

ALESSANDRA FAZIO - MICHELA MENGhini -
VALERIA TAMBURINI¹

EXPLORING EULCS SELF-(RE)PRESENTATIONS: WEBSITE STRUCTURES AND LEXICAL CONTENT

ABSTRACT:

This article shows the results of the first phase of the research carried out by the Foro Italico Research Unit (RU) within the project “European Language Centres as a multilingual community of practice: A multimodal discourse analysis of academic, cultural and social growth conveyed through the language of websites”. The aim of our research is twofold: first, to identify similarities and differences in the structures of the European University Language Centres (EULC) websites; second, to analyse the main linguistic features conveyed through them, with a focus on lexis. Three linguistic corpora related to each geographical language area studied, Italian, English and Spanish, have been developed in collaboration with the partner RUs to examine through a corpus-driven quantitative and qualitative analysis the differences in

¹ Author contribution: Alessandra Fazio, University of Rome “Foro Italico”; alessandra.fazio@uniroma4.it: 1. Introduction, 2. Literature review, 5. Conclusions and future steps. Michela Menghini, University of Rome “Foro Italico”; michela.menghini@uniroma4.it: 3.3 Corpus-driven quantitative and qualitative analysis, 4.2.1 Analysis of English language corpora, 4.2.2 Analysis of Italian language websites. Valeria Tamburini, University of Rome “Foro Italico”; valeria.tamburini@uniroma4.it: 3.1 Collecting website information, 3.2 Creating the corpora, 4.1 Structure and general content, 4.2.3 Analysis of Spanish language websites.

lexical content among the EULC websites.

ABSTRACT:

In questo studio si riportano i risultati della prima fase del lavoro svolto dall'unità di ricerca (UR) dell'Università degli Studi di Roma "Foro Italico" nell'ambito del progetto "European Language Centres as a multilingual community of practice: A multimodal discourse analysis of academic, cultural and social growth conveyed through the language of websites". Lo studio si propone, da un lato, di identificare le somiglianze e le differenze nelle strutture dei siti web dei Centri Linguistici Universitari (CLA) europei e, dall'altro, di analizzare le caratteristiche lessicali più rilevanti che emergono dall'analisi dei siti. A tal fine, sono stati creati, in collaborazione con le UR partner, tre corpora relativi alle aree linguistiche esaminate (italiano, inglese e spagnolo). L'analisi quantitativa e qualitativa dei corpora ha permesso di rilevare le specificità lessicali dei siti web dei CLA.

1. INTRODUCTION

The aim of this article is to report the outcomes of the initial phase of the research project "European Language Centres as a multilingual community of practice: A multimodal discourse analysis of academic, cultural and social growth conveyed through the language of websites"² carried out by the Foro Italico Research Unit (RU). The research objective, at this stage, was to identify similarities and differences in the structures of European University Language Centres (EULC) websites and to analyse their main linguistic features, focusing on lexical content.

Section 2. provides a review of the limited yet existing literature on institutional websites as multimodal communicative tools in higher education. The role of Language Centres (LCs) in promoting multilingualism and intercultural communication is highlighted, although more research on LCs' digital representation, which is not much investigated, is required.

Section 3. focuses on the methods used to collect website information, create corpora from the relevant EULC websites sections (in Italian, English and Spanish) to conduct a corpus-driven quantitative and qualitative analysis of their lexical content. The discussion of the results in section 4. is divided into two parts. The first part, subsection 4.1, examines the structure and general content of the EULC websites,

² This research project was conducted as part of the activities of PRIN (Projects of Relevant National Interest) 2022. The authors gratefully acknowledge the Italian Ministry of University and Research (MUR) for funding this study.

focusing on some key aspects, namely website types, original languages, translated versions and main sections. The second part, subsection 4.2, presents results related to the lexical features of the websites. An in-depth analysis of the English language corpora is shown, as well as a comparative analysis of the Italian and the Spanish language corpora with the English ones.

Section 5. summarises the findings and suggests follow-up developments for further research.

2. LITERATURE REVIEW

Regarding the existing research in this field, studies on institutional websites as multimodal communicative tools have gained significant attention in discourse analysis, particularly in higher education. This can be seen in the well-known works *Multimodal Discourse: The Modes and Media of Contemporary Communication* by Kress and van Leeuwen (2001) and *Analysing Discourse: Textual Analysis for Social Research* by Fairclough (2003). LCs, as key academic services promoting multilingualism, have been examined for their role in intercultural communication across disciplines, internationalisation, and as part of communication strategies aimed at enhancing universities' communication and global visibility (Triki 2022). However, the digital representation of LCs is an area that has not been investigated much. While previous research on university websites has largely focused on genre analysis of website sections such as the 'About' section (Caiazzo 2014) or marketing strategies (Chapleo 2015), limited attention has been given to their linguistic and discursive features.

Corpus-based approaches have proven to be effective for analysing institutional discourse, particularly to identify lexical patterns across different languages. This is especially relevant in comparative studies on EULC websites, where very few studies have applied a corpus-driven multimodal analysis to examine how LCs construct their identity online. Our study agrees with Tognini Bonelli's (2001: 17) assertion that "the corpus-driven approach builds up the theory step by step in the presence of the evidence". This methodological approach has been fundamental to shed light on the structural and lexical features of EULC websites across Italian, English, and Spanish language domains. In *Lexis in Contrast* (2008: 73-96), Tognini Bonelli emphasises the importance of comparative corpus approaches, illustrating how functionally complete units of meaning can be analysed across English and Italian.

Bondi (2010: 3) provides a thorough definition of corpus linguistics, stating that "in a quantitative perspective, keywords are those whose frequency (or infrequency) in a text or a corpus is statistically significant when compared to the standards set by a reference corpus." This perspective is highly relevant to our study as it examines multilingual corpora and their conceptual links to technology in language learning as well as LCs self-representations in comparison with more general reference corpora.

Finally, the paper shows many correlations and examples of key concepts from the multilingual corpora around the semantic area of technology in language learning where keywords in the field of blended learning through e-learning platforms rank high scores. This confirms the role of technology in language education as highlighted by Bax (2011) and Bonk (2012) and how it becomes an effective element in the language learning process without us being consciously aware of its crucial role and its impact on language learning.

This study aligns with recent trends in digital discourse studies such as the *British Council* BC report *The Future Demand for English in Europe: 2025 and Beyond*. The BC report addresses the need to enhance research on institutional multilingualism in higher education. In this respect, the findings contribute to the understanding of how EULCs position themselves through digital discourse, conceptualised in Wenger's terms as part of a community of practice (1999). This is also in line with the European Council Recommendation (2019) to “encourage research in and use of innovative, inclusive and multilingual pedagogies, including for example the use of digital tools, intercomprehension and ways to teach subjects through a foreign language (Content and Language Integrated Learning) and innovate initial teacher education.” Furthermore, the CEFR Companion Volume (2020: 28) emphasises the importance of both multilingualism and plurilingualism highlighting the difference between the former as the “coexistence of different languages at the social or individual level” and the latter as the “linguistic repertoire of an individual user/learner”. Both concepts are crucial and integral to the PRIN project, although they are not the main focus of this article, and are thus discussed in a marginal way.

3. METHODS

A three-step methodology has been used in this study: first, information collection to identify Italian, UK/Irish and Spanish websites' similarities and differences; second, corpora creation for each language area; third, quantitative and qualitative comparative analysis of lexical features between the corpora.

3.1. Collecting website information

To collect consistent data, an online form was designed by the University of Calabria RU and filled in for each website in collaboration with the other RUs. Four relevant key aspects were detected: website type, original language, translated versions and main sections. The University of Rome “Foro Italico” RU collected and analysed data from the Spanish websites to compare them with data from the Italian and UK/Irish websites collected by the other RUs.

Data on website type, original language, and translated versions were collected from all ACLES (Asociación de Centros de Lenguas en la Enseñanza Superior) members' websites ($N = 64$), while information related to main sections was collected only from

websites originally in Spanish and from those with an available Spanish version ($N = 62$). For the comparison of website type, original language and translated versions, all websites of the LCs that are members of AULC (Association of University Language Communities - UK / ÉIRE area) ($N = 72$) and AICLU (Associazione Italiana Centri Linguistici Universitari) ($N = 51$) were considered. However, a few websites are no longer available online at the moment of writing³, therefore only 50 Italian and 69 UK/Irish websites were considered for the comparison.

3.2. *Creating the corpora*

The synergy with the project RUs was crucial to create three main corpora, one for each language area (AICLU Italian, AULC English, ACLES Spanish). Textual content was extracted from key sections/pages (*e.g.*, Home, About, Courses) for each corpus.

For an in-depth cross-language analysis, two additional corpora were also created from the English versions of the Italian websites (AICLU English) and the Spanish websites (ACLES English). The “Foro Italico” RU manually extracted textual content from the UK/Irish websites and from the Spanish websites, including their English versions where available.

3.3. *Corpus-driven quantitative and qualitative analysis*

A corpus-driven quantitative and qualitative analysis was carried out on the lexical features of the Italian ($N = 50$), UK/Irish ($N = 69$), and Spanish ($N = 62$) websites. For the quantitative analysis, keywords (*i.e.*, single-word items) and terms (*i.e.*, multi-word items) were extracted to identify typical lexical units for each corpus, using the textual analysis tool Sketch Engine (Kilgarrieff *et al.* 2004, 2014). The English Web 2021 (enTenTen21), Italian Web 2020 (itTenTen20) and Spanish Web 2023 (esTenTen23) were used as reference corpora for English, Italian and Spanish language corpora, respectively. The score, which measures the statistical significance of the relative frequency of each keyword in the focus corpus compared to the reference corpus, was primarily considered for comparisons and analysis.

For the qualitative analysis, content-relevant items were selected from the most meaningful extracted keywords (at least the first 200 items) and analysed to identify patterns and language context of use. The following categories, related to the characteristics, services, and activities of LCs, were identified as the most recurring and relevant in the selected keywords and terms⁴ for the purposes of our project: ‘Certifications and language levels’, ‘Languages’, ‘Learning Management Systems (LMS)’, ‘Course organisation and procedures’, ‘Teaching activities/approach and

³ All websites were last accessed on 31/07/2024.

⁴ In this study, relevant single-word and multi-words items were assigned to categories and then analysed together as part of each category, with a focus on their semantic domain rather than on their word number.

assessment’.

Keywords and terms from the UK/Irish corpus, divided into the above-mentioned categories, were compared with those from the English versions of the Spanish websites ($N = 39$) and the corpus of the English versions of the Italian websites ($N = 24$). Scores were cross-checked across the three corpora when possible. The Italian and Spanish language corpora were analysed separately and compared only in broad terms.

4. RESULTS AND DISCUSSION

Results are discussed in two parts: 1) subsection 4.1 focuses on the structure and general content of the EULCs websites; 2) subsection 4.2 illustrates their main lexical features.

4.1. *Structure and general content*

Three types of websites emerged from the analysis: independent, embedded and departmental. As the data show (see Table 1), most UK/Irish and Spanish websites are embedded. In contrast, most Italian websites are independent. Furthermore, a fourth of the UK/Irish websites is part of a department website (*e.g.*, Department of Languages and Cultures).

Website type	AICLU	AULC	ACLES
Independent	61%	11%	27%
Embedded	37%	63%	73%
Departmental	2%	26%	0%

tab. 1. *Website type*

As far as the original language is concerned, all Italian websites are in Italian, and most UK/Irish websites are in English (only 1% is in Irish). Spanish websites show greater variation. The data in fact show that, while most of them are in Spanish, 19% are in Catalan and 3% in Galician.

Some websites are translated into one or more languages. The Spanish area has the highest number of translated versions ($N = 39$), followed by the Italian area ($N = 27$). In the UK/Irish area, only a few websites ($N = 7$) are translated. As shown in Table 2, the Spanish area offers a wider variety of languages for translation (Basque, Catalan, Chinese, English, French, Galician, Spanish) compared to the Italian area (Chinese, English, French, German) and the UK/Irish area (English, Irish, Welsh), although this may be due to the high number of versions in local languages. Spanish translations are provided for Spanish websites when the original language is Catalan or Galician. Similarly, English translations are provided for UK/Irish websites when

the original language is Irish.

Translated versions	AICLU	AULC	ACLES
Basque	0%	0%	3%
Catalan	0%	0%	9%
Chinese	2%	6%	3%
English	53%	1%	72%
French	2%	0%	3%
Galician	0%	0%	2%
German	2%	0%	0%
Irish	0%	1%	0%
Spanish	0%	0%	22%
Welsh	0%	4%	0%

tab. 2. *Translated versions*

English is the most used language for the translation of the Italian and Spanish websites. Focusing on the Spanish area, a distinction must be made between the websites that have an English version and those that offer a partial translation: 46% of Spanish websites are only partially translated into English, while 39% offer a full English version. Typically, on partially translated websites, only pages or paragraphs relevant to international students are translated.

Besides translations, to enhance accessibility, some EULCs offer features such as text-to-speech recognition, larger text sizes, and colour adjustments. However, probably due to technical problems, translation and other accessibility options may not function correctly. For instance, some websites allow users to switch to English, but the translation is not available.

As far as the structure of the EULC websites is concerned, we identified six main sections, that is, pages that are relevant and easy to identify because they are linked as items in the website menu. In some cases, clearly recognisable page segments were also treated as sections. As the data show (see Figure 1), most Italian, UK/Irish and Spanish websites have an 'About Us' section, which can be either a separate page or a paragraph on the homepage. In contrast, few have a 'Strategic Plan' section describing their mission. This section is rarely a separate page, but rather a paragraph in the 'About Us' section. A high percentage of Italian and UK/Irish websites have both a 'Courses' and a 'Languages' section. In the former, courses are usually presented by type (*e.g.*, pre-sessional or in-sessional), while in the latter they are listed by language. Unlike the Italian and UK/Irish websites, most Spanish websites have a 'Courses' section, which usually includes information about the type of courses and

the languages offered, but not a ‘Languages’ section. On the other hand, most Spanish websites have a section dedicated to assessment. This section, which usually refers to certification exams, is available on less than half of the Italian websites, where information on language certification is provided in a separate ‘Certification’ section. Although many Italian websites do not have an ‘Assessment’ section, 74% devote a section or part of a section to the placement test or at least mention it. As for the UK/Irish websites, only a few have an ‘Assessment’ section. Finally, some Spanish websites include a dedicated section on ‘Translation and mediation services’, while only about one-fifth of the Italian websites and a small number of UK/Irish websites offer such a section.

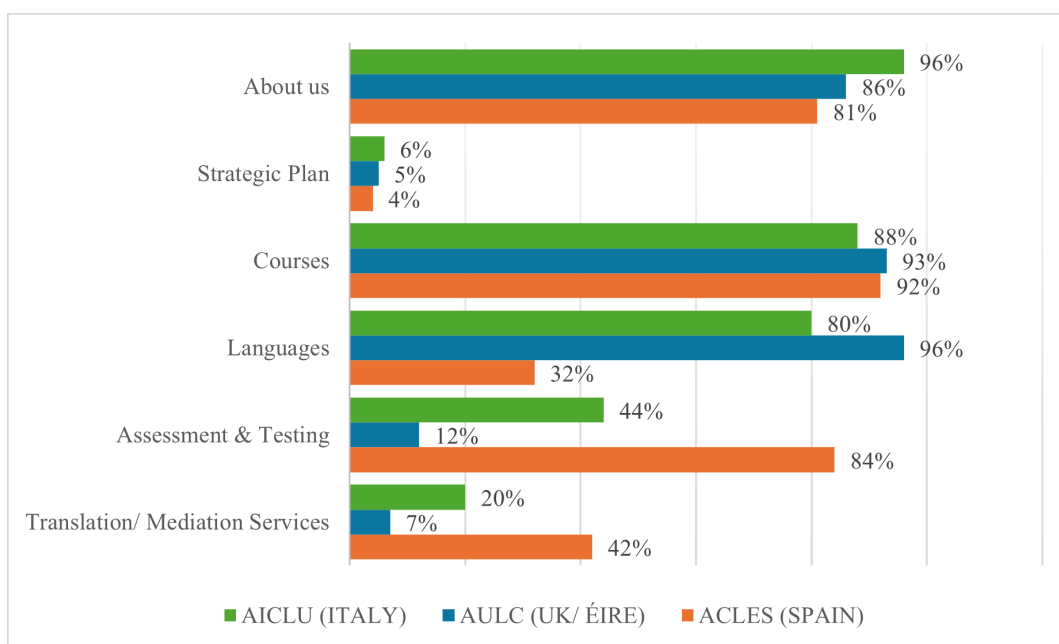


fig. 1. *Main sections*

As noted above, EULC websites provide information on language certifications, which can be either internal (i.e., issued by the language centre itself) or external (i.e., issued by specific internationally recognised organisations). As the data show (see Figure 2), most Spanish websites refer to internal and external certifications. Less than half of the Italian websites, instead, offer an internal certification, while a significant number mention external certifications. As for the UK/Irish websites, more than half offer internal certifications, while only a few mention external certifications. An average of 40% of the LCs that mention external certification on their websites are accredited international examination centres (AICLU = 46%; AULC = 40%; ACLES = 38%). These results seem to indicate that UK/Irish LCs are less focused on internationalisation and global mobility than their Italian and Spanish counterparts, as internationally recognised certifications are mentioned

significantly less often in their LCs than in the Italian and Spanish ones. While it is possible that students in the UK and Ireland have less need for international English certifications, an unexpected outcome is the comparatively limited offer of international certifications in other languages. This is confirmed by the data on the sections dedicated to internationalisation, as illustrated in the following paragraph.

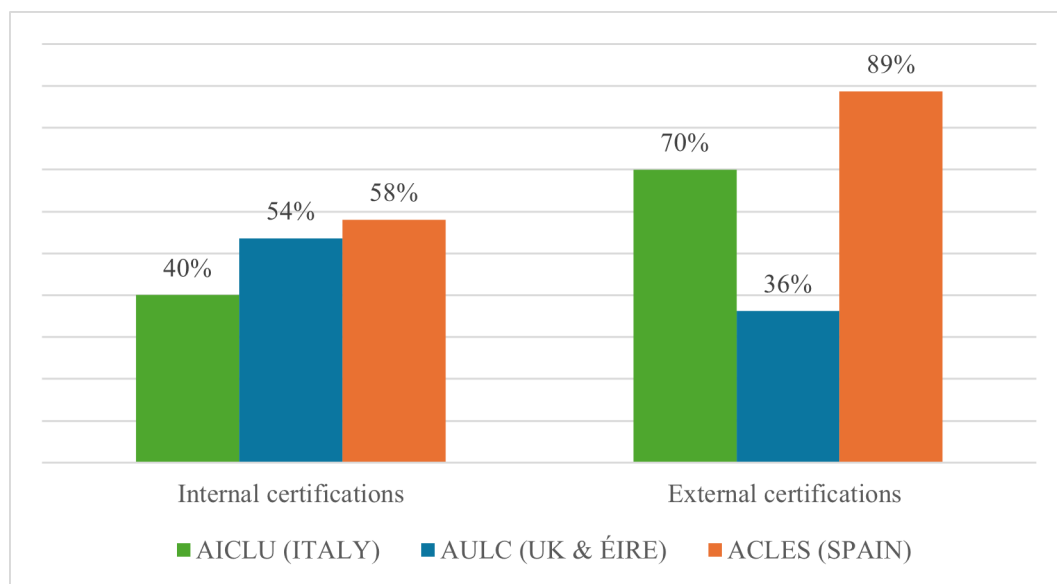


fig. 2. *Language certifications*

The analysis of the structure and general content of the EULC websites included an examination of some of the concepts relevant to the project, such as internationalisation and lifelong learning, to see whether and to what extent they are represented on the Italian, UK/Irish and Spanish websites. Interestingly, 86% of the Italian and 65% of the Spanish websites refer to international mobility programmes and students, whereas only 25% of the UK/Irish websites do. This might be expected, because the UK is no longer part of the EU and they do not participate in ERASMUS programmes. However, ‘Internationalisation’ is rarely a separate section, nor is ‘Lifelong learning’. The latter is mentioned by only a few EULC websites (AICLU = 12%; AULC = 13%; ACLES = 5%).

4.2. *Lexical features*

4.2.1. *Analysis of English language corpora*

An analysis of lexical features was first carried out for the corpora in English, focusing on their key lexical items. As a premise, it is important to keep in mind that English is the original language for most of the websites of UK/Irish language centres (LCs members of the AULC), while for the websites of the LCs members of the ACLES and the AICLU the English versions are either translations or separate versions (see

section 4.1. and Table 2 above) of the original Spanish and Italian websites. Such English versions are typically shorter than their Italian and Spanish counterparts, and are primarily aimed at international students and visiting scholars. Moreover, the Spanish and Italian LCs in general seem to focus on internationalisation and mobility programmes much more than the UK/Irish ones, as mentioned in the previous section. Such primary differences between the corpora inevitably affect their contents in general and their specific lexical features.

The first category identified in the qualitative analysis of the English corpora includes key items related to ‘Certifications and language levels’ (see Figure 3⁵). The *Common European Framework of Reference* (CEFR) and its levels (*A1*, *A2*, *B1*, etc.), along with a few international language certifications (*IELTS*, etc.), are frequently mentioned. It is to be expected that European LCs often refer to CEFR and its levels on their websites, making these lexical items emerge as keywords. However, Italian and Spanish LCs mention CEFR levels more often than UK and ÉIRE centres, and they feature more international certifications (*DELE*, *DELF*, etc.) among their keywords. Italian centres also frequently use terms like “*language certificate*” and “*language level*”. This emphasis on certifications and levels likely reflects the fact that their primary audience consists of international and mobility students, who need information on language requirements for university access. UK/Irish centres dedicate less space to certifications on their websites, even though English certifications are also required for incoming students to their universities.

The second category of lexical features includes the languages emerging as keywords, as well as the actual word *language* (relevant in all three corpora) and other connected terms such as *target language* (a top keyword only for the AULC corpus) and *foreign language* (see figure 4). The AULC corpus features a wider variety of languages among their keywords, as the UK/Irish LCs seem to describe their offer of courses in different languages more than the English version of the Spanish and Italian websites. As expected, the ACLES corpus features *Catalan* (one of the top-ranking keywords), *Spanish* and *Valencian* as noticeable keywords while the AICLU corpus features *Italian*, and, in a similar ranking, *English language* and *Français*.

5 In the figures of this section keywords and multi-word items are presented in alphabetical order.

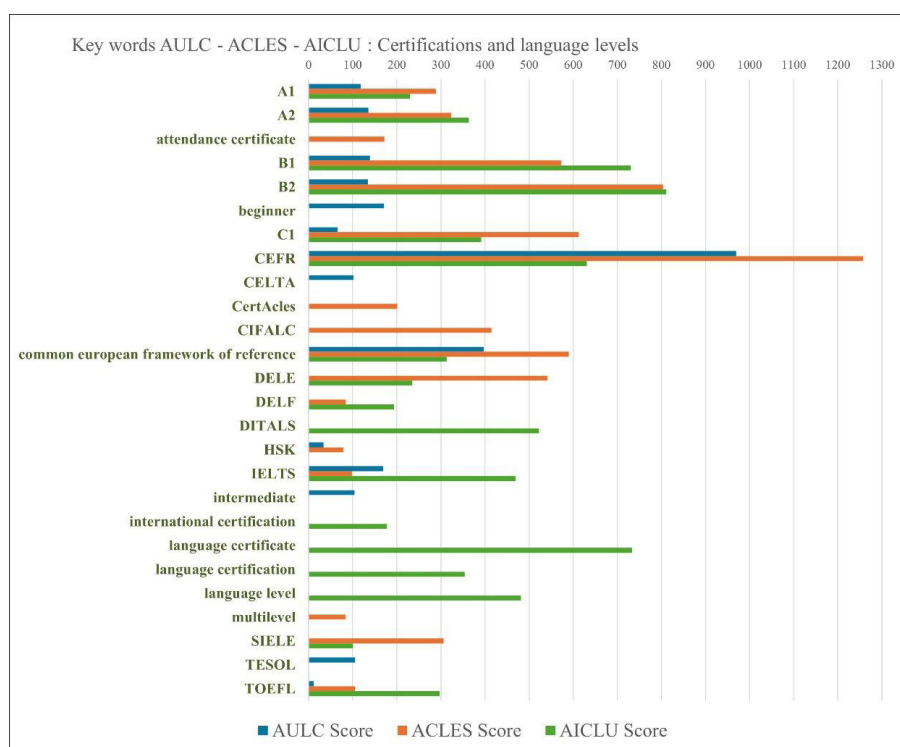


fig. 3. English language keywords: 'Certifications and language level' category.

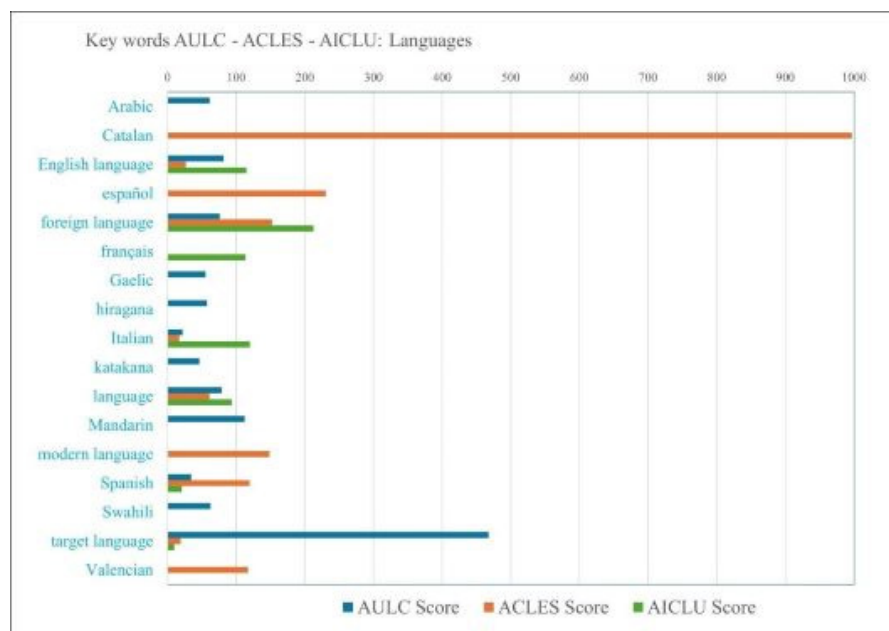


fig. 4. English language keywords: 'Languages' category.

‘Learning Management Systems (LMSs) and e-learning platforms’ emerged as the third category in this keywords analysis (see Figure 5). The names of such platforms of course emerge as keywords, as they are very rarely used outside higher education discourse. The only recurring platform across all three corpora in English is *Moodle*, as can be expected, since it is considered the most widely used open-source LMS (Gamage *et al.* 2022). The widespread use of e-learning platforms in general for language teaching is in line with the well-known effectiveness of the use of technology and blended modality in language teaching/learning (Bax 2011, Bonk/Graham 2012, Motteram 2013).

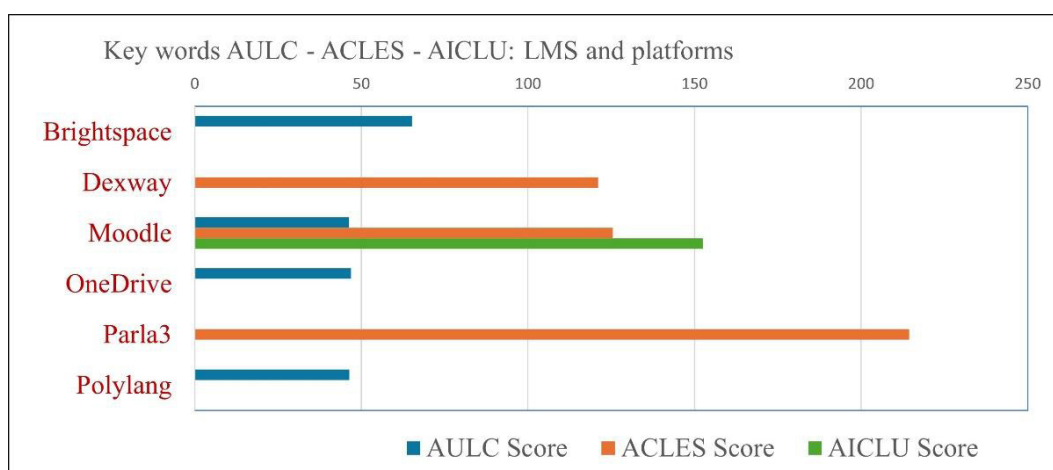


fig. 5. English language keywords: ‘Learning Management Systems and platforms’ category.

The remaining categories comprise the key items relating to the most typical contents for LCs websites, namely practical information about the courses and specific details about the course contents, activities and assessment. A selection of the most relevant items for this research project is presented here.

The category of ‘Course organisation’ features for all three corpora firstly terms such as *language centre* and *language course*, as explanations are given about the institution and its characteristics and teaching offer, and secondly keywords about important procedures and formal steps for students, such as *enrolment* (and the verb *enrol*) and *exam* (see Figure 6). The name of the language centres association for each corpus also emerges as keyword, *ACLES* for the Spanish websites, *AICLU* for the Italian ones, and (low ranking) *AULC* for the UK/Irish ones. High-ranking key terms for the AULC corpus also include *pre-sessional* (course or programme), *course book*, *evening course*, *student support*, as well as *non-credit* and *credit-bearing*, while a significant number of high-ranking key items for both the ACLES and the AICLU English corpora predictably relate to mobility and exchange students, such as the acronym *ECTS* (European Credit Transfer and Accumulation System), *credit recognition*, *ERASMUS student*, and mention of the courses for the local languages

(*Spanish course* for the ACLES corpus and *Italian course* for the AICLU corpus). *Preparation course* and *language test* are also significant key terms for the Spanish and Italian websites, seemingly related to necessary requirements for international students.

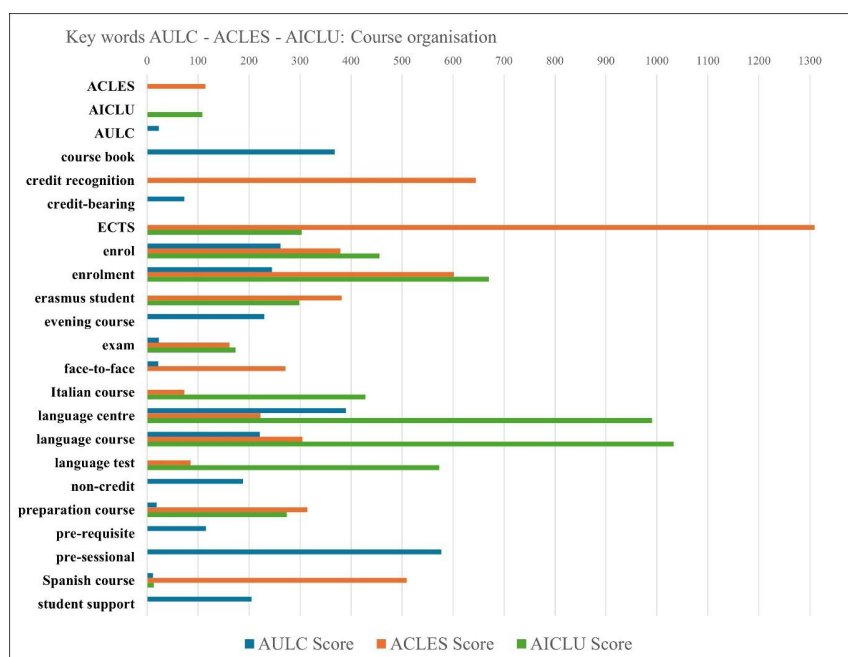


fig. 6. English language keywords: 'Course organisation and procedures' category.

The final keywords category has been labelled 'Teaching activities and assessment' and comprises the representation of the LC practices through the description of their courses and assessment modalities (see Figure 7). The AULC corpus features a wide range of keywords and multi-word key items related to teaching methods (*action-oriented communicative approach*, *inclusive learning environment*, etc.) and activities (*autonomous learning activity*, *whole-class activity*, *small-group activity*, etc.) as well as some about assessment (*self-assess / assessment*, *on-line assessment*, etc.), indicating a tendency to detailed descriptions of courses in the UK / Irish websites. The ACLES and AICLU corpora in English show fewer key items related to course descriptions, such as *language skill*, *communicative approach*, *self-study*, *language proficiency* (common key terms for all three corpora, although with quite different scores), and a few items related to assessment, such as *placement test*, *language assessment* and, exclusively in the AICLU corpus, *entry test*. Thus, the English versions of the Spanish and Italian websites seem to offer fewer details about courses and teaching activities. On the other hand, unlike the UK and ÉIRE websites, they feature *plurilingualism* and *multilingualism* as keywords (although the AICLU corpus shows lower scores for them), confirming their focus on an international audience.

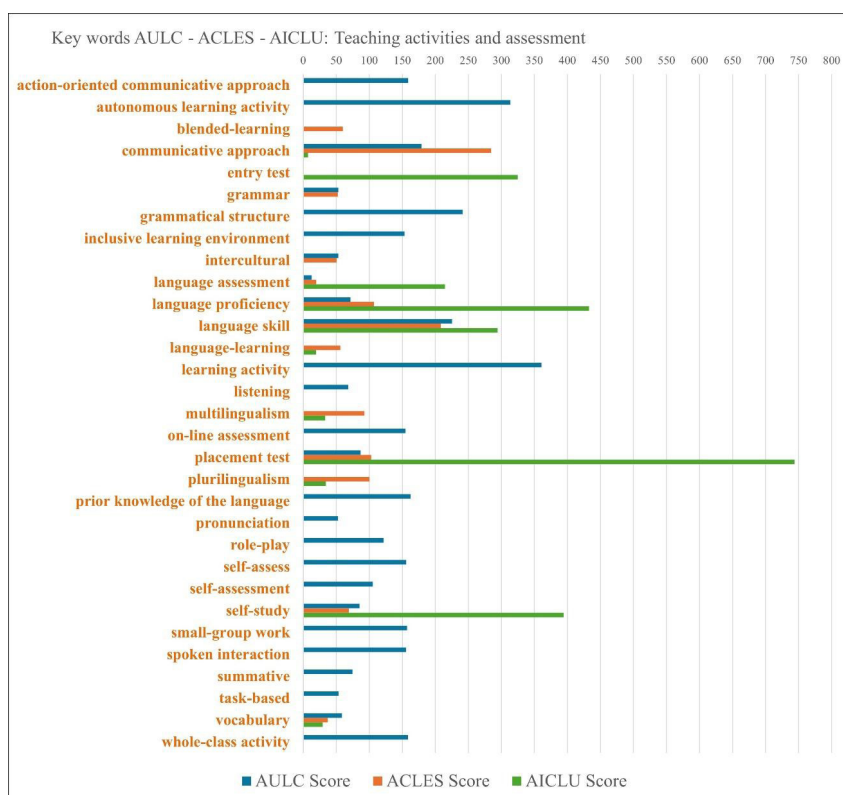


fig. 7. English language keywords: ‘Teaching activities and assessment’ category.

4.2.2. Analysis of Italian language websites

Key lexical features of the Italian and Spanish corpora were analysed using the same method and five categories described earlier. This subsection tackles the AICLU Italian corpus, while subsection 4.2.3 presents the analysis of the ACLES Spanish corpus.

The first category for the AICLU Italian corpus contains key items related to certifications and language levels, which closely resemble those found in the EULC English corpora. (see Table 3⁶). CEFR language levels (*B1*, *B2*, *livello linguistico*, etc.) are the top-ranking keywords, along with names of international certifications (*certificazione internazionale*, *IELTS*, *DITALS*, *DELTA*, etc.), while the CEFR itself features in its Italian version (*QCER*, *quadro comune di riferimento*) as well as in the English acronym.

⁶ In the tables in this section, key items are ranked by score. Items in English are in italics.

certifications	AICLU IT Score	certifications (cont.)	AICLU IT Score
B1	438	DELF	218
B2	411	TOEFL	181
IELTS	409	DELE	180
livello B1	407	A2	180
certificazione internazionale	379	CEFR	175
certificazione linguistica	360	C1	171
livello B2	338	livello C1	166
DITALS	326	certificazioni	165
QCER	312	livello di competenza	137
quadro comune europeo	308	livello di conoscenza della lingua	134
livello A1	294	A1	121
livello A2	265	C2	96
livello linguistico	245	principiante assoluto	91
esame di certificazione	224		

tab. 3: Italian language keywords: ‘Certifications and language level’ category.

For the second category, data show a much wider range of languages featuring as key items in this corpus than in its English version (*lingua francese, spagnola, inglese, italiano per stranieri, cinese*, etc.), including LIS (Italian Sign Language) (see Table 4). This variety suggests that Italian LCs provide more detailed descriptions of their courses on the Italian websites compared to their English versions, which mainly focus on necessary information for international visitors. The corpus also presents a small number of keywords in English (*English, language*, etc.), a recurring element for this corpus, due to the hugely widespread offer of English courses as well as of English-taught degrees (British Council 2018, Triki 2022).

Languages	AICLU IT Score
<i>English</i>	228
lingua francese	175
lingua straniera	164
lingua spagnola	147
<i>language</i>	128

lingua inglese	125
lingue	114
italiano per stranieri	91
lingua cinese	84
lingua tedesca	82
languages	72
lingua italiana	71
LIS	67

tab. 4: Italian language keywords: ‘Languages’ category.

Among Learning Management Systems, *Moodle* features as a keyword in the AICLU Italian corpus just like in the English version (see Table 5). Other platforms in relevant ranking positions are *esse3* (by CINECA, a widespread University administration system for students’ career, exams registration, etc.) and *Teams* (namely Microsoft MS Teams, more frequently used as a video conferencing tool in Italian universities than as a full-blown LMS).

LMS and platforms	AICLU IT Score
Moodle	163
esse3	141
piattaforma moodle	83
Teams	59

tab. 5: Italian language keywords: ‘Learning Management Systems and platforms’ category.

Most key items in the AICLU Italian corpus are included in the ‘Course organisation and procedures’ and in the ‘Teaching activities’ categories. A selection of the most interesting items was made for this study, following the same criteria used for the English corpora mentioned before. Italian websites offer detailed explanations and information about course organisation, logistics and procedures (see Table 6). The highest-ranking key items in the whole corpus are, predictably, the term *language centre* and the acronym for university language centre in Italian (*centro linguistico* and *CLA*), followed by the main services offered by LCs, *i.e.* language courses and exams (*corso di lingua*, *idoneità linguistica*, *esame di lingua*, *courses*). Interestingly, the Italian corpus confirms the focus on internationalisation as it features several key items related to foreign and mobility students, such as *stranieri*, *ERASMUS*, *mobilità internazionale*, *outgoing* (which is commonly used as in English, in Italian discourse about ERASMUS mobility, *e.g.* *studente ERASMUS outgoing*), *studente internazionale*, etc. The acronym for the Italian association of LCs (*AICLU*) is also a keyword, along

with that of the European association of LCs (*CERCLES*).

Course descriptions often mention native speakers as language instructors, explaining the keyword *madrelingua*. LCs facilities are predictably often mentioned, hence key items such as *laboratorio di esercitazioni* and *laboratorio linguistico*. The LC website (*sito del CLA*) is itself a key term.

Course organisation and procedures	AICLU IT Score
CLA	1844
centro linguistico	1724
corso di lingua	360
idoneità linguistica	212
esame di lingua	162
AICLU	148
stranieri	142
<i>courses</i>	130
iscrizioni	129
<i>ERASMUS</i>	126
sito del CLA	121
mobilità internazionale	118
madrelingua	116
prova di idoneità	113
<i>outgoing</i>	112
<i>students</i>	107
studente internazionale	106
corsi	103
laboratorio di esercitazioni	98
studente straniero	98
laboratori	73
<i>CERCLES</i>	72
laboratorio linguistico	69

tab. 6: Italian language keywords: 'Course organisation and procedures' category.

A wide range of key items, both in Italian and in English, refer to assessment and teaching activities and course contents, unlike in the AICLU corpus in English (see Table 7). Key terms related to language teaching and learning emerge, such as

competenza / abilità linguistica, apprendimento delle lingue / linguistico, as well as the traditional language skills (*speaking, listening, produzione / comprensione scritta, produzione / comprensione orale*, etc.). Assessment-related terms also emerge as high-ranking key items, such as *test di piazzamento, test finale, placement test, test di ingresso*. Autonomous online study methods are frequently presented in the Italian websites, as one of the top keywords is *autoapprendimento* and other related terms feature lower in the list, such as *blended, online in autoapprendimento, autovalutazione*. Keywords related to multilingualism (such as *plurilingue, plurilinguismo, multilinguismo*) can be found, although lower in the list, in line with the findings for the AICLU English corpus.

Teaching activities and assessment	AICLU IT Score	Teaching activities and assessment (cont.)	AICLU IT Score
competenza linguistica	345	<i>tandem</i>	116
autoapprendimento	302	intercomprensione	107
test di piazzamento	251	test valutativo	105
test finale	250	<i>blended</i>	100
abilità linguistica	209	apprendimento delle lingue straniere	92
<i>speaking</i>	204	esercitazioni assistite	88
<i>placement</i>	191	online in autoapprendimento	77
apprendimento delle lingue	176	grammaticale	71
<i>academic</i>	176	autovalutazione	67
<i>listening</i>	174	<i>grammar</i>	61
<i>placement test</i>	169	orale	56
produzione scritta	167	<i>vocabulary</i>	54
comprensione scritta	161	lessicale	53
produzione orale	146	<i>elearning</i>	44
<i>writing</i>	139	plurilingue	39
comprensione orale	131	plurilinguismo	36
test di ingresso	130	<i>self-access</i>	34
apprendimento linguistico	128	linguistico-comunicativa	34
test di livello	125	multilinguismo	27

tab. 7: Italian language keywords: ‘Teaching activities and assessment’ category.

4.2.3. Analysis of Spanish language websites

Most of the key terms in the ACLES Spanish corpus are found in the ‘Certifications and levels’ category (see Table 8). The acronym *MCER* (*i.e.*, the Spanish equivalent of CEFR, which stands for *Marco Común Europeo de Referencia*) tops the list, followed by the CEFR language levels (*B1*, *B2*, *C1*, etc.) and the names of some international language certifications (*DELE*, *SIELE*, *DELTA*, etc.). Moreover, one of the top-ranking items is *CertAcles*, which refers to an official nationally recognised certification issued by the LCs that are members of the ACLES. This data is consistent with the fact that Spanish websites are the ones that most often mention internal certifications, as described in 4.1.

Certifications and levels	ACLES ES Score	Certifications and levels (cont.)	ACLES ES Score
MCER	740	C2	207
B1	660	DELTA	189
B2	651	CCSE	180
CertAcles	508	TOEFL	148
C1	489	IELTS	142
A2	373	Cambridge	135
A1	311	HSK	112
examen oficial	269	certificado oficial	103
DELE	244	DALF	101
SIELE	233	multinivel	82

tab. 8. Spanish language keywords: ‘Certifications and language level’ category.

In the second category, the data show a limited number of languages featuring as top-ranking items in the ACLES Spanish corpus (see Table 9). Some of them (*idiomas*, *lenguas*, *lenguas extranjeras*) refer to languages in general, while others refer to specific languages (*lengua catalana*, *lengua inglesa*, *portugués-brasileño*). The Spanish language is included in the form of the acronym *ELE*, which stands for *Español como Lengua Extranjera*, probably because Spanish courses are primarily addressed to foreign students. The English language also features as *English*.

Languages	ACLES ES Score
idiomas	478
lenguas	448
English	144
lengua extranjera	140
lengua catalana	95
lengua inglesa	74
ELE	53
portugués-brasileño	43

tab. 9. Spanish language keywords: ‘Languages’ category.

For the third category, no table has been included in this paper, as no top-ranking items related to Learning Management Systems and platforms have emerged from the analysis (e.g., *Moodle* has a rather low score).

As for the key terms related to ‘Course organisation and procedures’ (see Table 10), they refer mainly to course recipients (*miebro de la comunidad universitaria*, *PDI*, *persona externa*, etc.), course format (*semipresencial*, *presencial*, *videoconferencia*, etc.), course registration (*matrícula*, *preinscripción*), exams and credits (*ECTS*, *reconocimiento de créditos*, *acreditación*, etc.). The acronym for the Spanish association of LCs (ACLES) is also a keyword.

Course organisation and procedures	ACLES ES Score	Course organisation and procedures (cont.)	ACLES ES Score
ECTS	300	videoconferencia	126
ACLES	296	ERASMUS	120
matrícula	199	persona externa	119
miembro de la comunidad universitaria	172	preinscripción	116
semipresencial	171	acreditación	107
PDI	163	PAS	106
curso de español	160	certificado de aprovechamiento	105
reconocimiento de créditos	158	curso de idiomas	91
presencial	151	estudiante internacional	81
curso intensivo	140		

tab. 10. Spanish language keywords: ‘Course organisation and procedures’ category.

Along with the first category, the last one features the highest number of key terms in the ACLES Spanish corpus (see Table 11). This category includes several multi-word terms, most of them related to traditional language skills (*expresión escrita, comprensión auditiva, interacción oral*, etc.), which also occur as single-word terms in English (*speaking, listening, writing*, etc.). Data show a few assessment-related items, such as *evaluación continua* and *prueba de nivel*, with the latter ranking top of the list. There are also some items related to teaching methods (*enfoque comunicativo, trabajo autónomo, autoaprendizaje*, etc.) and to teaching and learning in general (*competencia lingüística, destreza comunicativa, actividad oral*, etc.). Similarly to the AICLU Italian corpus, the ACLES Spanish corpus features *plurilingüismo* as a key term with a low score in this category. Unlike the Italian corpus, in the Spanish corpus the acronym CLIL (*Content and Language Integrated Learning*) emerges as a significant keyword.

Teaching activities	ACLES ES Score	Teaching activities (cont.)	ACLES ES Score
prueba de nivel	786	evaluación continua	92
expresión escrita	429	autoaprendizaje	86
comprensión auditiva	388	intercambio lingüístico	85
trabajo autónomo	360	participación activa del alumno	84
interacción oral	237	<i>reading</i>	80
<i>speaking</i>	219	enfoque comunicativo	77
<i>listening</i>	187	plurilingüismo	53
<i>writing</i>	169	gramática	45
CLIL	122		
actividad oral	105		

tab. 11. Spanish language keywords: ‘Teaching activities and assessment’ category.

5. CONCLUSIONS AND FUTURE STEPS

The analysis of EULC websites revealed significant structural differences, first of all Italian websites are predominantly independent, while UK/Irish and Spanish sites are more often embedded. Moreover, both the structural distinctions on the ‘Certifications’ and ‘Internationalisation’ sections and the lexical analysis reflect various degrees of attention from different linguistic areas on internationalisation, with Italian and Spanish centres emphasising mobility and international students more prominently than their UK/Irish counterparts.

In terms of original versus translated versions, Spanish websites, provide the widest range of translations, followed by Italian and then UK/Irish websites, mostly due to

the versions in the co-official languages in Spain. From a linguistic perspective, a closer examination of lexical features highlighted differences in the terminology used for language certifications, course structures, and assessment. The AICLU and ACLES English corpora emphasise international language certifications and the importance of the *Common European Framework* (CEFR), while UK/Irish websites place greater emphasis on internal certifications and course descriptions for a local audience - though this doesn't necessarily indicate a lack of interest in internationalisation. The Italian and Spanish websites provide a wider range of certifications, both internal and external, reflecting their greater emphasis on internationalisation. The Spanish and Italian corpora display a broader range of languages and multilingual elements, reflecting the diverse linguistic offerings and multilingual approach of institutions in these linguistic areas.

As a follow-up to this research, a detailed discourse analysis of the argumentative strategies employed on EULC websites could offer deeper insights into how these institutions promote their services and engage their different target audiences. To this purpose, the investigation on persuasive techniques and on the use of language to foster an inclusive and international academic community could be crucial. The analysis on internationalisation and multilingualism would allow for a deeper understanding of how these websites target international students, especially in terms of content accessibility, materials and language offerings.

Future research could also explore the layout of these websites, including their multimodal elements such as images and videos. This would provide an in-depth view of how visual elements complement the textual content to enhance users' engagement.

Finally, analysing the structure and lexical features of university sports centres could yield valuable insights into the specialised language and communication strategies used to represent this specific domain. By extending the analysis to this new dimension, the Foro Italico RU will contribute to a broader understanding of how European Universities use websites for specific centres of services to communicate with and support a multilingual, international audience, create communities of practice and convey socio-cultural values specific to their institutional environment.

REFERENCES

- Bax 2011 = Stephen Bax, *Normalisation Revisited: The Effective Use of Technology in Language Education*, in «International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)», 1(2), pp. 1-15
- Bondi 2010 = Marina Bondi, *Perspectives on keywords and keyness: An introduction*, in Marina Bondi / Mike Scott (eds.), *Studies in Corpus Linguistics*, Amsterdam, John Benjamins Publishing Company, pp. 1-18.
- Bonk/Graham 2012 = Curtis J. Bonk / Charles R. Graham, *The Handbook of Blended Learning: Global Perspectives, Local Designs*. San Francisco, CA, John Wiley & Sons.
- British Council 2018 = British Council, # EU2025English: *The Future Demand for English in Europe: 2025 and Beyond*, British Council, <https://www.britishcouncil.org/education/schools/support-for-languages/thought-leadership/research-report/future-of-english-eu-2025> (last access 14/03/2025)
- Caiazza 2014 = Caiazza Luisa, *Emerging Conventions in the Verbal Component of the 'About' Page of British University Websites*, in *Evolution in Genre, Emergence, Variation, Multimodality*, Bern, Peter Lang
- Chapleo 2015 = Chris Chapleo, *Brands in Higher Education*, in «Challenges and Potential Strategies», *International Studies of Management & Organization*, 45(2), pp.150–163. <https://doi.org/10.1080/00208825.2015.1006014>
- Council of Europe 2020 = *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing, Strasbourg, available at www.coe.int/lang-cefr.
- Council of Europe 2019 = *Council Recommendation of 22 May 2019 on a comprehensive approach to the teaching and learning of languages (2019/C 189/03)* [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019H0605\(02\)](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019H0605(02)).
- Fairclough 2003 = Norman Fairclough, *Analyzing Discourse: Textual Analysis for Social Research*, London, Routledge.
- Gamage et al. 2022 = Sithara HPW Gamage / Jennifer R. Ayres / Monica B. Behrend, *A Systematic Review on Trends in Using Moodle for Teaching and Learning*, in «International Journal of STEM Education», 9 (1), p. 9. <https://doi.org/10.1186/s40594-021-00323-x>
- Kilgarrieff et al. 2014 = Adam Kilgarrieff / Vít Baisa / Jan Bušta / Miloš Jakubíček / Vojtěch Kovář / Jan Michelfeit / Pavel Rychlý / Vít Suchomel, *The Sketch Engine: Ten Years On*, in «Lexicography», 1, pp. 7-36.
- Kilgarrieff et al. 2004 = Adam Kilgarrieff / Pavel Rychlý / Pavel Smrž / David Tugwell, *The Sketch Engine*, in *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004*, 1, pp. 105-116.
- Kress/van Leeuwen 2001 = Gunther Kress / Theo van Leeuwen. *Multimodal discourse: The modes and media of contemporary communication*, London, Arnold.
- McEnery/Hardie 2011 = Tony McEnery / Andrew Hardie. *Corpus Linguistics: Method, Theory and Practice*, Cambridge, Cambridge University Press.
- Motteram 2013 = Gary Motteram (ed.), *Innovations in Learning Technologies for English Language Teaching*, London, British Council.
- Sketch Engine, <http://www.sketchengine.eu>
- Stanford 2009 = Jeff Stanford, *Moodle 1.9 for Second Language Teaching*, Birmingham, Packt Publishing.
- Tognini-Bonelli 2001 = Elena Tognini-Bonelli, *Corpus Linguistics at Work*, Philadelphia/ Amsterdam, John Benjamin Publishing Company
- Tognini-Bonelli 2008 = Elena Tognini-Bonelli, *Functionally Complete Units of Meaning across*

- English and Italian: Towards a Corpus-Driven Approach*, in Bengt Altenberg / Sylviane Granger (eds.), *Lexis in Contrast: Corpus-based Approaches*, Amsterdam, John Benjamins Publishing Company, pp. 73-95.
- Triki 2022 = Nesrine Triki, *Language Policy and Planning: Aspects of Plurilingualism at an Italian University*, in «RiCOGNIZIONI. Rivista Di Lingue E Letterature Straniere E Culture Moderne», 9(17), pp. 9-28. <https://doi.org/10.13135/2384-8987/6667>
- Wenger 1999 = Etienne Wenger, *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, Cambridge University Press.