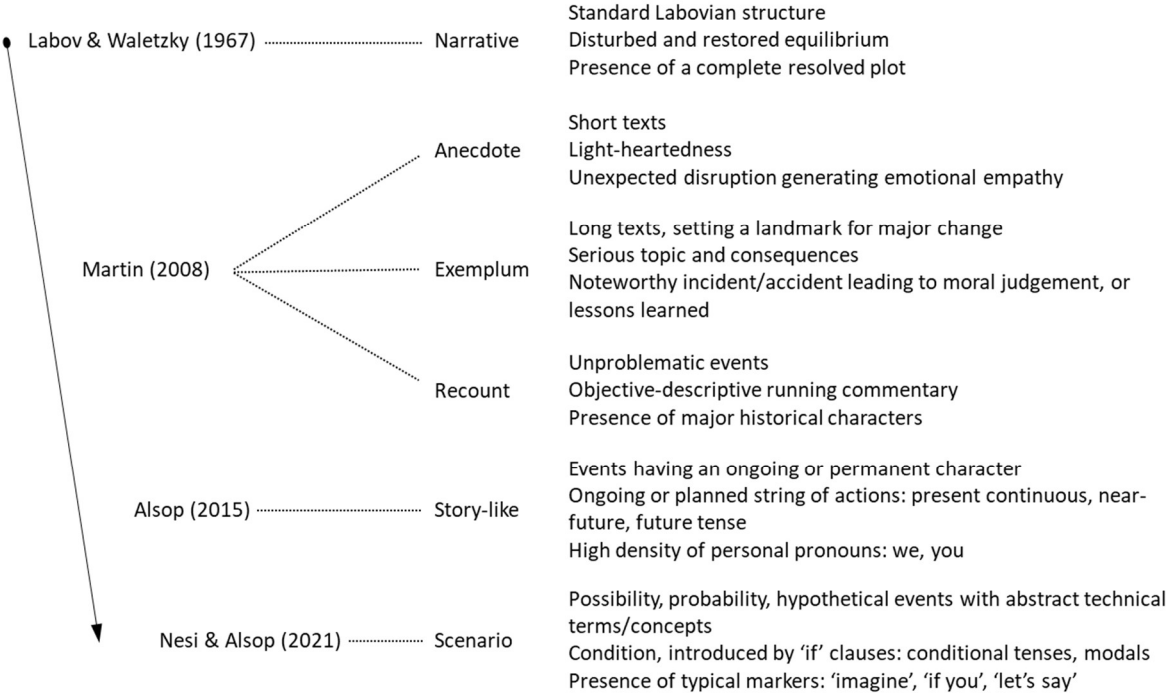
To conclude, decisions about the inclusion of a candidate story were made by combining the three steps above. The next section examines how those instances of storytelling were categorized.

5.4.3 *Categorizing stories in CoFEL*

An analysis grid was developed based on the literature and six types of stories were retained for their likely presence in EMI engineering lectures (Labov & Waletzky, 1967; Martin, 2008; Alsop, 2015; Nesi & Alsop, 2021). Figure 30 presents the grid, with specific

features for each type, and shows how researchers have built on each other’s work to enrich the story typology.

Figure 30
The six story types and their features



The three raters independently categorized the twenty lecture passages containing candidate stories, and then they compared the results in a series of four rounds. Firstly, two more candidate stories were eliminated. Although they did contain some typical story markers for a Scenario as in the example below (‘we can imagine’, 2nd line), the topic did not develop into a plot (Transcript 11).

the sedimentary rock comes from deposits er of different materials, so we have er layers er: so: we can imagine that- that if we: load the layers er: er: perpendicular to the layers or on the axis of the layers we don't have the same characteristics, ok?

Transcript 11. A candidate story (containing a Scenario marker) that was rejected.
 [20200213_ELO2_SCEN_10]

Secondly, the raters agreed on eight of the eighteen remaining candidate stories: three Exempla, two Anecdotes, two Recounts and one Scenario. Two types were absent from this first list: Narratives and Story-likes. Thirdly, seven candidate stories on which two out of the three raters agreed were discussed. In most cases, these candidate stories contained linguistic or semantic inaccuracies due to the use of EAL or to weak storytelling abilities. In some other cases, the

raters discussed how to apply the category definitions. For example, for one story, we decided that an object (an old farmhouse), rather than a person, could be the main character of a Narrative. After discussion, full agreement was reached on three Story-likes, two Anecdotes, one Narrative and one Scenario. Finally, the remaining three candidate stories were discussed and these cases also involved interpretation of the category definitions. For example, for one story, we decided that the UK government was the main historical character in a Recount. The final decisions were: one Exemplum, one Recount and one Story-like.

Thus, the corpus contains 18 stories: four Anecdotes, four Exempla, four Story-likes, three Recounts, two Scenarios and one Narrative. In the next section, selected storytelling excerpts from CoFEL will illustrate each of the six story types.

5.5 Exemplification of the typology: stories found in CoFEL

In this part, we comment upon sample stories found in CoFEL. The lecture recordings were transcribed orthographically without modifying divergent morphosyntactic forms. The sign ‘[...]’ indicates non-essential passages (repetitions, hesitations, etc.) that have been abridged for the reader’s convenience and lack of space. Titles have been given to stories for easier reference.

5.5.1 Narratives – *The Drôme Farmhouse*

A Narrative is defined by Martin (2008, p. 43) as “the relations between events and feelings”: this is the foundation of the structure, and function, of stories. A typical Labovian-Martinian ‘narrative’ comprises six stages: Abstract (optional), Orientation, Complication, Evaluation (floating), Resolution, Coda (optional). We will use this first sample⁵¹ from CoFEL to describe in detail each of these stages.

Abstract

This Narrative starts with an Abstract, answering the question ‘What is/was this about?’ (Gibbs, 2014). This provides a summary of the events or makes a general proposition, which the Narrative will exemplify. Although optional, this stage is present in the Soil & Rock Mechanics story below, as well as in most stories in CoFEL. The topic of this excerpt is the

⁵¹ All the excerpts in this part belong to the same corpus transcript [20200206_CAM2_NAR_10].

presence of water in soils, whether saturated or submerged, and its consequences (Transcript 12).

a soil which is under the groundwater level do not have the same weight as the soil which on the top of the groundwater level. [...] i call it effective specific weight because it's something which [...] is equal to the saturated volumic weight minus the volumic weight of water. [...] ok? i'm going to give you a short example of why it is super important. (8) {lecturer is erasing chalkboard}.

Transcript 12. Labovian Abstract example. [...] indicates abridged passages. [20200206_CAM2_NAR_10]

The lecturer stresses the importance of considering the presence of water in soils for the construction of houses or buildings, setting the theoretical backdrop for the ‘short example’ mentioned at the end of this first transcript, the story to come.

Orientation

The next stage, the Orientation, is indispensable and provides key information about the time, place, situation, and participants of the Narrative. In the next excerpt (Transcript 13), the date of the event is given, “in August 2003”, as well as the location at various levels of granularity, “in France”, “in the *département* (French for ‘county’ or ‘district’) called ‘la Drôme’” (southeastern France) and “close to here”. As for the participants, there is “a friend” of the lecturer, but the real object of the Narrative, a house, also becomes a character in itself, referred to “an old house” and “a farmer(’s) house”. However, most of this Orientation is dedicated to the description of the event, of ‘what’ happened there, while the lecturer is drawing a rough cross-section sketch of the scene on the chalkboard . This part of the Narrative prepares the audience for the Complication that is to come, by building tension or expectation.

ok well i should tell you a story that i know because it took place where i have a friend which is living there. {lecturer is drawing on chalkboard} there are mountains. this is the bedrock. ok? and here this is filled with some (sediment) soil. ok? it's close to here. it's er in the department called la drôme. ok? and here there is a house. it's an old house made of stone. [...]er then if you have a motion there is no resistance to tensile (in fact). so this is why we can see a lot of old houses made of stone with big crack- vertical (crack). [...] this was a lake in many- (well) many years ago. there is still (2) a groundwater (2) table here. ok? and you know what happened in august two thousand and three in france? (you're) maybe a bit too young.

Transcript 13. Labovian Orientation stage example. [...] indicates abridged passages. [20200206_CAM2_NAR_10]

Complication

Following the Orientation comes the Complication—or Complicating Action—stage. This stage is also essential and represents the climax of the Narrative, as it introduces the series of events that build the action itself, the problem. It can involve cycles, turning points, crises or problems, and shows how they were dealt with. In Transcript 14 below, the lecturer describes the heat-induced drought of summer 2003 (‘a climate situation where the temperatures were too high’), which provoked large cracks in many houses (‘you have cracks of twenty centimeters’), due to the evaporation of the groundwater contained in sedimentary soils (‘the groundwater level moved down’), presumably under the influence of climate change: ‘it was probably the first time in the history of this field’.

there was a climate er situation where the (standard) temperature were too high. ok? during all august. ok? so all the groundwater level moved down. ok? [...] er we don't know exactly maybe one meter something like that. ok? so what happened? what is the vertical stress that compress the soil which is here? it's the stress due to that soil which is let's say wet with a given specific weight and that soil which is saturated. [...] so let's say it is- it was here twenty kilo newton per meter cube and here it was er ten er kilo newton per meter cube. ok? [...] so the weight of this column of soil was higher due to the fact that the groundwater level moved down. so the soil which is under here was compressed by the weight that was higher. ok? so here it opened. and here {lecturer is pointing to the top of the farmhouse} you have cracks of twenty centimeters. ok? so the house was very close to- to crash.

Transcript 14. Labovian Complication stage example. [...] indicates abridged passages. [20200206_CAM2_NAR_10]

Resolution

The Evaluation, Resolution and optional Coda stages come next. The location of the Evaluation is not fixed: it can come before, after or coincide with the Resolution. In the excerpts below, it comes after. As for the Resolution stage, it answers the question ‘What finally happened?’ (Gibbs, 2014), thus describing the outcome of the events related in the Complication, or the solution of the problem. It serves as a conclusion to the Narrative. In Transcript 15, the Resolution details the rather far-reaching course of actions taken to deal with the issue at stake: the roof and part of the walls of the farmhouse had to be removed.

so to do that in fact they unbuild the roof. they removed half of the high of the- of the walls to reduce the weight. ok? and to- to preserve the- the farm.

Transcript 15. Labovian Resolution stage example. [20200206_CAM2_NAR_10]

Evaluation

The purpose of the floating Evaluation stage is to highlight the significance or the meaning of the Narrative. Here, the lecturer broadens the picture to the rest of the country, as the historical drought that year provoked a nation-wide lowering of the level of groundwater tables and aquifers (Transcript 16). The consequence for houses in many French regions was the appearance of sometimes serious structural cracks, when water evaporated, and soils dried out and lowered.

but this was- there was a lot of situation like that in france in fact. there was a lot of houses where there was a lot of cracks that appears due to this specific er increase in the temperature due to the fact that the groundwater table moved down. ok?

Transcript 16. Labovian Evaluation stage example.

[20200206_CAM2_NAR_10]

Coda

As in music, an optional Coda can adorn the end of the Narrative, answering the question ‘And now?’. Its purpose is to mark the end of the Narrative and to connect the text to the present or some other Narrative (Transcript 17). Here the lecturer finishes by recalling earlier references in the lecture to Archimedes’ principle (more commonly referred to as ‘buoyancy’ in English).

so this is due to the archimede(s) variation let's say of er the volumetric weight that changes due to archimede(s) force.

Transcript 17. Labovian Coda stage example.

[20200206_CAM2_NAR_10]

This extended, six-stage Narrative fits the “familiar equilibrium disturbed, then equilibrium restored motif” (Martin, 2008, p. 43) rather well: the events are resolved at the end of the ‘plot’ and the old, seriously endangered farmhouse is salvaged, albeit at great labor and cost, in the Resolution stage. The pedagogical potential of this story is high, as a cracked old farmhouse can be rather easily pictured by students. As successive stages of the story unfold, they may become slightly worried about the outcome: will the house eventually collapse? Some might even sympathize with the “poor, old house”, and be happy with the favorable outcome, especially if they had experienced a similar situation themselves. The reference at the beginning to a ‘friend’ of the lecturer anchors the story in reality. The lecturer’s appeal to the (too) young audience’s own memory about the 2003 drought in France may work in a similar fashion. There is some tension-building when the lecturer points to the gable end of the house on his drawing and evokes the enormous 20-centimeter cracks threatening its integrity. The lecturer-narrator gives shape to the event and sees himself as an expert in the domain, inviting his audience to

do the same. The rhetorical style may persuade current and future engineers of the need to take into consideration the nature of the soil in relation to climate change.

The next sample stories will be described using Labovian terminology, but none of them will be considered a Narrative, as explained.

5.5.2 *Anecdotes – The Untimely Airbag*

Martin (2008, p. 44) maintains that Anecdotes and Exempla are stories that contain an event that is problematized, but not resolved. Alsop (2015) specifies that the distinction is made at the level of reaction: Anecdotes elicit emotional empathy, whereas Exempla elicit a moral judgment. From a cultural viewpoint, Alsop found that UK-L1 speakers of English in the ELC often have recourse to the type, more than EAL speakers. Moreover, although Anecdotes may have quite serious topics, “the costs of the described complications are not grim and/or do not have serious/negative consequences” and they “introduce concepts in an entertaining and memorable way, rather than in elaborate or technical detail” (2015, p. 268).

The Anecdote reported below, from a Power Electronics lecture, corresponds to the above description: potentially serious incidents in a German car, due to the untimely inflating of airbag cushions (Transcript 18). We categorized it as an Anecdote, not an Exemplum, as although it could have had serious consequences (e.g., a car accident), none are reported. There is no ‘moral judgement’ either, contrary to the Exemplum about a burst dam described in the next part.

Abstract

ok er electromagnetic compatibility can be separated into emission and susceptibility. the emission represents the capability that the electrical device has to disturb some electrical equipments and this disturbance can be achieved through either the wires- that’s conducted emission- or through the air- through electromagnetic wave which is er radiated emission. on the other hand the electronic devices can be disturbed through either the wires- that’s conducted susceptibility or through the electromagnetic waves- that’s radiated susceptibility.

Orientation

so this er science of electromagnetic compatibility started in the late eighties er: or a bit earlier in the seventies er because of the first disturbance which were er encountered by people.

Complication

of course the world becomes more and more electrical and more and more electronics and we had a lot of er new issue er which (have) been er- er reported in the literature. er i told you already i think that er a german car had a lot of safety airbag(s) and when you turn your light(s) on er it generate(s) the- the starting of the airbags (@) which is of-obviously very er uncomfortable for the driver.

Evaluation

and that was er due to the proximity of two wires too close together. the wire of the sensors of the airbag and the wire of the light(s).

Transcript 18. Anecdote example.

[20210104_SAJ EMC]

This Anecdote opens with an abstract exposing general-purpose definitions and remarks on the electromagnetic compatibility of electronic devices. It is delivered in a light-hearted manner, with a humorous twist to it (‘the starting of the airbags [...] is obviously very uncomfortable for the driver’), as a car cannot be driven with inflating airbags. This Anecdote example lacks a resolution stage, a shared feature with *Exempla*. Like the preceding *Recount*, it illustrates the pedagogical potential of stories rather well: it is anchored in time (‘in the seventies’); it relates to commonly shared objects and knowledge (a ‘car’ as a vehicle and ‘airbags’ as standard equipment in all modern cars); it inserts situational humor into the lecture, perhaps to provide a brief, light-hearted interlude during an otherwise serious lecture.

5.5.3 *Exempla – The Malpasset Dam Disaster*

Exemplum is one of Martin’s four story types (2008), reporting on noteworthy incidents and entailing a moral judgement. In both the ELC and CoFEL corpora, that notion of judgement extends to matters which are scientific. In this type of story, lecturers tend to use longer strings of text, giving a greater wealth of detail, perhaps driven by the compelling concluding judgment of a serious incident or catastrophic disaster. The events reported in the type contain complex or far-reaching ‘complicating action’ stages. Moreover, such events spur major changes in engineering techniques or approaches. Classical themes developed in those stories include financial or health-and-safety responsibility. A key word analysis shows that *Exempla* are located where engineering structures are built or operated, staging “scenes in which the themes of death and mutilation are strongly linked to situations in which students may find themselves during their future careers” (Alsop, 2015, p. 264). It could be argued that lecturers use the *Exemplum* type of story to help students to picture themselves in a professional position, with great responsibility for themselves, their teams, or populations in general.

The lecture excerpt below is about the catastrophic dam burst at Malpasset, in southeastern France (Transcript 19). The Malpasset Dam was an arch dam on the Reyran River, located a few kilometers north of Fréjus, on the French Riviera. It collapsed in December 1959, killing over 400 people in the resulting flood and causing damage estimated at €56 million (US\$68 million). This lengthy story is presented using the standard Labovian breakdown to make it easier for the reader to grasp. However, its dramatic content, the fact that it set a landmark for major changes in rock mechanics approaches and led to the learning of lessons for future engineering projects does categorize it as an Exemplum.

Abstract

so rock mechanics is a quite young science and er: events in the sixties er highlight the lack of knowledge on rock mechanics and er because of a lot of large works that was built in the post war investment plan. [...] so one of the event is the malpasset dam failure.

Orientation

so mal- malpasset dam is an arc(h) dam that was built er by er andré coyne which was a famous engineer in the dam design. [...] so at this period er it was because a lot of dams that was built. it was a period of improvement on the design method and the arc(h) dam was thinner and thinner and the malpasset dam was really thin. [...] the aim of this dam is the water supply of fréjus and saint raphaël- the two towns in the south east of france. [...] the end of the building of malpasset dam was in nineteen fifty four. [...] one of the problem yes was that there weren't rains enough and really less than the one that was expected when they designed the volume of the water storage. so the level of the lake was quite low. [...]

Complication

and during december nineteen fifty nine er there were a lot of intense rainfalls and so the level grew up really fast and er: reached the: highest level of the dam. and unfortunately er: with this high level of water er the dam collapsed and especially on the left part of the dam. and millions of cube meter flew on the downstream of the dam and er: there were a lot of damage on roads- railways- building(s)- and there were more than four hundred er deaths in fréjus and saint raphaël.

Resolution

er after this failure er there were a lot of studies to: explain what happened. and the conclusion of this study was that there were a problem on the: foundation on the rock mass on the left bank of the dam. er there was a large fault that was not expected when the dam was built. [...] so then er the block on which the dam was er founded er slid and then failed. so all the water could er get out of the dam on the downstream part of the dam.

Evaluation

so at this period techniques to er: investigate the rock mass before building any dam were not really developed and er: the focus was more on the structural design of the dam than on the foundation characteristic of the rock mass.

Coda

so after this disaster er: there were a lot of studies on the rock mass characteristics and a lot of improvement in the investigation techniques er especially to know the characteristic of deep part of the rock mass. and today er in such situation the fault of malpasset would have been known before building the dam.

Transcript 19. Exemplum example. [...] indicates abridged passages.

[20200213_ELO_SRM]

In recounting the dreadful Malpasset disaster at length, the civil-engineering lecturer may be attempting to impress an ingenuous student audience and to convey a powerful sense of responsibility for the future decisions and actions. There is a clear moralizing intention, as the engineers of the time did not have the knowledge necessary to detect the presence of an underground rock fault, with dramatic, far-reaching complicating consequences and hundreds of casualties ('today [...] the fault of Malpasset would have been known before building the dam'). The lecturer highlights the potentially horrific consequences of a lack of expertise or preparation on the part of professionals. The story also insists on the rock mechanics experience acquired in the aftermath of the disaster ('after this disaster, there were a lot of studies on the rock mass characteristics and a lot of improvement in the investigation techniques'). This is another typical feature of Exempla: explaining how dramatic events may trigger major changes in the field.

5.5.4 *Recounts – The Seven Bridges of Königsberg*

Recounts tend to be straightforward and rather concise, usually containing neither complication nor resolution. This story type is meant to be informative, chronological, and merely descriptive of a series of events with which the lecturer will usually keep a distance. Martin states that the Recount type depicts 'unproblematic events', in a 'running commentary' (Martin, 2008). In Alsop's definition (2015, p. 259), the lecturer-engineer "conveys information about events in the past in chronological order". The type is characterized by any combination of the three basic components (places: Where? – people: Who? – engineering structures: What?) and Alsop insists on the frequent presence of renowned historical personalities.

Transcript 20 is a clear example of the Recount category, as it matches the criteria above perfectly. It invokes the famous 18th-century Swiss polymath, Leonhard Euler. Euler made influential discoveries in many branches of mathematics, such as infinitesimal calculus

or graph theory. As such, Euler fits rather well in the ‘major historical name’ (Who?) category. Likewise, engineering structures are referred to, here the “Königsberg Bridges” (What? – Where?). They are invoked more as a well-known mathematical problem (known as “The Seven Bridges of Königsberg”) than for their structural design. The city of Königsberg in Prussia (Kaliningrad, in today’s Russia) was established on the Pregel River over two large islands. These were connected to each other and to the mainland portions of the city by seven bridges. The problem is to devise a pathway through the city that crosses each of those bridges once and only once. In the example below, the lecturer addresses the issue of the complex management of the electrical-power grid through modern computer science, a metaphor derived from Euler’s graph theory. This is acknowledged in the last third of the story.

complex networks is an evolution of the graph theory. er all of you know er graph theory er that is starting from leonhard euler when he had to solve the problem of königsberg bridges. er: is he able to go to every part of the city er going through one bridge only. in fact he built a mathematical theory around it just to prove that it was impossible. this is what we did use er till nineteen sixty with random graph theory. somehow we had a lack of er- interest in graph theory just because the problem were becoming larger and larger and we didn't have the er- let's say- the computation ability to solve them. and so in parallel we need to add a new field that is computer. and thanks to the er big database the er processing power that was increasing more and more. in fact complex network theory appears er in the late nineties with small world networks and scale free networks. and this complex network theory er was really helpful for instance to study er typically social medias and pandemic situation.

Transcript 20. Recount example.

[20200820_PAC_NAR_30]

Transcript 20 illustrates the didactic value and pedagogical potential of stories for three reasons: a) it provides chronological anchoring (‘till 1960’); b) it contextualizes knowledge and experience (Euler’s graph theory evolving into modern complex networks, with the development of computer science); c) the metaphor (bridge route / complex networks) is likely to stick with the students, to favor new knowledge appropriation, with an emphasis on problem-solving.

5.5.5 *Story-likes – The Bubbles (a cable-car)*

According to Alsop (2015, Section 7.4.5), the pertinence of Story-likes to engineering domains is that they directly orient towards the future: these quasi-stories act as an invitation to learn and join the engineering community. Dealing with unreal events like Scenarios, they differ from them by using hypotheses or plausible predictions, not purely imaginary, fantastic

conceptions. Lecturers do this with caution, awareness, and scientific judgement, considering the huge responsibility that designing and building a dam, a nuclear-power plant, or a cantilever bridge represents. Alsop argues that Story-likes enable lecturers to embrace their expert roles vicariously, acting as mediators between the lay and professional worlds. Scientific and engineering knowledge is made accessible and young aspiring professionals are more likely to remember the content.

Transcript 21 does not qualify as a Narrative, because of its dominant present and future tenses, permanent character, and direct-address to audience style, with numerous personal pronouns (the transcript below maps the distribution of personal pronouns ‘I’, ‘you’ and ‘we’, showing their rather strong presence). Yet, the excerpt breaks down—imperfectly—into a Labovian structure and is thus a legitimate Story-like. We segmented parts of the excerpt using the Labovian approach and introduced three extraneous stages (Questioning, Humor and Exemplifying), to describe non-Labovian strings of text.

Abstract

er so one thing that can be interesting when you make surveys is to extract the soil to identify the different soils. [...] also because depending on the distance between the ground and the bedrock you can then select the kind of foundation that you can use for er- for your building.

Questioning

do you have an idea of the distance between our feet and the bedrock?

Orientation

[...] here we have two river(s). we have drac over there we have isere over there. the level is not the same. so drac is higher than isere. so there is a flux of water from one river to another one. this clean everything. ok? [...] we have some er- we have some clean gravel and after in fact we have big layers of silt and (clays). ok?

Questioning

do you have an idea of what there was here in that valley many many MANY years ago? [...]

Complication

yes there was some ice (4) it dig the rocks ok? [...] after everything melt there was some lakes and so on ok? so there was a deposit of sediment ok? this is why we have a lot of clay ok?

Questioning

do **you** have an idea of what **we** have here under our feet?

Humor

so **i** would suggest something. if **you** don't like sport **you** go to bastille with the bubbles {lecturer imitates sound of cable car going up} **you** climb up and **you** move down by walking. it's more easy than the opposite (actually). {SS: @@@}

Resolution

and when **you** will move down **you** will see that there are big panels that explain the geological context of the site. ok? and **you** will see that in the eighties **they** make a big survey close to the campus to see where is the bedrock. six hundred meters. ok? so **we** have something four hundred meters- or maybe more- five hundred meters of clay- silt and after- the last meters- it's gravel and so on. ok?

Exemplifying

so for example when **they** built the synchrotron [...] this is typically a building that **you** don't want any motion of the soil because if the electrons at three thousand kilometers go outside the ring **we** can have problems. so **they** made a very important survey concerning the soil and studying everything and **they**- in fact all the building is build on jack [...] to correct everything to be sure that if it moves [...] all the electrons are going in the good direction.

Evaluation

so it's very important to identify everything and to have a first idea of what is the site. [...] ok?

Transcript 21. Story-like example: a) Illustration of similarity of some non-Narrative text strings to the Labovian structure. b) Mapping of personal pronouns 'I', 'you' and 'we'. [20200206_CAM_NAR_10]

Using a finer-grained analysis, it becomes apparent that by using the present or future tenses, the lecturer seems to relate content directly to the students' own knowledge, interests, or environment. Analogy or direct reference are used, rather than hypothesizing or theorizing, for example, the passage "and when you will move down [sic], you will see that..." found in the Resolution section of the transcript. Like for Alsop (2015), in CoFEL sub-plots or sub-stories are sometimes embedded in the larger self-contained Story-like, as in this example. These can even be shorter, 'real' stories in which past tenses are used, such as in the Questioning-Complication, or Exemplifying, stages.

The strong interactional appeal is also evident in how the lecturer addresses the students in a very direct manner. For example, within the 681-word excerpt the lecturer uses the personal pronouns 'you' (20 occurrences) and 'we' (15). The 'you' pronoun is only found in its plural form here. It is much more frequent in the Story-likes than in the previous types.

The ‘you’ set can be divided into two subsets corresponding to two distinct functions of the pronoun: the first one is the generic form, close to the indefinite pronoun ‘one’ (“when you make surveys”); the second serves as a direct address to the student audience, collectively (“do you have an idea of...”). A limited number of the ‘I’ pronoun is found here (6 occurrences), signaling the “presence of the lecturer”, whereas the collective ‘you’ indicates the “presence of the audience” and ‘we’ marks the “presence of both”. All of these are, according to Luukka, “signals of interactional attitudes” in interpersonal metadiscourse (1992, pp. 79–80). That ‘metadiscourse’ describes elements that are used for interactional management and textual organization, in oral or written texts. Ädel (2010, p. 81) reinforces this point: ‘you’ “clearly refers to the audience, and the units are used to direct the audience’s attention and influence their interpretation of the ongoing discourse.” These frequent personal pronouns suggest a strong desire on the part of the lecturers to involve their students in the actual delivery of content, to make them feel that they belong to the community of engineers. The tentative use of humor serves a similar purpose, providing respite from concentrating on otherwise serious content.

5.5.6 Scenarios – The behavior of unsaturated soil & electromagnetic interference

According to Nesi and Alsop (2021), when lecturers have recourse to Scenarios, they choose to describe ‘imagined contexts’. They link the lecture subject matter to imaginary or surprising situations, inviting their audience to visualize scenes or actions. They argue that Scenarios engage the lecturers and their student audiences in the construction of mostly imaginary dialogues, via the use of imperatives, personal pronouns, conditionals and imagined direct speech. Imperatives like ‘imagine’, ‘suppose’, ‘let’s say’ involve the audience, as do first- and second-person pronouns like the non-generic ‘you’ or the inclusive ‘we’. If-clause conditionals, which may co-occur with ‘so’ or ‘due to’, guide listeners through a thought experiment. We emphasize that our acceptance of the term ‘scenario’ is not restricted to ‘fantastic’ events, and is thus broader than that of Nesi and Alsop.

Two samples of Scenarios from CoFEL are presented below. The first (Transcript 22) comes from a Soil Mechanics lecture on ‘soil moisture’—the water content of a soil—and ‘capillary action’—the behavior of a liquid flowing in narrow spaces, also called ‘capillary bridges’. Here the lecturer asks his students to use their imagination and picture the scene, the ‘hypothetical event’: the presence of some soil and the interstitial water it contains. The lecturer

supports his depiction with a drawing on the chalkboard. The last three sentences constitute the necessary coda, “functioning as a commentary on its current relevance” (Nesi & Alsop, 2021, p. 11). In this first example, almost all of the content and morphosyntactic Scenario markers enumerated above are present and have been boxed, for easier reference.

concerning the behavior of unsaturated soil it's something which is **very complex**. **if you imagine that you** have a soil. ok? {lecturer is drawing on chalkboard} this is the level of the soil. at a given depth **you** have groundwater. ok? **so** here **you** have the grains everywhere- ok? er this part of the soil will be said to be saturated. ok? here **we** have a zone of the soil that will be wet but unsaturated. ok? and maybe part of the soil that will be more or less dry **if you** assume that there wasn't any rain. while **you** have these in fact **you** have a kind of transition zone here **due to** the capillarity. [...] the- first specific point is that er when **you** have a soil which is saturated the water can flow and is free to move. when you have drains with a soil that is not saturated then **you** will have some capillarity effect and **you** will have liquid bridges- **let's say**. ok? <FRENCH> des ponts liquides </FRENCH> and this water in fact will be glued to the grain and will not be free to move. ok?

Coda

so this is why when **we** have a soil that is saturated **we** say there is some free water. and when **you** have a soil which is not saturated then **you** have some absorbed water- ok?

Transcript 22. First Scenario example.

[20200206_CAM1_SCEN_10]

The second sample of a Scenario from CoFEL (Transcript 23) comes from a Power Electronics lecture on ‘electromagnetic interference’ (here divergently referred to as ‘diaphony’ after the French term *diaphonie*). The lecturer is questioning his audience about the relevance of ‘shielding’ a circuit by asking them to “imagine” the potential use and effect of a metal shield. Here, the Scenario markers are ‘imagine’, the non-generic ‘you’, and recourse to the epistemic modals ‘may’ and conditional ‘could’, which reinforces the ‘hypothetical’ situation. The necessary coda can be found in the lecturer’s question and its potential resolution by the students.

in general application you have power routes which are generating a high (electric field) and to have signal routes which are sensitive to the capacitive diaphony. so very often (you) find some (cabling) rules which impose to separate power and signal routes- of course. if you cannot separate the routes- so here you have again the first (signal) line which is excited by disturbance voltage and here it's your victim. what about shielding? you could imagine put a metal shield between the two lines because er from maybe your basic (remembrance) of maxwell('s) equations you may imagine that it may be interesting for you to have a kind of shield which would prevent the electrical field er (from) flowing from the line one to line two. [...]

Coda

so the question is this one. is the metal shield effective in the situation here? and you will be asked to answer by yes or no.

Transcript 23. Second Scenario example.

[20201016_SAJ_SCEN_10]

By linking the lecture topic to imaginary or surprising situations, encouraging their audience to visualize scenes or actions, lecturers may be seeking to help learners to engage with the information and retain it (Nesi & Alsop, 2021, p. 24). Scenarios are also used to explain or illustrate technical vocabulary, or to introduce key, complex, or highly abstract concepts of the discipline. Their pedagogical potential is therefore very high. As Malmström et al. (2017, p. 4) remark, highlighting the role of terminology in content courses, “what content teachers teach is in a way the concepts underlying the terminology of the field”. Yet, they also add that the EMI lecturers they observed had “a limited repertoire” of strategies for drawing attention to new terminology (2017, p. 12).

5.6 Storytelling in CoFEL: conclusion and perspectives

This study of storytelling in French EMI lectures shows that some lecturers do explore the realm of storytelling, albeit infrequently, straying occasionally from authoritative facts or established findings in their lecture content. Using recognized story markers, twenty candidate stories were found in CoFEL and eighteen of them were identified as stories. Then, those stories were delineated and categorized by three raters into the six story types retained from the literature, with at least one representative for each type: Narrative, Anecdote, Exemplum, Recount, Story-like and Scenario. This typology could serve as a paradigm for more comprehensive data mining of CoFEL as well as other corpora, which would facilitate the comparison of stories in L1 and EAL lecturing contexts.

The excerpts which were analyzed illustrate how the lecturers strive to involve their audiences, for instance by using interactional language, exploiting direct-address personal pronouns, integrating engaging imperatives, like ‘imagine’ or ‘suppose’, and evoking personal or professional experiences. Additionally, divergent ELF lexical or morphosyntactic forms were found in the discourse of some of the lecturers. While this might abate the effect that a good plot, a captivating topic, or a humorous account can have on listeners, we nonetheless would argue that the use of stories in a lecture does not depend on perfect mastery of the language.

This exploratory study has its limitations, perhaps most notably the fact that switching to massive online teaching because of the 2020 pandemic may have impacted the use of stories in virtual classes, because of the loss of proximity and, therefore, lowered empathy. However, the primary goal has been met, as the results confirm the presence of several types of stories in the CoFEL corpus of EMI lectures. While the multimodal recordings have not yet been fully exploited, future analyses of audio and visual cues could provide more detail about when and how lecturers adapt their attitude, tone, and speech rate to tell a story. A finer analysis of the stories could inform a template for each story type, and this could be used by lecturers and their trainers to raise awareness of story types and to incite reflection on their pedagogical potential. The story templates could be adapted for specific subjects and be incorporated into lecturers’ skill sets. Moreover, the divergent morphosyntax encountered in the story samples also merits analysis, as EMI lecturers and international students may be disoriented by idiomatic, cultural, or idiosyncratic differences. There may also be a potential link between the lecturer’s language level, their degree of experience, and their propensity to tell stories. Finally, the lecturers who participated could be interviewed to explore their awareness of the potential of storytelling. Those who do not use stories could be invited to try, and linguistic or pedagogical assistance could be offered.

Future research could focus on storytelling skills in different contexts. It could be useful to compare similar corpus data from other countries, to explore cultural differences, with either L1 or EAL lecturers. Exploring storytelling at other educational levels and examining disciplinary fields other than engineering could be valuable. In a comparative approach between nomothetic (sciences) and idiographic (humanities) disciplines, Björkman (2008a, p. 120) suggests that, in the former “language is used to report results” whereas the latter “construct results through language”. Moreover, training programs could be developed to support an intentional approach to stories. Together with Ege et al. (2022), who analyzed seven Turkish

lecturers' "strategic discourse behaviors" and their functions, we found that lecturers "encountered many linguistic challenges due to a lack of fluency", especially concerning lexis, and that they could "adopt more conscious discourse strategies". The intentional recourse to storytelling could be one such strategy. Training programs could focus on difficult-to-explain terminology and concepts and consider stories to convey them more effectively. Such programs could be studied in engineering or broader higher education contexts to see whether stories help to connect with an audience and build rapport through emotional and social experiences.

To conclude, in today's world of rapid communication and shared science, engineering lecture content is likely to be quite similar across countries, regardless of the lecturer's origin or country in which they work. If research confirmed the pedagogical usefulness of storytelling in such instructional contexts, we would favor institutional support for professional development, so that lecturers could increase their awareness of the value of storytelling and exploit stories more frequently in their lectures. With some preparation and minimal practice, storytelling could be added to lecturers' skills, which could benefit students as well as lecturers.

6 Discussion

In this chapter, I discuss the relevance of the results of my study of the pragmatic functions of summarizing and storytelling in EMI engineering education in France, and suggest how this research could be extended.

6.1 Main findings

In order to make sense of the information that is delivered to them, students need to understand the pragmatic functions of the various parts of the lecture. EMI lecturers may or may not support learning by clearly signaling discursive and metadiscursive aspects for their audiences; students may learn to interpret the lecturers' cues relative to those aspects (Alsop, 2015). CoFEL analysis has shown that, although some location trends existed for summarizing (subsection 4.4.4), they are mere trends, and no such trends existed for storytelling. Therefore, students cannot anticipate the position of either function in the lecture. A fine-grain quantitative analysis of summarizing (section 4.4) shows that 8.2% of CoFEL duration is dedicated to the cumulated summarizing and storytelling functions, of which summarizing represents 5.4% of the total (104 summaries), and storytelling 2.8% (18 stories). These figures are lower than those of the Engineering Lecture Corpus (ELC, Nesi et al., 2014), as the breakdown of ELC quantitative data shows that the two functions represent 13.6% of the corpus, with 10.1% for summaries and 3.5% for stories. The higher ELC figures may be due to the predominance of L1 courses (UK and New Zealand, around three quarters of courses) over EAL courses (Malaysia).

Furthermore, the relative duration and token length of summarizing episodes according to their type did not differ much overall, but the type and location of summaries were found to be dependent. The beginning part of CoFEL lectures contains relatively more reviews of previous lecture content than of the other types: a large share of the entire summarizing activity (80%) is reviewing content which had been dealt with in past lectures, at the expense of reviewing current content, and previewing current and future content. Thus, in this part, lecturers focus on past sessions rather than current or future ones. This finding is consistent with Alsop's (2015) on the ELC, who found that, although reviews of previous content could be found at any point in the lectures, they tended to cluster in the first third of the lectures.

Concerning the midsection of lectures, 57% of all summarizing activity takes place there, and 78% of it is the reviewing of current content. CoFEL's findings seem consistent with Reece and Walker's (2007, pp. 19–20) prescriptive teacher training guide and their "skills lesson plan" of a successful lecture, with an introduction (beginning), a development (middle), and a conclusion (end). Our findings show a marked difference between such three lecture parts. In contrast, Yaakob (2013) argues that "announcements" (a term akin to CoFEL previews) made in the earlier part of lectures play a more important role than in their middle: students are likely to pay more attention and may take note at the beginning of lectures. In CoFEL, the midsection of lectures contains the majority (57%) of all summarizing activity, and the cumulated previewing (current and future) represents two thirds of midsection summarizing (64%). Clearly, whether for previews or reviews, lecturers favor the summarizing of the current lecture over that of past or future lectures.

Finally, in the ending of lectures, summarizing is the least present of the three parts (14% of all CoFEL summarizing) and mainly consists of reviewing current, and previewing future lecture content (respectively 33% and 40% of all ending summarizing). CoFEL results are thus quite consistent with Alsop's (2015): they show that lecturers tend to concentrate their reviewing of current content towards the end, and tend to conclude with previews of future lecture content.

Selected examples from CoFEL illustrated the pragmatic and linguistic variation in CoFEL summaries (section 4.5), according to their location in the lectures. Summaries, whether short or long, could play a lecture scaffolding role, from simple links between two parts of the lecture, to detailed enumerations of past or forthcoming lecture outlines. Learning is reinforced with reviews, when information is repackaged, whereas it is scaffolded by previews predicting new information (Alsop, 2015). Young (1994), referring to the metadiscursive "Discourse Structuring" function of indications and announcements of a new direction to be taken in the lecture, claims that the prediction of content enhances understanding. Young found that such announcements were roughly equivalent to the number of points a lecturer makes in a session. The scaffolding role is notably played by previews, as predicting new information scaffolds learning, both in the lecture hall or when reviewing lecture notes. The role is particularly salient with previews of future content, which, in terms of content processing, have a lecture structuring function at a higher level of temporality (Alsop, 2015). Thanks to previews, topics are put into the context of a framework for future lectures.

The phenomenon of summary chains was also investigated (section 4.5.5): for about a third of summary instances, lecturers combine two or more types, usually at strategic points in their lectures, for instance when closing a topic to begin a new one. The phenomenon can happen anywhere in the lectures, without any pattern of regularity, apart from their total absence from the ending part of lectures. The co-occurrence of summaries was highlighted by Alsop (2015), who also noted a tendency for chaining to occur early in the lectures, which could support the idea that the effect of summarizing can be considered greater at the beginning of lectures. Like summaries in general, and given their essential structuring value, the command of summary chains could form part of professional development programs for lecturers.

Concerning storytelling, the main finding was that engineering lecturers did tell stories to their international student audiences in the CoFEL lectures. Eighteen stories were identified and categorized into six story types, illustrated by representative examples. CoFEL stories are rather infrequent and underdeveloped. Nevertheless, the presence of these stories can be used as an argument for modifying teacher development, and promoting a more systematic use of storytelling, given its pedagogical potential.

With regard to ELF usage and EAL divergence, disfluencies, divergent pronunciations, or non-standard morphosyntactic forms were found in all the lecture recordings and transcripts, for both summaries and stories. Taken individually, these elements may seldom pose a problem for students, but when combined, they could lead to misunderstanding or incomprehension (Henderson, 2021). Thus, the effective use of summaries and stories in a lecture may not depend on a perfect mastery of the language, but rather on the lecturers' apprehension of the linguistic and pragmatic issues at stake in that specific communication situation.

6.2 Corpus patterns, relationships, and features

The summarizing pragmatic function is present in all lectures, and plays a significant role in lecture discourse; it occurs four times per hour and occupies almost six per cent of the lectures. Much variation appeared in the amount of summarizing in the lectures, from lecturer to lecturer, for example in the number of occurrences, or the duration of summaries. This variation can depend on the lecturers' idiosyncrasies, pedagogical approaches, or linguistic abilities, and sometimes it can be challenging for students. Additionally, a third of the summaries are involved in seamless chains of two summaries or more. Systematic co-

occurrences of summary types could not be identified, but these chains create relationships within the current lecture, and also across the previous, current and future lectures. Similarly to the ELC (Alsop, 2015), key lecture sections and their articulation may be highlighted in CoFEL, when lecturers form chains of related summaries to ensure that students are aware of the most important connections and junctures in the lecture, often with recapitulated key points informing the new content.

Storytelling patterns in the data showed that stories are much less frequent than summaries in CoFEL, with only 18 occurrences compared to 104 summaries, and stories are not found in all the lectures. Their main purpose is to illustrate and exemplify, not to develop new engineering concepts or to introduce new technical course content. CoFEL stories usually showcase experiences, either with the lecturers' positioning themselves as experts, or that of others, often to distance themselves from nonexperts, which is consistent with Dyer and Keller-Cohen findings (2000). As for Alsop (2015) in the ELC, this is done across the full range of story types found in CoFEL, with differing purposes. A narrative will explore real-life engineering situations in detail, outlining the problem after setting the scene, and giving its resolution and consequences. An anecdote will briefly describe an exceptional situation or event, usually with a touch of humor. An exemplum will insist on the moral aspect of a serious accident and the lessons learned from it, while recounts will simply describe a series of facts, usually revolving around famous figures in engineering or science and their theories or discoveries.

Lastly, person deixis is a salient feature of both summaries and stories, and plays a key role. The frequent use of the personal pronouns *you* and *we* in both pragmatic functions implies an interactional appeal on the part of the lecturers, as they address their students in a very direct manner, in an interpersonal, intersubjective metadiscourse (Ädel, 2010, 2012; Luukka, 1992). Through these interpersonal moves, the lecturers seek to build rapport with their students and engage them in a collective pedagogical effort. Yeo and Ting (2014) argue that personal pronouns are indicators of the engagement of students in lecture settings, as well as indicators of interactivity. They also found that personal pronouns are used frequently, notably at the beginning of lectures, when activating prior knowledge and giving instructions or announcements, which amounts to what we have called summarizing lecture content.

6.3 Implications

EMI and CLIL are different and this had important ramifications. EMI focuses solely on the transmission of content, through the medium of English as an additional language, but without any explicit intentional consideration of this language (Brown & Bradford, 2017). In contrast, CLIL seeks to combine the acquisition of content and language, with the added value of learning linguistic and communication skills applied to that content (Frigols-Martin et al., 2011). In the graduate school of engineering where this study was set, lecturing at Master's level is entirely EMI. The results of this study could inspire research or action research, by raising awareness that moving toward CLIL (i.e., the integration of language and content) would be beneficial. For example, researchers could build on the existing summarizing and storytelling of a given lecturer as a starting point to refine the lecturer's pragmatic and linguistic aspects of either or both functions, and study the impact on both lecturer and students. This could be done in language-and-content interdisciplinary approaches, or through team-teaching involving language specialists (Lasagabaster, 2018, 2022a, 2022b), and could represent a first step towards CLIL. The content-and-language challenge could also be partly met by preparing and using learning materials that have been adapted with language considerations, such as handouts or slideshows (Wilang et al., 2024).

The approach used to build and use CoFEL draws on previous research, notably Coventry University's Engineering Lecture Corpus (ELC) and their own method in using the pragmatic annotation of lectures (Alsop & Nesi, 2014; Alsop, 2015). However, despite their similarities, not the least of which is the fact that they both involve solely engineering lectures, there are significant methodological differences. For example, for CoFEL, unlike ELC, all transcription and annotation of audiovisual files was carried out on the ELAN platform. As a result, all annotations are time-aligned, enabling instant searches of strings of text in context, or the use of time codes and duration figures, whereas the ELC only provides researchers with token counts and the token length of text strings. While counting in tokens (rather than seconds, minutes or hours) can be relevant to researchers, it remains a rather abstract approach in more general communications and contexts, such as negotiations with institutional representatives.

Lecturers who are field specialists could benefit from this study of language use, as the results presented in this thesis give an insight into what happens in the lecture hall. Explicit macro-structuring features were found in lectures, as shown by the analysis of summarizing, and these play a key role in facilitating understanding. The presence of summarizing in CoFEL lectures confirms that lecturers scaffold their lectures, whether intentionally or not. Instances of storytelling were also found, and this is important because stories can be a powerful

pedagogical aid in conveying real-life situations in lectures, while contributing actively to the memorizing process. Of course, the results also reveal numerous imperfections in the rendering of the pragmatic functions, notably through the ELF-EAL prism, and possibly also through a lack of awareness of the key role of such functions. Nonetheless, the results show that building blocks are in place, on which a more systematic and well-thought-out use of these functions could be constructed, for example by raising lecturers' awareness of their fundamental roles in lectures. This could be done as part of their professional development, to prepare them to teach in the specific EMI context.

As with all types of divergence, an ELF perspective can make 'errors' acceptable, yet the major question is whether a given 'error' disturbs the communication situation (Björkman, 2008a, 2008b). More importantly, Björkman also suggests that, because engineering is primarily a nomothetic⁵² discipline, like economics or linguistics, in which language reports results (contrary to idiographic disciplines like philosophy or literature), it may reduce disturbance in communication. However, I would argue that engineering lectures are not solely about reporting procedures and results, and that they combine nomothetic and idiographic orientations. Pragmatic functions such as summarizing and storytelling may serve as a bridge between types of knowledge, thus reflecting the lecturers' desire to make their high-stakes content accessible. Analyzing those pragmatic functions recognizes their utility, as a locus for observing how ELF divergences can coexist with good, or at least well-intentioned, lecturing practices. Nativelike English is not obligatory to lecture well in the EMI-ELF context. Without downplaying the role of grammar, raising lecturers' awareness of key concepts such as the lexical or phonological accuracy of their specialist terminology would be a huge step forward in the intelligibility of engineering courses.

The findings of this thesis could help convince higher education institutions and policymakers that they have a key role to play in defining and implementing more interdisciplinary practices, particularly by bringing together content and language specialists. Greater integration of content and language could help students to become more competent communicators. Furthermore, lecturer preparation sessions to teach EMI classes seem to be either insufficient, or underused, perhaps due simply to a sheer lack of time in the lecturers' timetables. Yet, professional language development would help lecturers to avoid fossilization of their skills (Han & Selinker, 2005; Richter, 2019). In response to Airey's call for the

⁵² Cf., Windelband (1998, 1894 translated reprint), German philosopher (1848-1915) mainly remembered for introducing the terms *nomothetic* and *idiographic*.

development and implementation of disciplinary literacy (2011, 2020), I posit that lecturers, whose role is to create disciplinary literate graduates, must themselves be disciplinary literate in the additional language. Most lecturer training programs seem to remain general-purpose, whereas lecturers, as expert models, need to be appropriately involved in the communication practices of their discipline. Allowing more space for professional development, as well as encouraging team teaching and interdisciplinary approaches, could have a strong impact on the quality of EMI lectures. Summarizing and storytelling are good starting points.

6.4 Limitations of the study

There are noticeable limitations to this study. Although the panel of lecturers was fairly representative of the faculty in terms of background, age and gender, their number (11 lecturers) remains quite low, which limits generalizing. The same applies to the number of lectures (14) and the number of institutions where the study is set (1). While some EMI studies are based on a handful of participants, this thesis based on CoFEL is not comparable to studies exploiting larger corpora of academic lectures, such as the British Academic Spoken English (BASE)⁵³ corpus, or the Michigan Corpus of Academic Spoken English (MICASE)⁵⁴. This limitation in corpus size has direct implications on methodological choices, as a larger corpus would justify statistical analysis and provide richer qualitative data. Moreover, follow-up interviews would allow tentative causal relationships to be posited; for now, lecturers' intentions and students' experience of EMI lecturing remain unknown.

6.5 Further research and conclusion

Thanks to the observation and analysis of authentic lecturing situations from CoFEL, this study has provided some insight into French EMI engineering lectures, revealing the presence of summarizing and storytelling. This is in line with research on ELC, where the lecturers were both L1-English and EAL speakers (Alsop & Nesi, 2014; Alsop, 2015). A comparative analysis of the two corpora could provide a further avenue for research, with a few provisos. First, the amounts of summarizing and storytelling were found to be less important in CoFEL than in the ELC, which could be due to the fact that CoFEL is comprised solely of EAL lectures. Furthermore, CoFEL's fourteen lectures make it a smaller corpus than the ELC's

⁵³ https://www.reading.ac.uk/acadepts/ll/base_corpus/ [Last accessed June 4th, 2024]

⁵⁴ <https://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=simple> [Last accessed June 4th, 2024]

seventy-six. In further research, it could be of interest to compare CoFEL with the ELC's Malaysian subcorpus, a similar-sized set of engineering lectures in EAL. Since three quarters of ELC lectures are delivered in the lecturers' L1, any comparative pragmatic-linguistic analysis with CoFEL's EAL lectures runs the risk of being distorted. Comparisons would also need to consider the very different amounts of summarizing activity (in terms of the number of tokens), summarizing represents about ten per cent of all lecture discourse in the ELC, compared to a little less than seven per cent in CoFEL, and for storytelling, the ELC has 153 instances of stories (about two per lecture on average), as opposed to 18 instances in CoFEL (a little over one story per lecture).

Additionally, researchers could extend this research by exploiting the typologies and approaches to summarizing and storytelling. The original raw corpus created for this research remains largely unexploited, as CoFEL contains less than half of the total number of lectures recorded. CoFEL was integrally transcribed and annotated with a series of pragmatic and discursive features⁵⁵. For example, the Lecturing pragmatic function, which identifies all those parts of the lecture where the lecturers deliver their course content, represents a vast pool of untapped data.

It is hoped that this thesis will inspire future works in France and on French EMI. These could be done on the same pragmatic functions, but on a larger scale, preferably involving more than one researcher in one institution. Other pragmatic functions could be explored in the same vein, like the use of humor in lectures, as has already been done in some countries (Alsop, 2015, 2016; Berk, 1996; Ladilova & Schröder, 2022; Nesi, 2012; Pickering et al., 2009; Tzoannopoulou, 2016). Few instances of humor, or attempts at it, were found in CoFEL, unlike ELC, perhaps confirming that using humor in another language is one of the most difficult things to do.

CoFEL could be further exploited for a multitude of other types of surveys, notably by drawing on its multimedia dimension (video, sound, and full text transcriptions). This could open up new avenues of research into co-verbal gestures or phonological aspects, for example. As the corpus comprises less than half of all the multimedia recordings originally made, there is still much work to be done to exploit the material entirely. The lectures are all supported by slideshows, thus one could examine their content, the lecture-slide relationships, and the impact they may have on the lectures (González-Mujico & Lasagabaster, 2024).

⁵⁵ See Figure 6 (p. 50).

Enriching and triangulating the results with ancillary data, such as interviews with lecturers and students about their perceptions of the courses, would contribute to more robust interpretations. Comparative studies in other disciplinary subjects could shed light on the presence of specificities in engineering lecture discourse, as well as in other subjects. EMI engineering lectures around the world on a given subject may have much in common in terms of vocabulary or course content, yet investigating that which is context- and culture-specific could inform the internationalization of HE programs. Exploring the role played by code-switching in ELF lectures, as well as lecturer-student interaction, are also possible avenues for research.

Arguing that the use of English as a medium of instruction includes a wide range of curricular activities that constitute education, some researchers have suggested replacing EMI by English-Medium Education (EME), in a more positive and extensive vision of the concept, thus embracing both instruction and learning (Dafouz-Milne & Smit, 2016, 2020). While developing an operational framework to conduct planning, implementation, and assessment of EME, they furthered EME into English-Medium Education in Multilingual University Settings, or EMEMUS. Their Road-Mapping dynamic framework has become a useful research tool to analyze and compare EME instances, settings and practices within six comprehensive dimensions linked by discourse: the roles of English in relation to other languages; the academic disciplines taught; the management of language; the agents involved; the practices and processes; internationalization and glocalization (Dafouz-Milne & Smit, 2023). A promising spin-off of such moves in the past decade has been the successful rooting and growth of the concept of Integration of Content and Language in Higher Education (ICLHE), which could be roughly defined as a move towards the implementation of CLIL in HE (Dafouz-Milne, 2020; Dimova & Kling, 2020; Studer, 2018; Studer & Smit, 2021; Valcke & Wilkinson, 2017; Wilkinson, 2018; Wilkinson & Walsh, 2014; Wilkinson & Zegers, 2008). Such an approach could be envisaged and would benefit French higher education institutions.

References

- Abrahamson, C. E. (1998). Storytelling as a pedagogical tool in higher education. *Education*, 118(3), 440–452.
- Adedokun, F., & Tourbeaux, J. (2022). *Les personnels enseignants de l'enseignement supérieur du ministère de l'Enseignement supérieur et de la Recherche—Année 2021*. Ministère de l'enseignement supérieur et de la recherche – Direction générale des ressources humaines – Service des personnels enseignants de l'enseignement supérieur et de la recherche. <https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2022-11/note-dgrh-n-9-octobre-2022-ann-e-2021-25228.pdf>
- Ädel, A. (2010). Just to Give You Kind of a Map of Where We Are Going: A Taxonomy of Metadiscourse in Spoken and Written Academic English. *Nordic Journal of English Studies*, 9, 69–97. <https://doi.org/10.35360/njes.218>
- Ädel, A. (2012). “What I want you to remember is...”: Audience orientation in monologic academic discourse. *English Text Construction*, 5(1), 101–127. <https://doi.org/10.1075/etc.5.1.06ade>
- Airey, J. (2011). The disciplinary literacy discussion matrix: A heuristic tool for initiating collaboration in higher education. *Across the Disciplines*, 8(3), 1–9. <https://doi.org/10.37514/ATD-J.2011.8.3.18>
- Airey, J. (2012). I Don't Teach Language. The Linguistic Attitudes of Physics Lecturers in Sweden. *AILA Review*, 25, 64–79. <https://doi.org/10.1075/aila.25.05air>
- Airey, J. (2016). EAP, EMI or CLIL? In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 71–83). Routledge.
- 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., Lauridsen, K. M., Räsänen, A., Salö, 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), 561–576.
- Allen, W. (2017). Making corpus data visible: Visualising text with research intermediaries. *Corpora*, 12(3), 459–482. <https://doi.org/10.3366/cor.2017.0128>

- Allison, D., & Tauroza, S. (1995). The effect of discourse organisation on lecture comprehension. *English for Specific Purposes*, 14(2), 157–173. [https://doi.org/10.1016/0889-4906\(95\)00007-E](https://doi.org/10.1016/0889-4906(95)00007-E)
- Almeida, J. (2018). *Internationalisation at Home: An epistemology of equity*. European Association for International Education (EAIE). <https://www.eaie.org/>
- Almeida, J., Robson, S., Morosini, M., & Baranzeli, C. (2019). Understanding Internationalization at Home: Perspectives from the Global North and South. *European Educational Research Journal*, 18(2), 200–217. <https://doi.org/10.1177/1474904118807537>
- Alsop, S. (2015). *The Pragmatic Annotation of Academic Lectures: An Exploratory Study Applied to Engineering Disciplines* [Coventry University, UK]. https://pure.coventry.ac.uk/ws/portalfiles/portal/31310787/Alsop_Pure.pdf
- Alsop, S. (2016). The ‘Humour’ Element in Engineering Lectures Across Cultures: An Approach to Pragmatic Annotation. In M. J. López-Couso, B. Méndez-Naya, P. Núñez-Pertejo, & I. M. Palacios-Martínez (Eds.), *Corpus linguistics on the move*. Brill - Rodopi. https://doi.org/10.1163/9789004321342_016
- Alsop, S., Moreton, E., & Nesi, H. (2013). The uses of storytelling in university Engineering lectures. *ESP Across Cultures*, 10, 8–19.
- Alsop, S., & Nesi, H. (2014). *The Pragmatic Annotation of a Corpus of Academic Lectures*. http://www.lrec-conf.org/proceedings/lrec2014/pdf/323_Paper.pdf
- Alterio, M., & McDrury, J. (2003). *Learning Through Storytelling in Higher Education: Using Reflection and Experience to Improve Learning*. Routledge. <https://doi.org/10.4324/9780203416655>
- Andrews, D., Hull, T., & Donahue, J. (2009). Storytelling as an Instructional Method: Definitions and Research Questions. *Interdisciplinary Journal of Problem-Based Learning*, 3(2), 6–23. <https://doi.org/10.7771/1541-5015.1063>
- Archibald, A., Cogo, A., & Jenkins, J. (2011). *Latest Trends in ELF Research*. Cambridge Scholars Publishing.
- Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 51(5), 267–272. <https://doi.org/10.1037/h0046669>
- Baltaru, R.-D., Manac, R.-D., & Ivan, M.-D. (2022). Do rankings affect universities’ financial sustainability? – Financial vulnerability to rankings and elite status as a positional good.

- Studies in Higher Education*, 47(11), 2323–2335.
<https://doi.org/10.1080/03075079.2022.2061447>
- Baratta, A. (2019). *World Englishes in English Language Teaching*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-13286-6>
- Beacco, J.-C. (2019). L'internationalisation des formations d'enseignement supérieur: La question des langues, de politique en didactique. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliu*, 38(2), Article Vol. 38 N°2. <https://doi.org/10.4000/apliut.7129>
- Beacco, J.-C., Bertrand, O., Herreras, J. C., & Tremblay, C. (Eds.). (2022). *La gouvernance linguistique des universités et établissements d'enseignement supérieur*. Palaiseau: Éditions de l'École Polytechnique.
- Beelen, J., & Jones, E. (2015). Europe Calling: A New Definition for Internationalization at Home. *International Higher Education*, 83, 12–13. <https://doi.org/10.6017/ihe.2015.83.9080>
- Benito, M., Gil, P., & Romera, R. (2019). Funding, is it key for standing out in the university rankings? *Scientometrics*, 121(2), 771–792. <https://doi.org/10.1007/s11192-019-03202-z>
- Berk, R. (1996). Student Ratings of 10 Strategies for Using Humor in College Teaching. *Journal on Excellence in College Teaching*, 7.
- Bernstein, B. (1999). Vertical and horizontal discourse: An essay. *British Journal of Sociology of Education*, 20(2), 157–173. <https://doi.org/10.1080/01425699995380>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Birch, P., Baïdak, N., De Coster, I., & Kocanova, D. (2023). *Key data on teaching languages at school in Europe: 2023 edition*. European Commission, European Education and Culture Executive Agency. Publications Office of the European Union. <https://data.europa.eu/doi/10.2797/529032>
- Björkman, B. (2008a). English as the Lingua Franca of Engineering: The Morphosyntax of Academic Speech Events. *Nordic Journal of English Studies*, 7(3), Article 3.
- Björkman, B. (2008b). 'So where we are?' Spoken lingua franca English at a technical university in Sweden. *English Today*, 24(2), 35–41. <https://doi.org/10.1017/S0266078408000187>

- Björkman, B. (2013). English as an Academic Lingua Franca: An Investigation of Form and Communicative Effectiveness. In *English as an Academic Lingua Franca*. De Gruyter Mouton. <https://doi.org/10.1515/9783110279542>
- Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge University Press.
- Blattès, M. (2018a). *English-Medium Instruction in French Higher Education: Competing Ideologies, Policies and Practices* [PhD Thesis, King's College London]. <https://kclpure.kcl.ac.uk/portal/en/studentTheses/english-medium-instruction-in-french-higher-education>
- Blattès, M. (2018b). Policy development for English-medium instruction in French universities. *European Journal of Language Policy*, 10(1), 13–38. <https://doi.org/10.3828/ejlp.2018.2>
- Block, D., & Khan, S. (2021). *The Secret Life of English-Medium Instruction in Higher Education: Examining Microphenomena in Context*. Routledge.
- Block, D., & Moncada-Comas, B. (2019). English-medium instruction in higher education and the ELT gaze: STEM lecturers' self-positioning as NOT English language teachers. *International Journal of Bilingual Education and Bilingualism*, 25(2), 401–417. <https://doi.org/10.1080/13670050.2019.1689917>
- Blum-Kulka, S., & Olshtain, E. (1986). Too Many Words: Length of Utterance and Pragmatic Failure. *Studies in Second Language Acquisition*, 8(2), 165–179. <https://doi.org/10.1017/S0272263100006069>
- Borri, C., Guberti, E., & Melsa, J. (2007). International dimension in engineering education. *European Journal of Engineering Education*, 32(6), 627–637. <https://doi.org/10.1080/03043790701520586>
- Bowles, H., & Murphy, A. C. (Eds.). (2020). *English-medium instruction and the internationalization of universities*. Palgrave Macmillan.
- Breetvelt, I. (2018). English-medium instruction in Dutch higher education: A policy reconstruction and impact study. *L1-Educational Studies in Language and Literature*, 1–24. <https://doi.org/10.17239/L1ESLL-2018.18.03.10>
- Briggs Baffoe-Djan, J., & Smith, S. (2017). English Medium Instruction and Idiomaticity in English as a Lingua Franca. *Iranian Journal of Language Teaching Research*, 5(3), 27–44.

- Brown, H., & Bradford, A. (2017). *EMI, CLIL, & CBI: Differing Approaches and Goals*.
https://www.academia.edu/34155098/EMI_CLIL_and_CBI_Differing_Approaches_and_Goals
- Bruner, J. S. (1990). *Acts of Meaning: Four Lectures on Mind and Culture* (Kindle e-edition).
 Harvard University Press.
- Cagliari, R., & Jenkins, J. (Eds.). (2011). *Discourses, Communities, and Global Englishes*.
<https://www.peterlang.com/document/1051903>
- Cambridge English. (2021). *Certificate in EMI Skills course*. Cambridge University Press &
 Assessment [Accessed April 11th, 2024]. <https://www.cambridgeenglish.org/teaching-english/teaching-qualifications/institutions/certificate-in-emi-skills/>
- Cambridge English. (2022). *Linguaskill test*. Cambridge University Press & Assessment.
<https://www.cambridgeenglish.org/exams-and-tests/linguaskill/>
- Campus France. (2022). *In France, you can also study in English!* EPIC Campus France,
 l'Agence Française Pour La Promotion de l'enseignement Supérieur, l'accueil et La
 Mobilité Internationale [Accessed April 10, 2024].
<https://www.campusfrance.org/en/in-france-you-can-also-study-in-english>
- Campus France. (2023). *Le coût des études supérieures en France*. EPIC Campus France,
 l'Agence française pour la promotion de l'enseignement supérieur, l'accueil et la
 mobilité internationale [Accessed April 10, 2024].
<https://www.campusfrance.org/fr/cout-etudes-superieures-france-frais-inscription>
- Campus France. (2024a). *Chiffres Clés 2024 de la mobilité étudiante*. EPIC Campus France,
 l'Agence française pour la promotion de l'enseignement supérieur, l'accueil et la
 mobilité internationale [Accessed April 10, 2024].
<https://www.campusfrance.org/fr/actu/chiffres-cles-2024-de-la-mobilite-etudiante-plus-de-412-000-etudiants-internationaux-en-france>
- Campus France. (2024b, April 10). *Programs Taught in English catalog—VI.4.7*. EPIC
 Campus France, l'Agence Française Pour La Promotion de l'enseignement Supérieur,
 l'accueil et La Mobilité Internationale [Accessed April 10, 2024].
<https://taughtie.campusfrance.org/tiesearch/#/catalog>
- Card, S.-K., Mackinlay, J.-D., & Shneiderman, B. (1999). *Readings in Information
 Visualization: Using Vision to Think*. Morgan Kaufmann Publishers In.
- Carrell, P. L. (1987). Content and Formal Schemata in ESL Reading. *TESOL Quarterly*, 21(3),
 461–481. <https://doi.org/10.2307/3586498>

- CGE. (n.d.). Commission Relations internationales—Conférence des Grandes Écoles. *Conférence des Grandes Écoles*. Retrieved May 16, 2024, from <https://www.cge.asso.fr/commissions/relations-internationales/>
- Chafe, W. L. (1985). Some reasons for hesitating. In D. Tannen & M. Saville-Troike (Eds.), *Temporal Variables in Speech* (pp. 169–180). Ablex Pub. Corp. <https://doi.org/10.1515/9783110816570.169>
- Chanquoy, L., Tricot, A., & Sweller, J. (2007). La théorie de la charge cognitive. In *La charge cognitive: Théorie et applications* (pp. 131–188). Armand Colin.
- Chapple, J. (2015). Teaching in English Is Not Necessarily the Teaching of English. *International Education Studies*, 8(3), 1–13.
- Chaudron, C., & Richards, J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7(2), 113–127. <https://doi.org/10.1093/applin/7.2.113>
- CIP-KU. (2009). *English language certification: TOEPAS*. Centre for Internationalisation and Parallel Language Use (CIP); University of Copenhagen. <https://cip.ku.dk/english/certification/>
- Coleman, J. A. (2006). English-medium teaching in European Higher Education. *Language Teaching*, 39(1), Article 1.
- Collins, P. C. (1987). Cleft and pseudo-cleft constructions in english spoken and written discourse. In G. Sampson & D. McCarthy (Eds.), *Corpus Linguistics: Readings in a Widening Discipline (2004)* (pp. 85–94). Continuum International Publishing Group.
- Connor, U., & Mauranen, A. (1999). Linguistic Analysis of Grant Proposals: European Union Research Grants. *English for Specific Purposes*, 18(1), 47–62. [https://doi.org/10.1016/S0889-4906\(97\)00026-4](https://doi.org/10.1016/S0889-4906(97)00026-4)
- Connor, U., & Upton, T. A. (2004). *Discourse in the Professions: Perspectives from corpus linguistics*. John Benjamins Publishing Co.
- Cosnefroy, L., Ketele, J.-M. D., Hugonnier, B., Palomba, D., Parmentier, P., & Uvalic-Trumbic, S. (2020). *L'internationalisation de l'enseignement supérieur: Le meilleur des mondes ?* De Boeck SUP.
- Costa, F., & Mastellotto, L. (2022). The role of English for specific purposes (ESP) in supporting the linguistic dimension in English-medium instruction (EMI). *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 5(2), Article 2. <https://doi.org/10.5565/rev/clil.91>

- Council of Europe. (2020). *Common European Framework of Reference for Languages (CEFR)* [Institutional]. <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>
- Council of Europe. (2024). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Common European Framework of Reference for Languages (CEFR). <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Crawford Camiciottoli, B. (2004). Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *Journal of English for Academic Purposes*, 3(1), 39–54. [https://doi.org/10.1016/S1475-1585\(03\)00044-4](https://doi.org/10.1016/S1475-1585(03)00044-4)
- Crystal, D. (2003). *English as a Global Language* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486999>
- Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics* (6th ed.). Wiley-Blackwell. <https://www.wiley.com/en-br/A+Dictionary+of+Linguistics+and+Phonetics%2C+6th+Edition-p-9781405152969>
- Crystal, D. (2018). *The Cambridge Encyclopedia of the English Language* (3rd ed.). Cambridge University Press. <https://www.cambridge.org/highereducation/books/the-cambridge-encyclopedia-of-the-english-language/2B10AC8766B73D09955C899572C1E7EC#overview>
- Csomay, E. (2007). A corpus-based look at linguistic variation in classroom interaction: Teacher talk versus student talk in American University classes. *Journal of English for Academic Purposes*, 6(4), 336–355. <https://doi.org/10.1016/j.jeap.2007.09.004>
- CTI. (n.d.). *UE & Processus de Bologne – CTI – Commission des Titres d’Ingénieur*. Retrieved May 16, 2024, from <https://www.cti-commission.fr/en/international/eees/ue-processus-de-bologne>
- CTI. (2006). *Références et Orientations 2006*. Commission des Titres d’Ingénieur. <https://www.cti-commission.fr/references-et-orientations>
- Curaj, A., Deca, L., & Pricopie, R. (Eds.). (2020). *European Higher Education Area: Challenges for a New Decade*. Springer Nature Switzerland AG. <https://link.springer.com/book/10.1007/978-3-030-56316-5>
- Dafouz-Milne, E. (2020). Postscript: Moving Forward in Integrating Content and Language in Multilingual Higher Education. In S. Dimova & J. Kling (Eds.), *Integrating Content and Language in Multilingual Universities* (pp. 179–183). Springer International Publishing. https://doi.org/10.1007/978-3-030-46947-4_11

- Dafouz-Milne, E., & Smit, U. (2016). Towards a dynamic conceptual framework for English-medium education in multilingual university settings. *Journal of Applied Linguistics*, 37(3), 397–415. <https://doi.org/10.1093/applin/amu034>
- Dafouz-Milne, E., & Smit, U. (2020). *ROAD-MAPPING English Medium Education in the Internationalised University*. Palgrave Pivot.
- Dafouz-Milne, E., & Smit, U. (Eds.). (2023). *Researching English-Medium Higher Education: Diverse Applications and Critical Evaluations of the ROAD-MAPPING Framework*. Routledge. <https://doi.org/10.4324/9781003193852>
- Dahlstrom, M. (2014). Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences of the United States of America*, 111. <https://doi.org/10.1073/pnas.1320645111>
- Dakhli, J. (2008). *Lingua franca: Histoire d'une langue métisse en Méditerranée*. Actes Sud.
- Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). *Language Use and Language Learning in CLIL Classrooms*. John Benjamins Publishing.
- Dalton-Puffer, C., & Smit, U. (2013). Content and Language Integrated Learning: A research agenda. *Language Teaching*, 46(4), 545–559. <https://doi.org/10.1017/S0261444813000256>
- Davies, A. (2003). *The Native Speaker: Myth and Reality*. <https://www.multilingual-matters.com/page/detail/The-Native-Speaker/?k=9781853596223>
- Davies, A. (2013). Is the native speaker dead? *Histoire Épistémologie Langage*, 35(2), 17–28. <https://doi.org/10.3406/hel.2013.3455>
- de Ketele, J.-M., Hugonnier, B., Parmentier, P., & Cosnefroy, L. (2016). *Quelle excellence pour l'enseignement supérieur ?* De Boeck Supérieur. <https://unr.ra.scholarvox.com/book/88841262>
- de Wit, H., Hunter, F., Howard, L., & Egron-Polak, E. (2015). *Internationalisation of higher education*.
- Dearden, J. (2014). *English as a medium of instruction—A growing global phenomenon*. British Council. <https://ora.ox.ac.uk/objects/uuid:4f72cdf8-b2eb-4d41-a785-4a283bf6caaa>
- Deneire, M., & Benmokhtar, H. (2024). English-medium instruction in higher education in France. In *The Routledge Handbook of English-Medium Instruction in Higher Education*. Routledge.
- Deroey, K. L. B. (2012). What they highlight is...: The discourse functions of basic wh-clefts in lectures. *Journal of English for Academic Purposes*, 11(2), 112–124. <https://doi.org/10.1016/j.jeap.2011.10.002>

- Deroey, K. L. B. (2013). *Relevance marking in lectures: A corpus-based study* [Universiteit Gent].
https://www.academia.edu/3821030/Relevance_marking_in_lectures_a_corpus_based_study
- Deroey, K. L. B., & Taverniers, M. (2012). Just remember this: Lexicogrammatical relevance markers in lectures. *English for Specific Purposes*, 31(4), 221–233.
<https://doi.org/10.1016/j.esp.2012.05.001>
- Derwing, T. M., & Munro, M. J. (2009a). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review*, 66(2), 181–202.
- Derwing, T. M., & Munro, M. J. (2009b). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476–490.
- Dewey, M. J., & Patsko, L. (2017). ELF and teacher education. In *The Routledge Handbook of English as a Lingua Franca*. Routledge.
- Dimova, S. (2017). Life after oral English certification: The consequences of the Test of Oral English Proficiency for Academic Staff for EMI lecturers. *English for Specific Purposes*, 46, 45–58. <https://doi.org/10.1016/j.esp.2016.12.004>
- Dimova, S. (2021). Certifying lecturers' English language skills for teaching in English-medium instruction programs in higher education. *ASp. La Revue Du GERAS*, 79, Article 79. <https://doi.org/10.4000/asp.7056>
- Dimova, S., Hultgren, A. K., & Jensen, C. (2015). *English-Medium Instruction in European Higher Education* (Vol. 3). De Gruyter. www.degruyter.com
- Dimova, S., & Kling, J. (2018). Assessing English-Medium Instruction Lecturer Language Proficiency Across Disciplines. *TESOL Quarterly*, 52(3), 634–656.
<https://doi.org/10.1002/tesq.454>
- Dimova, S., & Kling, J. (Eds.). (2020). *Integrating Content and Language in Multilingual Universities* (Springer, Vol. 44). <https://link-springer-com.gaelnomade-1.grenet.fr/book/10.1007%2F978-3-030-46947-4>
- Doiz, A., Lasagabaster, D., & Pavón, V. (2019). The integration of language and content in English-medium instruction courses: Lecturers' beliefs and practices. *Iberica -Journal of the European Association of Languages for Specific Purposes*.
https://www.academia.edu/41661258/The_integration_of_language_and_content_in_English_medium_instruction_courses_Lecturers_beliefs_and_practices

- Dubois, S., & Sankoff, D. (2005). The Variationist Approach Toward Discourse Structural Effects and Socio-Interactional Dynamics. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The Handbook of Discourse Analysis* (1st ed., pp. 282–303). Wiley. <https://doi.org/10.1002/9780470753460.ch16>
- Dyer, J., & Keller-Cohen, D. (2000). The Discursive Construction of Professional Self Through Narratives of Personal Experience. *Discourse Studies*, 2(3), 283–304. <https://doi.org/10.1177/1461445600002003002>
- ELAN. (2023). *Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive* (Version 5.7 to 6.7) [Computer software]. <https://archive.mpi.nl/tla/elan>
- EMINSA. (2024). *Développer des enseignements scientifiques en anglais* (fr) [Text]. INSA Toulouse. <https://c2ip.insa-toulouse.fr/fr/pedagogies/eminsa.html>
- EOP. (2024). *European Observatory for Plurilingualism*. Last Accessed April 15, 2024. <https://www.observatoireplurilinguisme.eu/en-us/>
- Eslami, Z. R., & Eslami-Rasekh, A. (2007). Discourse markers in academic lectures. *Asian EFL Journal*, 9(1), 22–38.
- Ethnologue. (2023). *What is the most spoken language?* Ethnologue. <https://www.ethnologue.com/insights/most-spoken-language/>
- European Commission. (1999). *The Bologna Process and the European Higher Education Area* [Text]. Education and Training - European Commission. https://ec.europa.eu/education/policies/higher-education/bologna-process-and-european-higher-education-area_en
- European University Association. (2023, October 17). *Key considerations for the use of rankings by higher education institutions*. European University Association asbl. <https://eua.eu/resources/publications/1080:key-considerations-for-the-use-of-rankings-by-higher-education-institutions.html>
- Eurydice. (2006). *Content and language integrated learning (CLIL) at school in Europe*. Eurydice European Unit.
- Fehringer, C., & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica, Mouton de Gruyter – Societas Linguistica Europaea*, 41(1), 37–72.
- Fillmore, C. J. (1997). *Lectures on Deixis* (Internet Archive). Stanford, Calif.: CSLI Publications. <https://press.uchicago.edu/ucp/books/book/distributed/L/bo3645785.html#anchor-anchored-nav>

- Flowerdew, J. (Ed.). (1994a). *Academic Listening: Research Perspectives*. Cambridge University Press.
- Flowerdew, J. (1994b). Research of relevance to second language lecture comprehension: An overview. *Academic Listening: Research Perspectives*, 7, 29.
- Flowerdew, J., & Tauroza, S. (1995). The Effect of Discourse Markers on Second Language Lecture Comprehension. *Studies in Second Language Acquisition*, 17(4), 435–458. <https://doi.org/10.1017/S0272263100014406>
- Francis, G. (1994). Labelling discourse: An aspect of nominal-group lexical cohesion. In M. Coulthard (Ed.), *Advances in Written Text Analysis*. Routledge.
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31(7), 931–952. [https://doi.org/10.1016/S0378-2166\(98\)00101-5](https://doi.org/10.1016/S0378-2166(98)00101-5)
- Frath, P. (2014). L'anglicisation des formations dans l'enseignement supérieur / coordonné par Pierre Frath. *Les Langues Modernes*, N°1. <https://www.aplv-languesmodernes.org/spip.php?article5430>
- Frath, P. (2017). *Anthropologie de l'anglicisation de l'enseignement supérieur et de la recherche*. Sapientia Hominis.
- Frigols-Martin, M. J., Marsh, D., Mehisto, P., & Wolff, D. (2011). *European Framework for CLIL Teacher Education*. European Centre for Modern Languages, Council of Europe Publishing, Strasbourg. <https://www.ecml.at/ECML-Programme/Programme2008-2011/EuropeanFrameworkforCLILTeacherEducation/tabid/4444/Default.aspx>
- Frost, D., Henderson, A., Wilhelm, J., & Abou Haidar, L. (2021). Perception du discours académique de locuteurs non anglophones: Étude expérimentale. *Recherches en didactique des langues et des cultures*, 3(18–3). <https://doi.org/10.4000/rdlc.10119>
- Gabillon, Z. (2020). Revisiting CLIL: Background, Pedagogy, and Theoretical Underpinnings. *Contextes et Didactiques. Revue Semestrielle En Sciences de l'éducation*, 15, Article 15. <https://doi.org/10.4000/ced.1836>
- Gabriëls, R., & Wilkinson, R. (2020). Resistance to EMI in the Netherlands. In *English-medium instruction and the internationalization of universities* (pp. 49–75). Palgrave Macmillan.
- Galloway, N., Kriukow, J., & Numajiri, T. (2017). *Internationalisation, higher education and the growing demand for English: An investigation into the English medium of instruction (EMI) movement in China and Japan*. British Council. https://www.pure.ed.ac.uk/ws/portalfiles/portal/79051481/h035_eltra_internationalisation_he_and_the_growing_demand_for_english_a4_final_web.pdf

- Genette, G. (1983). *Nouveau discours du récit*. Seuil. <https://www.seuil.com/ouvrage/nouveau-discours-du-recit-gerard-genette/9782020066273>
- Gibbs, G. R. (2014). Analyzing Biographies and Narratives. In J. Hughes & J. Goodwin (Eds.), *Documentary & Archival Research*. Sage. <http://eprints.hud.ac.uk/id/eprint/19253/>
- Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics* (Vol. 10). Cambridge University Press.
- González-Mujico, F. de L., & Lasagabaster, D. (2024). PowerPoint presentations in L1-medium and English-medium lectures: Examining knowledge-building practices in multimodal slide content. *Language and Education*, 0(0), 1–22. <https://doi.org/10.1080/09500782.2024.2346527>
- Graddol, D. (2006). *English next*. British Council. <https://www.teachingenglish.org.uk/publications/case-studies-insights-and-research/english-next>
- Graziani, J.-F. (2014). De la loi Toubon à la loi Fioraso: Quel cadre légal pour les formations en anglais dans les universités françaises? *European Journal of Language Policy*, 6(2), 159–175. <https://doi.org/10.3828/ejlp.2014.10>
- Graziani, J.-F. (2022). L’anglicisation des formations universitaires est-elle une politique linguistique? *European Journal of Language Policy*, 14(1), 53–66. <https://doi.org/10.3828/ejlp.2022.4>
- Grin, F. (2013). L’anglais dans l’enseignement académique: Le débat s’égare dans les clichés. *Le Temps*. <https://archive-ouverte.unige.ch/unige:38201>
- Halliday, M. A. K., & Martin, J. R. (1993). *Writing Science: Literacy And Discursive Power*. Routledge. <https://doi.org/10.4324/9780203209936>
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday’s Introduction to Functional Grammar* (4th ed.). Routledge. <https://www.routledge.com/Hallidays-Introduction-to-Functional-Grammar/Halliday-Matthiessen/p/book/9780203431269#>
- Han, Z., & Selinker, L. (2005). Fossilization in L2 learners. In *Handbook of research in second language teaching and learning* (pp. 455–470). Routledge.
- Hands, P., Mohideen, K., Moore, J., & Williams, D. (2017). *Collins Cobuild English Grammar* (J. Sinclair, M. Airlie, R. Scrimgeour, L. Todd, & C. Wigley, Eds.; 4th ed.). Collins CoBUILD. <https://collins.co.uk/products/9780008135812>
- Hanks, W. F. (2011). Deixis and indexicality. In W. Bublitz & N. R. Norrick (Eds.), *Foundations of Pragmatics* (Vol. 1, p. 725). De Gruyter Mouton. <https://doi.org/10.1515/9783110214260>

- Hansen, C., & Jensen, C. (1995). Evaluating lecture comprehension. In J. Flowerdew (Ed.), *Academic Listening: Research Perspectives* (pp. 241–268). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524612.019>
- Henderson, A. (2019). Training for English-Medium Instruction in French Higher Education: Keeping Pronunciation in Focus. *Recherche et Pratiques Pédagogiques En Langues de Spécialité. Cahiers de l'ApliuT, Vol. 38 N°2, Article Vol. 38 N°2.* <https://doi.org/10.4000/apliut.7402>
- Henderson, A. (2021). *Intelligibility and identity: From teaching pronunciation to training for spoken language variation (document to obtain the “Habilitation à Diriger des Recherches” qualification in France).* Université Grenoble Alpes. <https://hal.science/tel-03295473>
- Héran, F. (2013). L’anglais hors la loi ? Enquête sur les langues de recherche et d’enseignement en France. *Population & Sociétés, 501(6), 1–4.* <https://doi.org/10.3917/popsoc.501.0001>
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language.* Cambridge University Press. <https://doi.org/10.1017/9781316423530>
- Hugonnier, B., De Ketele, J.-M., Cosnefroy, L., & Parmentier, P. (2020). L’internationalisation de l’enseignement supérieur: Quelles opportunités, quels risques? In *L’internationalisation de l’enseignement supérieur: Le meilleur des mondes?* (pp. 187–203). De Boeck Supérieur.
- Hulstijn, J. H., & Laufer, B. (2001). Some Empirical Evidence for the Involvement Load Hypothesis in Vocabulary Acquisition. *Language Learning, 51(3), 539–558.* <https://doi.org/10.1111/0023-8333.00164>
- Hunston, S., & Francis, G. (2000). *Pattern Grammar: A corpus-driven approach to the lexical grammar of English* (Vol. 4). John Benjamins Publishing.
- IDIL. (2022). *Inter-Disciplinary and In-Lab graduate program (IDIL).* Université de Montpellier, France. <https://idil.edu.umontpellier.fr/en/>
- Jablonkai, R. R. (2021). Corpus linguistic methods in EMI research: A missed opportunity? In J. K. H. Pun & S. Curle (Eds.), *Research Methods in English Medium Instruction* (pp. 92–106). Routledge.
- Jenkins, J. (2000a). *The phonology of English as an international language: New models, new norms, new goals.* Oxford University Press. <http://books.google.fr/books?id=0TvHACfrUjEC&printsec=frontcover#v=onepage&q&f=false>

- Jenkins, J. (2000b). *The phonology of English as an international language: New models, new norms, new goals*. Oxford University Press. <http://books.google.fr/books?id=0TvHACfrUjEC&printsec=frontcover#v=onepage&q&f=false>
- Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford University Press.
- Jenkins, J. (2014). *English as a Lingua Franca in the International University: The Politics of Academic English Language Policy*. <https://www.routledge.com/English-as-a-Lingua-Franca-in-the-International-University-The-Politics/Jenkins/p/book/9780415684644>
- Jenkins, J. (2018). *English medium instruction in Higher Education: The role of ELF*. ResearchGate. https://www.researchgate.net/publication/326211635_English_medium_instruction_in_Higher_Education_The_role_of_ELF
- Jung, E. H. (Sarah). (2003). The Role of Discourse Signaling Cues in Second Language Listening Comprehension. *The Modern Language Journal*, 87(4), 562–577.
- JyU Board. (2015). *The University of Jyväskylä language policy*. University of Jyväskylä. <https://www.jyu.fi/en/about-us/organisation-and-management/regulations-and-principles/principles-and-policies>
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: English language in the outer circle. *World Englishes. Critical Concepts in Linguistics*, 241–270.
- Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes*, 9(1), 3–20. <https://doi.org/10.1111/j.1467-971X.1990.tb00683.x>
- Kachru, B. B. (1997). World Englishes and English-Using Communities. *Annual Review of Applied Linguistics*, 17, 66–87. <https://doi.org/10.1017/S0267190500003287>
- Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, 24(3), 269–292. <https://doi.org/10.1016/j.esp.2004.08.003>
- Kappa, K. (2016). Exploring solidarity and consensus in English as lingua franca interactions. *Journal of Pragmatics*, 95, 16–33. <https://doi.org/10.1016/j.pragma.2016.01.015>
- Khan, S., & Aguilar-Pérez, M. (2023). So What Do We Have Here? An Engineering Lecturer's Metadiscursive Use of Rhetorical Questions in L1 and English-Medium Instruction. In B. Bellés-Fortuño, L. Bellés-Calvera, & A.-I. Martínez-Hernández (Eds.), *New Trends on Metadiscourse: An Analysis of Online and Textual Genres* (pp. 11–37). Springer International Publishing. https://doi.org/10.1007/978-3-031-36690-1_2
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1, 7–36.

- Kirkpatrick, T. A. (2011). Internationalization or Englishization: Medium of instruction in today's universities. *Hong Kong: Centre for Governance and Citizenship, The Hong Kong Institute of Education*.
<https://repository.eduhk.hk/en/publications/internationalization-or-englishization-medium-of-instruction-in-t-3>
- Klaassen, R. G. (2001). *The international university curriculum: Challenges in English-medium engineering education* [Doctoral thesis, Delft University of Technology].
<http://resolver.tudelft.nl/uuid:dea78484-b8c2-40d0-9677-6a508878e3d9>
- Klaassen, R. G., & Räsänen, A. (2006). Assessment and staff development in higher education for English-medium instruction: A question-raising article. In *Bridging the assessment gap in English-medium higher education* (pp. 235–255). AKS-verlag.
- Kling, J. (2012). *The development of the Test of Oral English Proficiency for Academic Staff (TOEPAS)* (p. 28). Centre for Internationalisation and Parallel Language Use (CIP), University of Copenhagen.
http://cip.ku.dk/forskning/cip_publicationer/CIP_TOEPAS_Technical_Report.pdf
- Labov, W., & Waletzky, J. (1967). Narrative Analysis: Oral Versions of Personal Experience. *Journal of Narrative and Life History*, 7(1–4), 3–38.
<https://doi.org/10.1075/jnlh.7.02nar>
- Ladilova, A., & Schröder, U. (2022). Humor in intercultural interaction: A source for misunderstanding or a common ground builder? A multimodal analysis. *Intercultural Pragmatics*, 19(1), 71–101. <https://doi.org/10.1515/ip-2022-0003>
- Landrum, R. E., Brakke, K., & McCarthy, M. A. (2019). The pedagogical power of storytelling. *Scholarship of Teaching and Learning in Psychology*, 5(3), 247–253.
<https://doi.org/10.1037/stl0000152>
- LanQua. (2010). *LanQua Toolkit: The Quality Model*. European Commission.
<https://www.lanqua.eu/>
- Lanvers, U., & Hultgren, A. K. (2018). The Englishization of European education: Foreword. *European Journal of Language Policy*, 10(1), 1–11.
- Lasagabaster, D. (2018). Fostering team teaching: Mapping out a research agenda for English-medium instruction at university level. *Language Teaching*, 51(3), 400–416.
- Lasagabaster, D. (2022a). *English-Medium Instruction in Higher Education*. Cambridge University Press. <https://doi.org/10.1017/9781108903493>
- Lasagabaster, D. (2022b). Teacher preparedness for English-medium instruction. *Journal of English-Medium Instruction*, 1, 48–64. <https://doi.org/10.1075/jemi.21011.las>

- Lasagabaster, D., & Doiz, A. (2021). *Language Use in English-Medium Instruction at University: International Perspectives on Teacher Practice*.
- Lasagabaster, D., & Pagèze, J. (2018). Teacher Development for Teaching and Learning in English in a French Higher Education Context. *L'analisi Linguistica e Letteraria*, 25, 289–310.
- Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64(4), 367–375. <https://doi.org/10.1093/elt/ccp082>
- Le Lièvre, F. (2008). *L'Anglais en France, une langue multiple: Pratiques, représentations et postures d'apprentissage: une étude en contexte universitaire* [These de doctorat, Tours]. <http://www.theses.fr/en/2008TOUR2015>
- Le Lièvre, F. (2021). An ambivalent picture: The Englishization of higher education in France. In R. Wilkinson & R. Gabriëls (Eds.), *The Englishization of Higher Education in Europe* (pp. 97–119). Amsterdam University Press. https://doi.org/10.5117/9789463727358_ch05
- Le Lièvre, F., Anquetil, M., Derivry-Plard, M., Fäcke, C., & Verstraete-Hansen, L. (Eds.). (2018). *Langues et cultures dans l'internationalisation de l'enseignement supérieur au XXIe siècle*. <https://www.peterlang.com/document/1055789>
- Le Lièvre, F., & Forlot, G. (2014). Éducation et plurilinguisme: Quel rôle pour l'anglais? *Cahiers internationaux de sociolinguistique*, 5(1), 161–171.
- Lebauer, R. S. (1984). Using Lecture Transcripts in EAP Lecture Comprehension Courses. *TESOL Quarterly*, 18(1), 41–54. <https://doi.org/10.2307/3586334>
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Levinson, S. C. (2001). Pragmatics. In N. Smelser & P. Baltes (Eds.), *International Encyclopedia of Social and Behavioral Sciences* (Vol. 17, pp. 11948–11954). Oxford: Pergamon. <https://www.mpi.nl/publications/item66467/pragmatics>
- Lightbown, P. (2014). *Focus On Content-Based Language Teaching*. Oxford University Press.
- Luukka, M.-R. (1992). Metadiscourse in academic texts. *Text and Talk in Professional Contexts: Selected Papers from the International Conference "Discourse and the Professions" Uppsala, 26 - 29 August, 1992*, 77–88.
- Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge University Press.
- Lyons, J. (1977). *Semantics: Volume 2*. Cambridge University Press.
- Lyons, J. (1995). *Linguistic Semantics: An Introduction*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511810213>

- Lyster, R. (2017). Integrating Content and Language in Higher Education: Preface. In J. Valcke & R. Wilkinson (Eds.), *Integrating Content and Language in Higher Education: Perspectives on Professional Practice* (pp. 7–14). Peter Lang Publishing Group.
- 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. <https://doi.org/10.1017/S0261444817000350>
- Malavska, V. (2016). Genre of an Academic Lecture. *International Journal on Language, Literature and Culture in Education*, 3. <https://doi.org/10.1515/llce-2016-0010>
- Malmström, H., Mežek, Š., Pecorari, D., Shaw, P., & Irvine, A. (2017). Engaging with Terminology in the Multilingual Classroom: Teachers' Practices for Bridging the Gap Between L1 Lectures and English Reading. *Classroom Discourse*, 8(1), 3–18. <https://doi.org/10.1080/19463014.2016.1224723>
- Malmström, H., & Pecorari, D. (2021). Epilogue: Disciplinary Literacies as a Nexus for Content and Language Teacher Practice. In *Language Use in English-Medium Instruction at University* (p. 9). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003134534-11/epilogue-hans-malmstr%C3%B6m-diane-pecorari>
- Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press. <https://mitpress.mit.edu/9780262133609>
- Marginson, S., & van der Wende, M. (2011). Européanisation, classements internationaux et mobilité des universitaires: Trois exemples de mondialisation de l'enseignement supérieur. In OCDE (Ed.), *L'enseignement supérieur à l'horizon 2030 – Volume 2: Mondialisation*. Éditions OCDE. <http://dx.doi.org/10.1787/9789264075405-fr>
- Marmaridou, S. (2000). Pragmatic Meaning and Cognition. In *Pbns.72*. John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.72>
- Martin, J. R. (2008). Negotiating Values: Narrative and Exposition. *Journal of Bioethical Inquiry*, 5, 41–55. <https://doi.org/10.1007/s11673-008-9091-8>
- Matthews, P. H. (2014). The Concise Oxford Dictionary of Linguistics. In *The Concise Oxford Dictionary of Linguistics* (3rd ed., 1–1). Oxford University Press. <https://www.oxfordreference.com/display/10.1093/acref/9780199675128.001.0001/acref-9780199675128>
- McArthur, T. (1987). The English languages? *English Today*, 3(3), 9–13. Cambridge Core. <https://doi.org/10.1017/S0266078400013511>

- McNett, G. (2016). Using Stories to Facilitate Learning. *College Teaching*, 64(4), 184–193. <https://doi.org/10.1080/87567555.2016.1189389>
- Mehisto, P., Marsh, D., & Frigols-Martin, M. J. (2008). *Uncovering CLIL*. Macmillan Education. <http://archive.org/details/284060456-uncovering-clil-1-pdf>
- Miller, L. (2009). Engineering Lectures in a Second Language: What Factors Facilitate Students' Listening Comprehension? *The Asian EFL Journal Quarterly*, 11(2), 8–30.
- Minto, B. (2021). *The Pyramid Principle: Logic in Writing and Thinking* (3rd ed.). FT Publishing International.
- Modiano, M. (1999). International English in the global village. *English Today*, 15(2), 22–28. <https://doi.org/10.1017/S026607840001083X>
- Moore, B., & Détourbe, M.-A. (2019). *EMINSA - English Medium Instruction at INSA ou comment développer des enseignements scientifiques en anglais à l'INSA Toulouse*. 6ème colloque “pédagogie et formation.” <https://pedago-insa2019.sciencesconf.org/>
- Musselin, C. (2009). The Side Effects of the Bologna Process on National Institutional Settings: The Case of France. In A. Amaral, G. Neave, C. Musselin, & P. Maassen (Eds.), *European Integration and the Governance of Higher Education and Research* (pp. 181–205). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9505-4_8
- Nesi, H. (2012). Laughter in university lectures. *Journal of English for Academic Purposes*, 11(2), 79–89. <https://doi.org/10.1016/j.jeap.2011.12.003>
- Nesi, H., & Alsop, S. (2021). Stories and scenarios: Lecturers' use of fantastic hypothetical events. *Journal of English for Academic Purposes, Special Issue JEAP 101022*. <https://doi.org/10.1016/j.jeap.2021.101022>
- Nesi, H., Alsop, S., Ahmad, U. K., & Grant, L. (2014). *Engineering Lecture Corpus (ELC)*. Coventry, U.K.: Coventry University. [Last Access: 25/04/2023]. <https://www.coventry.ac.uk/research/research-directories/current-projects/2015/engineering-lecture-corpus-elc/>
- Nesi, H., & Moreton, E. (2012). EFL/ESL writers and the use of shell nouns. In R. Tang (Ed.), *Academic Writing in a Second Or Foreign Language: Issues and Challenges Facing ESL/EFL Academic Writers in Higher Education Contexts* (pp. 126–145). Continuum Publishing Corporation.
- Newbold, D. (2023). Framing the Framework: Four Decades of Change in Language Teaching (and the Long March of ELF). In *La linguistica educativa tra ricerca e sperimentazione Scritti in onore di Carmel Mary Coonan* (pp. 153–164). <https://doi.org/10.30687/978-88-6969-683-1/014>

- O'Connell, A.-M., & Chaplier, C. (2014). Quelle anglicisation des formations ? *Les Langues Modernes, n°1*, Article n°1.
- O'Dowd, R. (2018). The training and accreditation of teachers for English medium instruction: An overview of practice in European universities. *International Journal of Bilingual Education and Bilingualism*, 25/06/2018. <https://doi.org/10.1080/13670050.2018.1491945>
- Olsen, L. A., & Huckin, T. H. (1990). Point-driven understanding in engineering lecture comprehension. *English for Specific Purposes*, 9(1), 33–47. [https://doi.org/10.1016/0889-4906\(90\)90027-A](https://doi.org/10.1016/0889-4906(90)90027-A)
- Onysko, A. (Ed.). (2021). *Research Developments in World Englishes*. Bloomsbury Academic. <https://doi.org/10.5040/9781350167087>
- Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive Load Theory: New Conceptualizations, Specifications, and Integrated Research Perspectives. *Educational Psychology Review*, 22(2), 115–121. <https://doi.org/10.1007/s10648-010-9133-8>
- Pagèze, J. (2017). *Défi international*. Université de Bordeaux [Accessed April 11th, 2024]. <http://langues.u-bordeaux.fr/Sciences-de-l-Homme-et-sante/Defi-international>
- Paikeday, T. M. (1985). *The Native Speaker Is Dead: An Informal Discussion of a Linguistic Myth With Noam Chomsky and Other Linguists, Philosophers, Psychologists, and Lexic*. Paikeday Pub Co.
- Patel, M., Solly, M., Copeland, S., O'Sullivan, B., & Jin, Y. (2023). *The Future of English: Global Perspectives*. British Council. <https://british-council.shorthandstories.com/the-future-of-english-publication-launch/>
- Patsko, L. (2013). Using the Lingua Franca Core to promote students' mutual intelligibility in the multilingual classroom: Five teachers' experiences. *MA Thesis, ELT & Applied Linguistics*. https://www.academia.edu/17003943/Using_the_Lingua_Franca_Core_to_promote_students_mutual_intelligibility_in_the_multilingual_classroom_five_teachers_experiences
- Patsko, L. (2014). Pronunciation in English as a lingua franca: A brief introduction. *The Journal of the English Language Teachers' Association of Switzerland. Volume, 31*, 28–29.
- Pérez Cañado, M. (2021). *CLIL-ising EMI: An Analysis of Student and Teacher Training Needs in Monolingual Contexts* (pp. 171–191). https://doi.org/10.1007/978-3-030-70095-9_9

- Pérez, M. A., & Macià, E. A. (2002). Metadiscourse in Lecture Comprehension: Does it Really Help Foreign Language Learners? *Atlantis: Revista de La Asociación Española de Estudios Anglo-Norteamericanos*, 24(1), 7–22.
- Peters, P. (2004). *The Cambridge guide to English usage*. Cambridge University Press (CUP).
- Phillipson, R. (2006). English, a Cuckoo in the European Higher Education Nest of Languages? *European Journal of English Studies*, 10, 13–32. <https://doi.org/10.1080/13825570600590846>
- Phillipson, R. (2008). English, panacea or pandemic? *International Conference: Language Issues in English-Medium Universities: A Global Concern*. University of Hong Kong.
- Picavet, F. (2022). *CoFEL - A Corpus of French Engineering Lectures*. Université Grenoble Alpes, Grenoble, France. Unpublished.
- Picavet, F., Henderson, A., & de Vries, E. (2023). Storytelling in L2 English-medium engineering lectures: A typology. *Journal of English for Academic Purposes*, 62. <https://doi.org/10.1016/j.jeap.2022.101214>
- Pickering, L., Corduas, M., Eisterhold, J., Seifried, B., Eggleston, A., & Attardo, S. (2009). Prosodic Markers of Saliency in Humorous Narratives. *Discourse Processes*, 46(6), 517.
- Pitzl, M.-L., Seidlhofer, B., & Breiteneder, A. (2006). English as a lingua franca in Europe: Challenges for applied linguistics. *Annual Review of Applied Linguistics*, 26. https://www.academia.edu/5796331/English_as_a_lingua_franca_in_Europe_challenges_for_applied_linguistics
- Räsänen, A. (2011). *The Promise and challenge of CLIL (Content and Language Integrated Learning) as a mediator for Internationalisation*. Conference: Multilingual Competences for Professional and Social Success in Europe, Warsaw, Poland. <http://www.celelc.org/archive/Projects-and-Reports/052-Raesaenen-CLIL/index.html>
- Räsänen, A., & Klaassen, R. G. (2006). From learning outcomes to staff competences in integrated content and language instruction at the higher education level. In *Bridging Assessment Gap in English-Medium Higher Education* (pp. 256–278). AKS-verlag.
- Reece, I., & Walker, S. (2007). *Teaching Training and Learning: A practical guide* (6th edition). Business Education Publishers Ltd.
- Richards, J. C., & Pun, J. K. H. (2023). A Typology of English-Medium Instruction. *RELC Journal [First Published Online February 8, 2021]*, 54(1), 216–240. <https://journals-sagepub-com.gaelnomade-1.grenet.fr/doi/full/10.1177/0033688220968584>

- Richards, J. C., & Schmidt, R. (2010). *Longman Dictionary of Language Teaching and Applied Linguistics* (4th ed.). Routledge.
- Richter, K. (2019). *English-Medium Instruction and Pronunciation* (Multilingual Matters | Channel View Publications). http://www.multilingual-matters.com/display.asp?sort=sort_date/d&sf1=title_exact&st1=englishmediuminstructionandpronunciation
- Robertson, R. (1992). *Globalization: Social Theory and Global Culture* (1er édition). SAGE Publications Ltd.
- Robson, S., Almeida, J., & Schartner, A. (2017). Internationalization at home: Time for review and development? *European Journal of Higher Education*, 8(1), 19–35. <https://doi.org/10.1080/21568235.2017.1376697>
- Robson, S., & Wihlborg, M. (2019). Internationalisation of higher education: Impacts, challenges and future possibilities. *European Educational Research Journal*, 18, 127–134. <https://doi.org/10.1177/1474904119834779>
- Roussel, S., Joulia, D., Tricot, A., & Sweller, J. (2017). Learning subject content through a foreign language should not ignore human cognitive architecture: A cognitive load theory approach. *Learning and Instruction*, 52, 69–79. <https://doi.org/10.1016/j.learninstruc.2017.04.007>
- Roussel, S., & Tricot, A. (2017). Emile/CLIL, EMI vs TCC: Des espaces de recherche si proches et si lointains. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication, Vol. 20, n° 3*. <https://doi.org/10.4000/alsic.3126>
- Saarinen, T., & Rontu, H. (2018). University language policies: How does Finnish constitutional bilingualism meet the needs for internationalisation in English? *European Journal of Language Policy*, 10(1), 97–119. <https://doi.org/10.3828/ejlp.2018.5>
- Samarin, W. J. (1987). *Lingua Franca*. Walter de Gruyter. <https://tspace.library.utoronto.ca/handle/1807/70765>
- Sancho Guinda, C. (2013). *Teacher Targets: A model for CLIL and ELF teacher education in polytechnic settings*. https://www.academia.edu/5544662/Teacher_Targets_A_model_for_CLIL_and ELF_teacher_education_in_polytechnic_settings
- Schourup, L. (1999). Discourse markers. *Lingua*, 107(3), 227–265. [https://doi.org/10.1016/S0024-3841\(96\)90026-1](https://doi.org/10.1016/S0024-3841(96)90026-1)

- Schutte, C., Klein, L., & Roth, S. (2020). *English for staff: Lecturers, researchers and administration*. Hochschule RheinMain [Accessed April 11th, 2024]. <https://www.hs-rm.de/de/english-for-staff>
- Scott, M. (2020). *WordSmith Tools (Version 8.0) [Computer Software]* [Computer software]. Lexical Analysis Software Ltd., Stroud, UK. <http://lexically.net/wordsmith/index.html>
- Seidlhofer, B. (2000). Mind the gap: English as a mother tongue vs. English as a lingua franca. *VIEWS – Vienna English Working Papers*, 9(1), 51–68.
- Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, 24. <https://doi.org/10.1017/S0267190504000145>
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59(4), 339–341.
- Seidlhofer, B. (2013). *Understanding English as a Lingua Franca—Oxford Applied Linguistics*. Oxford University Press.
- Seidlhofer, B., & Jenkins, J. (2003). English as a Lingua Franca and the Politics of Property. In *The Politics of English as a World Language* (pp. 139–154). Brill. https://doi.org/10.1163/9789401200929_012
- Siirtola, H., Säily, T., Nevalainen, T., & Rähä, K.-J. (2014). Text Variation Explorer: Towards interactive visualization tools for corpus linguistics. *International Journal of Corpus Linguistics*, 19(3), 417–429. <https://doi.org/10.1075/ijcl.19.3.05sii>
- Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse* (R. Carter, Ed.). Routledge. <https://doi.org/10.4324/9780203594070>
- Smit, U. (2023). English-medium instruction (EMI). *ELT Journal*, ccad018, 6. <https://doi.org/10.1093/elt/ccad018>
- Smit, U., & Dafouz-Milne, E. (2012). Integrating content and language in higher education: An introduction to English-medium policies, conceptual issues and research practices across Europe. *Aila Review*, 25(1), 1–12.
- Smith, K. (2004). Studying in an additional language: What is gained, what is lost and what is assessed. In R. Wilkinson (Ed.), *Integrating content and language: Meeting the challenge of a multilingual higher education* (pp. 78–93). Universitaire Pers Maastricht Maastricht, The Netherlands.
- Straker Cook, R. H. (1975). *Communicative approach to the analysis of extended monologue discourse and its relevance to the development of teaching materials for English special purposes* [University of Edinburgh]. <https://era.ed.ac.uk/handle/1842/7249>

- Studer, P. (2018). English in the age of comprehensive internationalization: Defining competence parameters for teachers in higher education. *Bulletin Suisse de Linguistique Appliquée*, 107, 27–47.
- Studer, P., & Smit, U. (2021). English-medium education in internationalised universities: Sketching a way forward. *European Journal of Language Policy*, 13, 129–140. <https://doi.org/10.3828/ejlp.2021.8>
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory and Cognition*, 5(4), 360–367.
- Sweller, J. (2018). Measuring cognitive load. *Perspectives on Medical Education*, 7(1), 1–2. <https://doi.org/10.1007/s40037-017-0395-4>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer.
- Tadros, A. A. (1989). Predictive categories in university textbooks. *English for Specific Purposes*, 8(1), 17–31. [https://doi.org/10.1016/0889-4906\(89\)90004-5](https://doi.org/10.1016/0889-4906(89)90004-5)
- Tadros, A. A. (1994). Predictive categories in expository text. In *Advances in Written Text Analysis*. Routledge.
- Taillefer, G. (2009). EMILE en France: Quelle offre ? *Les Après-midi de LAIRDIL*, 15, 59.
- Taillefer, G. (2011). Developing Romanian International Master's Level Education: Reflections on Change Management. *Quality Assurance Review for Higher Education*, 3(2), 126–133.
- Taillefer, G. (2013). CLIL in higher education: The (perfect?) crossroads of ESP and didactic reflection. *ASp. La Revue Du GERAS*, 63, 31–53. <https://doi.org/10.4000/asp.3290>
- Thompson, S. E. (2003). Text-structuring metadiscourse, intonation and the signalling of organisation in academic lectures. *Journal of English for Academic Purposes*, 2(1), 5–20. [https://doi.org/10.1016/S1475-1585\(02\)00036-X](https://doi.org/10.1016/S1475-1585(02)00036-X)
- Tremblay, C. (2022). *L'impératif plurilingue: 18 ans avec l'Observatoire européen du plurilinguisme*. Bookelis.
- Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905–916.
- Truchot, C. (2005). L'anglais comme « lingua franca »: Observations sur un mode de majoration. *Cahiers de sociolinguistique*, 10(1), 167–178. <https://doi.org/10.3917/csl.0501.0167>

- Truchot, C. (2010). *L'enseignement supérieur en anglais véhiculaire: La qualité en question*. Diploweb.Com : La Revue Géopolitique. <https://www.diploweb.com/L-enseignement-superieur-en.html>
- Truchot, C. (2013). Enseigner en anglais dans les universités françaises: Quels effets prévisibles? *Éducation et Sociétés Plurilingues*, 35, 87–90. <https://doi.org/10.4000/esp.2792>
- Tudor, I. (2009). Content and language integrated learning (CLIL) in higher education in Europe. *Les Après-Midi de LAIRDIL*, 15, 9–27.
- Tzoannopoulou, M. (2016). Humor in academic lectures: A case study in an ELF context. *ELF: Pedagogical and Interdisciplinary Perspectives*, 78–86.
- UiB Board. (2019). *Language policy guidelines*. Regelsamlingen Ved UiB, Universitetet i Bergen, Norge/Norway. https://regler.app.uib.no/regler_en/Part-2-Research-Education-Dissemination/2.3-Dissemination/Maalbruk/Language-policy-guidelines/
- Valcke, J., & Wilkinson, R. (2017). Integrating Content and Language in Higher Education: Perspectives on Professional Practice. In *Peter Lang Publishing Group*. Peter Lang Publishing Group. <https://doi.org/10.3726/978-3-653-07263-1>
- Vilkancienė, L. (2011). CLIL in Tertiary Education: Does it Have Anything to Offer? *Studies about Languages*, 18, Article 18. <https://doi.org/10.5755/j01.sal.0.18.418>
- Vinke, A. A. (1995). *English as the medium of instruction in Dutch engineering education* [PhD Thesis, Delft University of Technology]. <https://repository.tudelft.nl/islandora/object/uuid%3A491b55f9-fbf9-4650-a44d-acb9af8412a8>
- VOICE. (2021). *VOICE - The Vienna-Oxford International Corpus of English (version VOICE 3.0 Online)*. <https://voice.acdh.oeaw.ac.at/>
- VOICE Project. (2007). *VOICE Transcription Conventions [2.1]* (Version Accessed 11/05/2022) [Computer software]. <https://voice.acdh.oeaw.ac.at/transcription-conventions/>
- Wächter, B. (2018). *Internationalization. A question of quality?* University of Antwerp. <https://blog.uantwerpen.be/impala/internationalisation-a-question-of-quality/>
- Wächter, B., & Maiworm, F. (2002). *English-Language-Taught Degree Programmes in Higher Education: Trends and Success Factors*. Academic Cooperation Association (ACA), Lemmens.
- Wächter, B., & Maiworm, F. (2008). *English-Taught Programmes in European Higher Education: The Picture in 2007*. Academic Cooperation Association (ACA), Lemmens.

- Wächter, B., & Maiworm, F. (2014). *English-Taught Programmes in European Higher Education: The State of Play in 2014*. Academic Cooperation Association (ACA), Lemmens.
- Walker, R. (2021). *Teaching the pronunciation of English as a lingua franca*. Oxford University Press.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.
- Wijasuriya, B. S. (1971). *Occurrence of discourse markers and inter-sentence connectives in university lectures and their place in the testing and teaching of listening comprehension in English as a foreign language* [Ph.D., University of Manchester]. <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.497170>
- Wilang, J. D., Ngamchaturat, T., & Seepho, S. (2024). Bridging Content and Language in English-medium Engineering Programs. *International Journal of Learning, Teaching and Educational Research*, 23(4), Article 4. <http://mail.ijlter.org/index.php/ijlter/article/view/9978>
- Wilkinson, R. (2013). English-medium instruction at a Dutch university: Challenges and pitfalls. In A. Doiz, D. Lasagabaster, & J. M. Sierra (Eds.), *English-medium instruction at universities: Global challenges* (pp. 3–24).
- Wilkinson, R. (2018). Content and language integration at universities? Collaborative reflections. *International Journal of Bilingual Education and Bilingualism*, 21(5), 607–615. <https://doi.org/10.1080/13670050.2018.1491948>
- Wilkinson, R., & Gabriëls, R. (2019). *Quality of disciplinary content learning in EMI degree programmes* [Keynote Communication]. 6th ICLHE Conference: Multilingualism and Multimodality in Higher Education, Universitat Jaume I, Castellón, Spain.
- Wilkinson, R., & Gabriëls, R. (Eds.). (2021). *The Englishization of Higher Education in Europe*. Amsterdam University Press. <https://doi.org/10.5117/9789463727358>
- Wilkinson, R., & Walsh, M. L. (Eds.). (2014). *Integrating Content and Language in Higher Education*. <https://www.peterlang.com/document/1050371>
- Wilkinson, R., & Zegers, V. (Eds.). (2008). *Realizing Content and Language Integration in Higher Education*. Maastricht University, Maastricht, Netherlands.
- Windelband, W. (1998). History and Natural Science. (*Original Publication: 1894*) *Theory & Psychology*, 8(1), 5–22. <https://doi.org/10.1177/0959354398081001>

- Wingate, U. (2022). Student support and teacher education in English for Academic Purposes and English Medium Instruction: Two sides of the same coin? *Language Teaching*, 1–12. <https://doi.org/10.1017/S0261444822000465>
- Yaakob, S. B. (2013). *A genre analysis and corpus based study of university lecture introductions* [PhD Thesis, University of Birmingham]. <https://etheses.bham.ac.uk/id/eprint/5241/>
- Yeo, J.-Y., & Ting, S.-H. (2014). Personal pronouns for student engagement in arts and science lecture introductions. *English for Specific Purposes*, 34, 26–37. <https://doi.org/10.1016/j.esp.2013.11.001>
- Young, L. (1994). University lectures – macro-structure and micro-features. In J. Flowerdew (Ed.), *Academic Listening: Research Perspectives* (pp. 159–176). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524612.013>
- Yule, G. (1996). *Pragmatics*. Oxford University Press.

Appendices

Appendix A	Recording consent form: lecturers, French version.....	170
Appendix B	Recording consent form: students, English version	171
Appendix C	Examples of stacked summary displays across three 2-hour lectures	172
Appendix D	Summarizing attributes in the ELC: exclusive qualitative features	173
Appendix E	Summarizing attributes in the ELC: exclusive lexicogrammatical features...	174
Appendix F	Summarizing attributes in the ELC: exclusive lexicogrammatical features (cont.)	175
Appendix G	Example of a RevPRE summary in initial position.....	176
Appendix H	Longest RevPRE summary in CoFEL.....	177
Appendix I	Longest PreCUR summary in CoFEL	179
Appendix J	Longest RevCUR summary.....	180
Appendix K	Example of a chain of four summaries combining two types	181

Appendix A. Recording consent form: lecturers, French version



Formulaire d'autorisation d'enregistrement audio ou vidéo d'une personne majeure

Version française – Mise à jour : 01/2022 – English version overleaf

Consent Form_FDP_VIDEO_FR-ENG_PROFs_v1

Je, soussigné(e) : Nom: Prénom :

- déclare appartenir au personnel enseignant de (ou dans cette autre institution partenaire:)
- déclare avoir été informé-e que des enregistrements ont lieu dans mon cours aujourd'hui, et certains cours suivants de cette année, dans le cadre du projet de recherche de doctorat entrepris par M. Francis Picavet: «**L'intégration d'une matière et d'une langue : étude de l'enseignement en langue anglaise dans la formation d'ingénieurs**».

J'autorise l'Université Grenoble Alpes (ci-après UGA) et (ci-après), en la personne de Francis Picavet, doctorant attaché administrativement au LaRAC- E.A. 602 (un laboratoire de recherche de l'UGA) et professeur d'anglais à l'INP, à utiliser les enregistrements réalisés de manière **anonyme**.

Cette autorisation couvre tous les aspects du projet de chercheur, au sein des établissements publics d'enseignement supérieur et à des fins non commerciales, notamment pour:

- la thèse de doctorat du chercheur;
- la participation à d'autres projets scientifiques pertinents;
- les publications scientifiques de toutes sortes;
- des communications lors d'événements scientifiques;
- l'élaboration de supports de formation.

Conformément à la loi n°. 78-17 du 6 janvier 1978, telle que modifiée par la loi no. 2004-801 du 6 août 2004, je dispose d'un droit d'accès, de rectification et de suppression des informations personnelles. Toute demande d'accès, de correction, d'opposition ou de suppression doit être adressée par courrier électronique au chercheur du projet, Francis Picavet, à l'adresse suivante : Francis.Picavet@grenoble-inp.fr

Le présent formulaire de consentement est valable pendant SEPT ANS ; au-delà de cette période, un nouveau formulaire de consentement doit être signé pour continuer à utiliser les enregistrements.

L'UGA et Grenoble INP s'engagent à ne pas reproduire ni diffuser les enregistrements pour tout usage susceptible de porter atteinte à ma réputation, mon honneur et ma dignité.

Ville: Date:

Signature:



Laboratoire de Recherche sur les Apprentissages en Contexte -



Appendix B. Recording consent form: students, English version



Consent form for audio or video recording of an adult

English version – Last update: **04/2020** – **Version française au verso**

Consent Form_FDP_AUDIO_FR-ENG_v2

I, undersigned: Last Name: Cliquez ou appuyez ici pour entrer du texte. First Name: Cliquez ou appuyez ici pour entrer du texte.

declare that I am a student at
(or this other partner institution: Cliquez ou appuyez ici pour entrer du texte.)

- My student number: Cliquez ou appuyez ici pour entrer du texte.

I declare that I have been informed that recordings take place in class today, and some sessions later in the year, as part of the doctorate research project undertaken by Mr Francis Picavet: **“Integrating content and language – a study of English-medium instruction in engineering education”**.

I authorize the *Université Grenoble Alpes* (hereafter UGA) and (hereafter), in the person of Francis Picavet, a PhD researcher administratively attached to the LaRAC– E.A. 602 (one of the UGA laboratories) and a teacher of English at INP, to use those recordings **anonymously**.

This authorization includes all the aspects of the researcher’s project, within public higher education institutions and for non-commercial purposes, in particular for:

- the researcher’s doctorate thesis;
- participation in other relevant scientific projects;
- scientific publications of all kinds;
- communication in scientific events;
- the development of training materials.

In compliance with the French Data Protection Act no. 78-17 of 6 January 1978, as amended by Act no. 2004-801 of 6 August 2004, I have a right of access, rectification and deletion of personal information. Any request for access, correction, opposition or deletion shall be sent by e-mail to the project researcher, Francis Picavet, at the following address: Francis.Picavet@grenoble-inp.fr

The present consent form is valid for SEVEN YEARS, beyond which a new consent form must be signed to continue to use the recordings.

UGA and Grenoble INP undertake to neither reproduce nor distribute the recordings, for any use susceptible to infringing upon my reputation, honor and dignity.

City: Cliquez ou appuyez ici pour entrer du texte. Date: Cliquez ou appuyez ici pour entrer une date.

Signature: Cliquez ou appuyez ici pour entrer du texte.

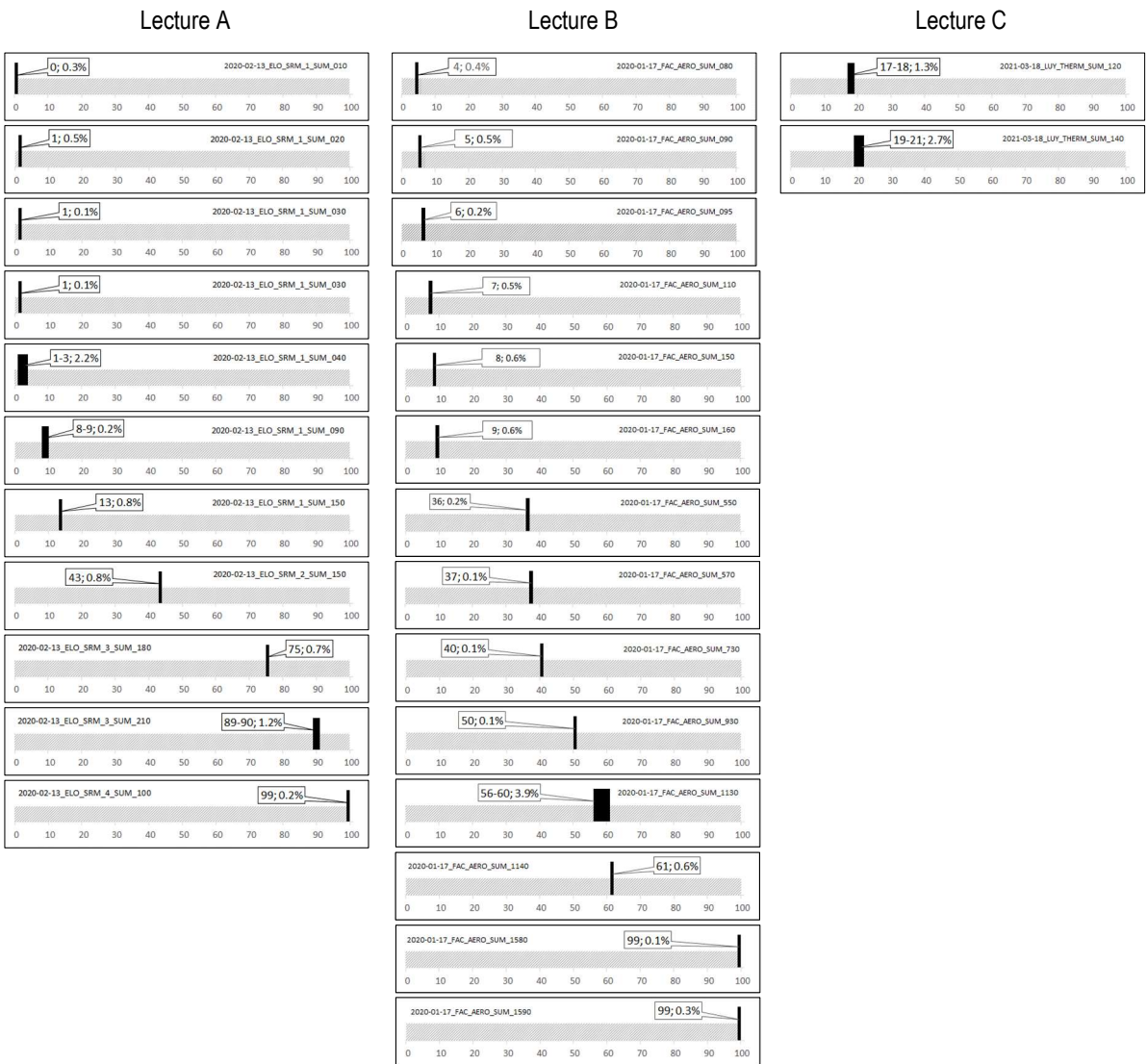


Laboratoire de Recherche sur les Apprentissages en Contexte -



Appendix C. Examples of stacked summary displays across three 2-hour lectures

All the summary graphs of a lecture can be stacked in chronological order, providing an overview of all the session summaries at once. The table below shows three concrete examples of stacked summaries found in three distinct two-hour lectures. This visualization shows clearly that the practices of lecturers can be drastically different: some lecturers use summaries very regularly from the beginning to the end of the session, while others use summarizing much less regularly, and summary locations and durations can vary greatly.



Note. The three lectures: Soil & Rock Mechanics (A – 10 summaries), Aeroelasticity (B – 14 summaries), and Thermodynamics (C – 2 summaries). The stacking of summaries facilitates descriptive and comparative approaches, highlighting differences between lectures or lecturer practices.

Appendix D. Summarizing attributes in the ELC: exclusive qualitative features

	Review previous (RevPRE)	Review current (RevCUR)	Preview current (PreCUR)	Preview future (PreFUT)
Function	<p>Previously delivered information:</p> <ul style="list-style-type: none"> – is identified – is recapitulated – is fixed in time – is acknowledged as being complex <p>For students, RevPRE acts as:</p> <ul style="list-style-type: none"> – a reference point – a cognitive anchor – groundwork for future learning 	<p>Recent or just given information is:</p> <ul style="list-style-type: none"> – highlighted – reminded – repackaged <p>For students, RevCUR seeks to:</p> <ul style="list-style-type: none"> – improve assimilation & learning – function both as a reminder and signal of importance, esp. when immediate 	<p>For students, PreCUR:</p> <ul style="list-style-type: none"> – announces or outlines upcoming content – performs an explicit <i>metadiscursive</i> signaling function (Crawford Camiciottoli, 2004) – has a strong discourse-structuring function (Young, 1994) – scaffolds learning in the current lecture – uses down-toning strategies to minimize imposition 	<p>For students, PreFUT:</p> <ul style="list-style-type: none"> – has a higher-level information-structuring function – announces information which can be put on the backburner – reassures that complex problems are not for now – explains that what is done now fits into the overall learning outcomes
Constituents	<p>Concepts or processes can be either:</p> <ul style="list-style-type: none"> – simply named – reiterated in detail <p>RevPRE includes reviews of information (in order of frequency):</p> <ol style="list-style-type: none"> 1. in the immediately preceding lecture of “last week” 2. any other preceding lectures across in the module 3. any other preceding lectures across years of study 	<p>Specific & non-specific time references mostly absent</p>	<p>At the beginning of lectures:</p> <ul style="list-style-type: none"> – plots the course of the upcoming class in detail – tends to include references to more than one topic – often provides a rationale for the upcoming content <p>Later in the class:</p> <ul style="list-style-type: none"> – immediate topic is briefly outlined prior to delivery – later topic can include a time reference 	<p>In PreFUTs:</p> <ul style="list-style-type: none"> – minimal detail is given about upcoming lecture content – lecturers are brief and direct in delivering information – topics are mostly named and not elucidated – the distance in time to delivery allows lecturers be more explicit about difficulty (no minimizing)
Distribution	<p>Some clustering in the 1st third of lectures, but can occur in any part of the lecture.</p>	<p>Some clustering towards the end of lectures, but can occur in any part of the lecture. Regular uniformity of occurrence in the central 60%. Short bursts are dominant.</p>	<p>Some clustering towards the beginning of lectures, but can occur in any part of the lecture. By far the most common form of summarizing. Happens most often, but for the shortest periods.</p>	<p>Some clustering towards the end of lectures, but can occur in any part of the lecture. The least common form of summarizing Relatively short and infrequent: lecturers do not prioritize forward-scaffolding of content.</p>
Miscellaneous	<p>Double requirement for attention:</p> <ul style="list-style-type: none"> – lecturer deems information worthy of review – information explicitly marked as complex through negative evaluation (<i>more difficult; quite complicated; not easy</i>) 	<p>Lecturers emphasize the importance of summarized information:</p> <ul style="list-style-type: none"> – with comparisons to equally complex concepts or processes – by acknowledging the challenges in the delivery/reception of the information 	<p>Sometimes:</p> <ul style="list-style-type: none"> – functions as quick alerts embedded within longer stretches of content – lecturers present upcoming content as a puzzle that the class will solve – lecturers raise questions as if the solution was not already determined <p>The longer the distance in previewing, the more elaborate the attributes become.</p>	<p>Typically contains less information than reviews: students are alerted to upcoming topics, but the level of detail is minimal. Lecturers usually name the topic and when it will be delivered, often with an anaphoric reference to the current topic.</p>

Note. Adapted from the Engineering Lecture Corpus (ELC - Alsop, 2015; Nesi et al., 2014)

Appendix E. Summarizing attributes in the ELC: exclusive lexicogrammatical features

	Review previous (RevPRE)	Review current (RevCUR)	Preview current (PreCUR)	Preview future (PreFUT)
Discourse functions	<p>Discourse functions are frequently named (e.g., <i>recap</i> and <i>summary</i>).</p> <p><i>Advance labelling</i> is used (Tadros, 1989)</p> <p>Meta-statements concerning what is about to occur are used (<i>let's recap</i>, etc.)</p>	<p>Occasional naming of discourse acts and specifying occurrence signals immediacy (<i>going back to what we said a few moments ago...</i>)</p> <p>The message is “notice should be taken”, as content is worth immediate or near repetition</p>	<p>Transitions between topics discursively marked by <i>right</i>, <i>now</i> and <i>ok</i>:</p> <ul style="list-style-type: none"> – as boundary markers at the end of previous topic – as short alerts providing a ‘mini-introduction’ to new topics – as authorial commitment to perform a discourse act, occurring cyclically throughout lectures, in <i>advance labelling</i> (Tadros, 1989) 	<p>Refers to upcoming content that expands on the current work, with temporal discourse markers, vague or precise: <i>later</i> or <i>next</i> + temporal deixis (<i>weekday</i>, <i>week</i>, <i>month</i>, <i>term</i>, <i>semester</i>, <i>year</i>, etc.)</p>
Morphosyntactic patterns	<p>Strongly characterized by: temporal deixis*/specific time frame + pronoun + Simple Past tense verb** + topic reference/object-concept of review (e.g., <i>last week we looked at shear design</i>).</p> <p><small>*precise or vague/loose</small></p> <p><small>**except in the ELF Malaysian subcorpus</small></p> <p>Also: prompt-explanation patterns (e.g., <i>If you remember, I described it – it's effectively a new material.</i>).</p>	<p>Fronting of non-specific chronological references to information delivery (e.g., <i>I said earlier that...</i>).</p> <p>Causal logical connector (mainly conjunction <i>so</i>) frequently preceding the pattern: conjunction + anaphoric demonstrative reference to topic (or vice-versa) (e.g., <i>So, that is how we measure things on site.</i>)</p> <p>Temporal deixis often not specified at all: focus is on reiteration of information or facts.</p>	<p>Fronting of important information with simple pseudo-clefts (e.g., <i>What we are going to move on is look at a more powerful technique called the method of sections.</i>)</p> <p>Patterns: pronoun + modal auxiliary (e.g., <i>I will do a demonstration</i>) is more common than: pronoun + semi-modal (e.g., <i>I'm going to do a demonstration</i>), contrary to PreFUTs.</p>	<p>Common pattern: temporal deixis + pronoun + auxiliary verb + lexical verb (e.g., <i>In your third year, you're going to start dealing with fatigue.</i>)</p> <p>Patterns: pronoun + semi-modal (e.g., <i>I'm going to do a demonstration</i>) is more common than: pronoun + modal auxiliary (e.g., <i>I will do a demonstration</i>), contrary to PreCURs.</p>

Note. Adapted from Alsop (2015). The Engineering Lecture Corpus (ELC, Alsop, 2015; Nesi et al., 2014)

Appendix F. Summarizing attributes in the ELC: exclusive lexicogrammatical features (cont.)

	Review previous (RevPRE)	Review current (RevCUR)	Preview current (PreCUR)	Preview future (PreFUT)
Salient pronouns	<i>You</i> is the most common pronoun (pmw ^a).		<i>We</i> is the most common pronoun (pmw ^a).	
	<i>We</i> is also salient and predominantly inclusive (when discussing work that has already been done). <i>We</i> accompanies and softens a negative evaluation.		<i>We</i> is almost equally frequent in PreCUR and PreFUT, which is over 3 times higher than in non-summary discourse. Lecturers much favor the inclusive <i>we</i> , said with an inclusive tone. This contributes to <i>togetherness</i> : students and lecturer will <i>figure out</i> , <i>examine</i> then <i>determine</i> the <i>how</i> of the engineering problem <i>together</i> , often through an <i>example</i> .	
Other salient lexicogrammatical features	Two most salient tokens: <i>last + week</i> (= bigram in 1/3 of instances – temporal deixis). <i>Remember</i> is salient (relevance marker #1 by far) + <i>if you– / do you–</i> (3-grams). Often contain negative evaluation (e.g., <i>a bit more difficult</i>).	Use of minimizing language and down-toning strategies (e.g., <i>a little bit, just</i>): minimizes imposition on students when faced with a difficult concept.		<i>We</i> and <i>'ll</i> rank as the 2nd and 3rd most common key words.
		<i>Earlier</i> is the key-most token (mark of distance from delivery). <i>Mentioned</i> is salient (e.g., <i>I just mentioned, as I mentioned just now</i>) Less tense and aspect marking in the ELF divergent (Malaysia) subcorpus.	Correlation between levels of difficulty and informality. High frequency of 4-grams, especially <i>we're going to</i> and <i>I'm going to</i> when outcome is certain. High keyness also of <i>look, today, will</i> and <i>going</i> . Frequent use of <i>later</i> when current content is not delivered immediately in lesson outlines. Specific softening-minimizing language: imposition of daunting concepts mitigated or offset through lexis and structure: <i>a little problem comes next; that should be looked at carefully; the crux of what is actually being measured</i> .	Markers of temporality are salient, with typical key words: <i>next, week, tomorrow, year, semester</i> and weekdays. <i>Later</i> is a key word, but not instrumental to understanding the <i>how</i> of the content (unlike PreCURs): . Few hedge words* or softeners occur, as processes of knowledge building are presented as simple and incremental. * <i>hedge words</i> (or <i>downtoners</i>) are a quick way to soften the impact of a statement, e.g., <i>almost, rather, barely, hardly</i> , etc. (Peters, 2004, p. 259)
Misc. other features	Lecturer's tone: mostly reassuring or gentle [identified by using the multimedia files]	Distance from original delivery to review tends to be short. Can mark a shift in lecturer's tone. Often no <i>temporal deixis</i> at all: focus is on reiteration of information or facts.	Distance to the delivery of PreCUR is commonly identified (e.g., <i>move onto now, last five minutes</i>). Difference between <i>will</i> and <i>(be) going to</i> indicates the speaker's orientation to the future event. (Peters, 2004, p. 494)	Processes of knowledge building are presented as simple and incremental (e.g., <i>the theory builds on very easily from one lecture to the next</i>).

Note. Lexicogrammatical features after Halliday & Matthiessen (2014) Other salient lexicogrammatical features and discursive features adapted from Alsop (2015). The Engineering Lecture Corpus (ELC, Alsop, 2015; Nesi et al., 2014)

a. pmw: per million words

Appendix G. Example of a RevPRE summary in initial position

Appendix G. Example of a RevPRE summary in initial position, in a lecture on electro-magnetic compatibility. Duration: 3.9 minutes (460 tokens). RevPRE type markers are italicized. Discourse markers are boxed. Personal pronouns are highlighted. Relevance markers are in bold.

[L10 - 20201023_SAJ_EMC1_SUM_0010]

5
10
15
20
25
30

ok. (2) **so**- **basically** *last- er last week* **we** started with some introduction er on er how the (2) er electrical equipment can be disturbed by others. *it was* **the core of the definition** of the electromagnetic compatibility. **so**- *we reviewed* er: some coupling modes which were divided into five groups. the first coupling mode is a- well- er part of a circuit. **you** find an impedance- which is common between er several subsystems. (2) er: **so**- here for instance- **you** have an impedance- each time the system number one er: is er (2) generating some current there is a voltage drop here- (i figure there's) a voltage drop here- of course- but also on this one. (3) er: the second group is er diaphony. **so**- capacitive diaphony is a (stray) capacitance between two lines and this generate an induced voltage on a victim line. to also investigate the shielding effect and- as **you may remember**- it is not sufficient to insert a metal layer between the two wires- because this just divide or split the capacitances into two capacitors- but each of them is higher than the initial one. therefore- **you** don't change anything in fact. **so**- to be efficient **your** shield has to facilitate evacuation of current to a stable potential. it is either the groove or the d-c plus- if **you** want. (4) er another diaphony is er inductive diaphony- which is the magnetic coupling between two wires. **so**- one wire is generating a h field and this h field is generating a (xx) and induced voltage on the other line. (3) again- er shielding er: must be either avoiding- in- in- in all cases a shield must avoid the magnetic field generated by the disturbance line to be coupled to the victim line. **so**- either it is a magnetic shield- in this case that's **ok**- it's (xxx). but if it's just a conduction layer then- er to cancel the incoming field- **we** must er facilitate the circulation of currents inside the tube. **so**- either **you** are in frequency- in high frequency range and **you** play with any currents or **you** facilitate the circulation of currents through two connections er in order to allow the current flowing inside the shield. (3) and the two last er coupling modes are er related to higher frequency **so**- **you** have a magnetic field arising er: in the vicinity of a system which defines here a coupling loop and the magnetic field generated induce voltage in the loop- or er the electric field generates some er induced voltage on the line. (4) **ok** **so**- that was the first point where i tried to illustrate this with the measurement issue- but er i will not detail more than this. (4) er:: then i think **we** also started to define the standards.

Appendix H. Longest RevPRE summary in CoFEL

Appendix H. Example of a long RevPRE summary (the longest in CoFEL: 873 tokens, 425 seconds). Typical RevPRE verbal phrases are *italicized*. Discourse markers are so. Personal pronouns are i.

[L10 – 20201023_SAJ_EMC1_SUM_0040]

(2) so- this is- er the low frequency: standards in which the harmonic of the currents are limited. er: so- that's the european standard one thousand dot three dot two (2) and as *we discussed*- i *remember*- er it is not depending on the power- on the amplitude of the fundamental. it's not- the harmonic limitation is er provided in absolute values- not in relative values and that's normal from the grid point of view- because what is important for the grid is to reduce the voltage drop on the impedance here. so- it is a harmonic of the current circulating in the grid which are under concern and i don't care about your- er your fundamental. of course- er as i mentioned before- this standard apply for devices which power is greater than seventy-five watts. that's the trade-offs that have been found when- er when defining the standard. in higher frequency range *we* monitor the line to ground voltage and compare it to the standard (3) er (4) and for this *we* have to really define er the: (2) impedance between the electrical equipment and the ground as well as the s- impedance between the source and the ground as well as the impedance between the source and the electrical equipment. *you* could (x) ground currents which are originating this er line to ground voltage which is er under concern for the standard. er this ground currents are of course impacted by this line (of) impedance- this stray capacitance- this stray capacitance. so- if *you* change the distance- for instance- from the source to the electrical equipment- or if *you* change the distance from the source to the ground- or if *you* change er the distance from the equipment to the ground- *you* will change the result. and of course- *you* cannot afford having results depending on the way *you* are er making them in a standard. reproducibility is a- is a key point. so- therefore *we* insert between the source and the device and the test a- a specific network which force the current- the ground current- and also the line current flowing inside this equipment- this measuring equipment and not allow it to go in other ways. (3) so- basically this is called the line impedance stabilization network. er so- this is the most complete one and *we make* some comments in this one. so- here is the source side and here is the device on the test side. *you* can see here that- if *you* have a ground current- he will not flow through this path because this is a high frequency- er high frequency- this is a high impedance in the high frequency range. so- the line current er or the ground current will flow to this impedance here. and this impedance is a specific and defined by the standard as er varying from 5 ohm to 50 ohm. (4) so- this is the full l-i-s-n^a in a simplifi- -fied mode. (2) so- *you* can see here (3) that the disturbance generated from the circuits and the test- if they are between lines they will flow this path- therefore it is er crossing the calibrated impedances z-n (2) and if it's ground currents- (4) the ground current coming here will split into two different paths- one here and one here but still crossing the impedance here. there is no reason for ground current for flowing here because here it's a high frequency- a high value impedance. same for here (3) and of course er: the same for the- the current between lines. (3) so- this is the good point of the l-i-s-n which is really here to er focus the current in er given impedance path and ensure reproducibility of the- (2) of the- of the measurement. the value of the impedance is er: not- er: he has not been defined randomly. the five ohm in low frequency range corresponds to existing impedance

45 and the fifty ohm also is a kind of trade-off in the high frequency between the actual
impedance of the grid and the fact that high frequency impedance measurement
systems- like spectrum analyzer have input impedance of 50 ohms. (5) then er for
high frequency you use antenna and- and that. er: then we started to define the units
which are used er electromagnetic compatibility. so- er since the disturbance is-
50 which are generated by the switching devices or electrical equipments- are very low
amplitude we use er micro- er we- we use reference to microvolts or microamp and
to express this quantity in er deci- decibel. (2) so- the:- the voltage expressed in d-b
micro volts is 20 times the logarithm of the voltage divided by one micro volts. the
same for d-b microamp. careful with the power which is er more a telecom unit than
an e-m-c unit. er: it's 10 times logarithm of p divided by one milliwatt. (3) so- since
55 input impedance of an e-m-c measurement system is 50 ohm- that's a standard- er:
there is some links between voltage and current- and voltage and power- and current
and power- and you can express this. (3) so- we also er had some reflection about
er: (5) attenuator- sorry- and we use the ten d-b attenuator and you understood what
(that's) here. and i think we stopped here.

a. A line impedance stabilization network (LISN) is a device used in conducted and radiated radio-frequency emission and susceptibility tests, as specified in various EMC test standards.

Appendix I. Longest PreCUR summary in CoFEL

Appendix I. Example of a long PreCUR summary, the longest in CoFEL. Summary parts appear in square brackets (e.g., [1]). PreCUR type markers are *italicized*. Discourse markers are boxed. Personal pronouns are highlighted. Relevance markers are in **bold**.

[20201016_SAJ_EMC1_SUM_0020]

[1] so- *the lecture of today is focused on* electromagnetic compatibility. it's part of the module design models for power electronics so it's er a new part of the power electronics design. [2] so- you started with er passive components design or choice. er you started already with er doctor (xxx) [a colleague] as er computation of losses and er how to manage the- the heat dissipation. [3] *now we are working* on the electromagnetic compatibility which is er **something which is quite important** in power electronics. er it is really er the consequence of switching er which is used in power electronics and i will try to- to help you in understanding er the basis of e-m-c electromagnetic compatibility which is e-m-c. in french it is c-e-m. [4] so- after the basis which will be normally *today* but if it's too long especially because er you are at the end of the day and you may be tired. if it's too long we will go on next week. er but normally we should also be able to know how to model er a- (xxx) matter in order to provide the e-m-c analysis. [5] and finally- **the main goal of this lecture** is to help you in er understanding how to design an e-m-c filter **which is mandatory regarding the standards** which are- er which have been er er which have been created to ma- to mitigate this disturbance that you can see er from the- (xxx). and finally- because it's high frequency phenomena the filter are composed of passive elements but they are excited in the megahertz range and maybe er ten megahertz range. er **we need to know** the high frequency behavior of this kind of components and also the (cabling). [6] so- *that will be the-* **the main points that will be addressed**. [7] er we have er four lectures and also two different labworks. [8] so- normally *today* we'll finish the introduction but as i told you it may be a bit too long. so- maybe be we'll manage to (xxx) in another way [9] and then we will go with disturbance sources and how to er- and how to er to er simulate er in an efficient way this disturbances and how to design the filter and how to address the high-frequency behavior of the components.

Appendix J. Longest RevCUR summary

Appendix J. Example of a long RevCUR summary (the second longest summary in CoFEL: 4.75 min, 503 tokens). Type markers are *italicized*. Personal pronouns are **highlighted**. DMs are **boxed**. Other elements of interest are underlined or in **bold**. The numbers in brackets indicate pauses, in seconds. Letters or words in brackets indicate an uncertain transcription. Curly brackets {...} contain transcriber's comments.

[L01 - 20200117_FAC_AERO_SUM_1130]

all right. **so**- *actually*:- *it's not very clear probably the way i've- i've explained it* but the flutter phenomenon is hidden here. {lecturer points at one of the equations on the chalkboard} **so**- it **basically** says that *we have studied* the case delta posi- is positive at first and *we have seen* that *we* have either er just er (a) harmonic solution. **so**- pure imaginary roots- or *we* have another case where *we* do not have vibration and *we* have just either an amplification or attenuation of the- the- the motion but without vibration. and here {lecturer points at another equation on the chalkboard} *then we have considered* the case delta negative in which *we* have complex roots. and in this case- if *we* have the positive er real part of the root- then *we* have an amplification of the vibration which corresponds to the flutter phenomenon. and here *what we have said* is that delta is equal to zero is actually the transition between the two cases. there is no risk of flutter at first. then for delta positive- delta equals zero is the transition. and then delta is negative corresponds to the flutter case. (2) **all right?** {lecturer points at another equation on the chalkboard} and *what we have said here* is that the case delta equals zero corresponds to a positive value of b and at which will be corresponding a given value of the dynamic pressure that *we* will call q-f and this q-f will be the dynamic pressure at the flutter er er point. **right?** **so**- it's similar to the: static aeroelasticity. *we had* a divergence for the specific value of the velocity and here *we* can define a: flutter velocity that corresponds to the: velocity beyond which *we* will have a risk of flutter. **ok?** and if *we*: look as well at the solutions *we have seen* that *we* start with delta positive to real roots and then *we* have the transition at delta equals zero which correspond(s) to any one root and that **basically** mean(s) that the: phenomenon at the- the beginning will have two different frequencies- omega one and omega two- and they will step by step go closer one to the other- and the divergence phenomenon corresponds to the windspeed at which those two frequencies will merge and there will be only one frequency corresponding to the flutter phenomenon. *if you remember the explanation that we have given at the beginning of this part*- dynamic aeroelasticity- *we have seen* that if the two motion-torsion and the flexion- they operate at the same frequency and in a given configuration in phase then *we* will have an implication of the phenomenon. **so**- *that's basically what is said in this*. **so**:- (oops) (2) {chalk slips from lecturer's hands} if *we* have the velocity here and value of omega- *we'll* have one (2) frequency that will change with the wind speed and the other one as well. and when they merge- (3) *we* have the flutter speed. (7) *we* can say that- (35) {lecturer resumes writing formulas on chalkboard} **all right**. **so**- *we* have the two natural frequencies that merge with the increasing wind speed and *we* have the flutter point which is here. (it) corresponds to this point. (4)

Appendix K. Example of a chain of four summaries combining two types

Appendix K. One instance of a chain of four summaries combining two types: RevCUR + PreFUT + RevCUR + PreFUT. *Italics* emphasize type markers. Double vertical bars indicate the labelled type boundaries. Numbers in brackets indicate pauses in seconds.

[L07 – 20200820_PAC_SPI_0180-0182-0184-0190]

||RevCUR|| ok so- er: *we already discussed* about betweenness centrality. *er let's jump on the take-back-home information.* (3) so- er please *remember that* er: studying a- er: (3) at least two infrastructure together- being able to separate their services into two different worlds- so- it's already complicated to solve. and er:: *it's-er what we did propose* is one way to solve risk analysis for complex and independent- interdependent infrastructures (2) er which usually are er studied by different experts (2) and- er:: *as i present(ed)* you- this is a way to find some er er (1) critical interdependence (1) and typically the ones you need to secure at first to be- er: to be able to have the more resilient system of system- or smart energy system. (4) so- if i jump now in- into a- the- conclusion and- and further work before the evaluation of this presentation. so- let me summarize what- what we- what we discussed today. so- there is really a need of original and innovative methodologies to be able to study the interdependencies and the vulnerabilities of interconnected systems. (2) i was presenting you- but it's not the only one- complex networks- which is promising to study the topological weaknesses of large interconnected systems. ||PreFUT|| the proce- the proposition is flexible and *you will see in- er in further work* how we can extend it. it's flexible and should be able to be used to study other infrastructure- different (1) er to i-c-t^a and power systems- but you can add- for sure- can add gas- water- recycling- we discuss already- and so on. (2) ||RevCUR|| and er: *i was- i was pushing toward you* the use of hermitian matrix and eigenspectral analysis to study the structure of coupled infrastructures. (3) ||PreFUT|| so- *further work. where are we going?* of course we need to validate the models and so- er: at first to validate topological approach we can think about combined simulator. so- having simulators able to discuss with f-m-u- (f-m-e)- f-m-i- for instance- er process. make sure that what we are assessing as most critical nodes are- with a simulation approach- indeed- very important. we can- of course- rely on what we create in predis {a local experimental platform}. so- reduce scale networks in which you have representation of both (1) power system and communication. is a predis center that is hosted in my lab- in my school. (2) er: we want to incorporate dynamic aspect. so- dynamic- not as oscillation that the colleagues tranquil control system are- er: are- are doing- but er other type of dynamic aspects. as i have a er power system problem that will lead to blackouts. all these different er: trajectory can happen. so- dynamic aspects. the same as we are er er typically analyzing in a pandemia er:: situation- as- as the covid we are now having. we need to have cross-disciplinary approach- of course- and so- we need to incorporate people from control system- markets- human behavior- to have this extended towards smart cities. we can extend the- the modeling er based on smart grid architecture modeling. this smart grid architecture modeling is proposing different layers. so- three-dimension er way of just stu- studying electric- electrical components and i-c-t^a components- but you can as well extend this to multi-infrastructure. (3) and so- there is a real potential in this area of research. (3) er: so- *my- my presentation is over.*

a. ICT: Information and Communications Technology.

