We are living in a 'new linguistic dispensation' revolving around multilingualism and multiculturalism, technology galore and an unprecedented mobility of people and linguistic artefacts. New opportunities have opened up to L2 users via the Internet, satellite communication and wireless technologies, while English language learning is moving out of the language classroom and becoming more and more informal ⬚ unguided, naturalistic, mainly incidental. The present volume engages with this new English-centred linguistic ecology by probing university students' private worlds and undetected learning processes across Italy. The linguistic affordances available 'in the wild' are also explored with the aim of identifying the features that are likely to impact informal language learning today. A few key thematic threads run across the volume. They concur to defining shared features of the novel modalities of access to and appropriation of English as an additional language in Italy. At the same time, they detail Italian learner-users' increasingly diverse language experiences, individual language learning trajectories, as well as intrinsic and extrinsic motivations in engaging with English during leisure time. The impact of the (new) technologies and the digital revolution on the expanding informal learning of English in the country is highlighted in all twelve contributions, with the access to audiovisual, digital, and interactional media emerging as a new form of language immersion freely available to English learner-users worldwide. Overall, this volume contributes to depicting the varied landscape of English language learning in the wild and provides new etic and emic insights into the relation that links youths to the language in online (social) media and registers, and pop cultural artefacts.

Chapters by:

Maria Pavesi, Francesca Bianchi
Maria Pavesi
Elisa Ghia
Serena Ghiselli
Paola Leone, Emanuela Paone
Cristina Mariotti

Camilla De Riso
Gloria Cappelli, Nicoletta Simi
Elena Manca
Francesca Bianchi
Maicol Formentelli, Raffaele Zago
Eliana Galiano

vol. 66 – Special Issue   2024

'In the Presence of English' Informal English Language Learning in Italy

Edited by

Maria Pavesi
Francesca Bianchi

Lingue e Linguaggi

vol. 66 - Special Issue
2024

Università del Salento

Lingue e Linguaggi

66/2024

Special issue

'In the Presence of English' Informal English Language Learning in Italy

edited by

Maria Pavesi

Francesca Bianchi

LINGUE E LINGUAGGI

Indice

CHAPTERS

'IN THE PRESENCE OF ENGLISH': INFORMAL ENGLISH LANGUAGE LEARNING IN ITALY
An introduction

MARIA PAVESI[1], FRANCESCA BIANCHI[2]
[1]UNIVERSITÀ DI PAVIA, [2]UNIVERSITÀ DEL SALENTO

1. Setting the scene

We are living in a 'new linguistic dispensation' (Aronin *et al.* 2013) centred around multilingualism and multiculturalism, technology galore and an unprecedented mobility of people and linguistic artefacts. Concomitantly, English language learning is moving out of the language classroom and becoming more and more informal – i.e., unguided, naturalistic, mainly incidental – as L2 users have extensive individual experience with English-language media during their leisure time (Sockett 2014). They also engage in joint activities online by getting involved in communities of practice, contact zones and affinity spaces in English (Vasquez-Calvo 2021). These dramatic changes in linguistic ecology require "a new model of what constitutes a linguistic environment for learning" (Arnbjörnsdóttir, Ingvarsdóttir 2018, p. 7), to account for the complexification of the linguistic landscapes in which second language acquisition (SLA) is dynamically embedded.

Quite recently, Benson (2021) has put forward the strongest argument in favour of the centrality of space in SLA: space comes first as it subsumes and conditions all the other variables involved in L2 learning. By adopting a spatial perspective, in the present volume we thus espouse the view that additional languages are learnt informally in language learning environments, or spaces, that exceed the tangible and intangible boundaries of institutional settings and are populated by a great variety of language-bearing assemblages – most of which include a non-human component. That is, new opportunities have opened up to L2 learner-users while language learning is increasingly becoming a disembodied experience in the wild (Wagner 2015), carried out via the Internet, satellite communication and wireless technologies. The primacy of informal contexts in English SLA does not however exclude that formal and informal practices may intersect, as spaces blend in the L2 speakers' experience over time (Reinhardt 2022). Indeed, given young users' extensive schooling in English worldwide, no learning of L2 English can be uniquely informal. From an early age onwards,

formal and informal learning interweave, although each of the two processes may be predominant at different stages in an individual's life.

The informal learning of English has been investigated on a large scale in several European and Asian countries, extending to the other continents as well (Dressman, Sadler 2020; Toffoli *et al.* 2023; Cole, Vanderplank 2016, among many). In Italy, however, empirical data on the current changes in L2 informal learners' behaviour and attitudes are still limited and restricted to a few case studies (Aiello 2018; Pavesi, Ghia 2020). More fieldwork is thus urgently needed, to create a large empirical database and provide reliable pictures of the constantly changing learning landscapes. This broad research agenda also calls for a critical assessment of the methodological approaches and empirical tools required to move beyond local observations and surface-level analyses. It further requires SLA-oriented accounts of the language registers to which L2 learner-users are most frequently exposed during their favoured leisure activities.

The present volume engages with this new English-centred linguistic ecology by presenting research conducted within the project "The informalisation of English language learning through the media: Language input, learning outcomes and sociolinguistic attitudes from an Italian perspective", funded by the Italian Ministry for University and Research. The overarching aim of the national investigation is to probe students' private worlds and undetected learning processes across Italy, represented by four universities geographically distributed across the country. The project hence qualifies as the first large-scale investigation of Italian university students' modalities of access to English in the wild, their evolving L2 motivation and attitudes, as well as their media-induced L2 development. To comply with its general aim, the project is organized in three stages and implements a mixed research design. It combines cross-sectional and longitudinal data collection, coupled with an array of empirical tools and quantitative and qualitative approaches to data analysis. The national investigation began with the creation of a fine-grained questionnaire – IECoL[1] – focusing on language behaviour outside formal settings, which was administered to selected and comparable samples of students enrolled at the universities of Pavia, Pisa, Salento and Catania; IECoL was coupled with a vocabulary test assessing participants' receptive lexical knowledge. The first phase of the project was followed by ethnographic investigations via emically-orientated semi-structured interviews. These were meant to illuminate behavioural patterns over time, as well as students' beliefs, attitudes and language learning motivations associated with English-language media. Finally, the project has gained access to the development of media-triggered L2 competence by observing learners over time. In parallel with the three main stages, it also engages with corpus-based descriptions of relevant media genres and registers. These are necessary to inform hypotheses about the impact of different input types upon the learning of L2

---

[1] The acronym stands for "Informal English Contact and Learning" (Pavesi *et al.* 2023).

English beyond the classroom. Although linguistic research on media registers is extensive, it mainly addresses systemic, sociolinguistic, textual and pragmatic aspects, rarely explicitly exploring register usability for SLA purposes.

2. Looking inside the volume

The present collection of papers aims to address some of the issues highlighted above, which were debated at the panel "'In the presence of English': Informal English language learning in Italy" during the 31st AIA Conference (Rende, 13-16 September 2023). By borrowing part of the title from Berns *et al.*'s (2007) research monograph, titled "In the Presence of English. Media and European Youth", we manifested our indebtedness to the scholars who pioneered the investigation of media-driven informal use and learning of L2 English among the youth. Since the time when Berns and colleagues collected their data among German, Belgian, French and Dutch secondary school students, access to entertainment, information and disciplinary content in online media, social media, and pop culture in English appears to have bourgeoned in Italy. The time is thus ripe to look at the informal learning of the language in this country.

A few key thematic threads run across the volume. They concur to defining shared features of the novel modalities of engagement with and appropriation of English as an additional language in Italy. At the same time, they detail learner-users' increasingly diverse language experiences and language learning trajectories. The constructs of informal language learning and the related ones of formal and non-formal language learning are variously explored in all the contributions to this volume. Two different perspectives are adopted – the etic one of external observers and the emic one of the L2 users, in both cases focusing on the defining spatial dimensions of the learning process. The spatial perspective entails a temporal dimension as well, in that each individual's learning path changes in response to the contact with different learning environments but is also the result of previous encounters with the language. The impact of the (new) technologies and the digital revolution on the expanding informal learning of English is also highlighted in all contributions, with the access to audiovisual, digital, and interactional media emerging as a new form of language immersion freely available to English learner-users worldwide. By shifting the research focus from reported behavioural patterns to motivation, language attitudes and reasons for accessing English-language media, the inner factors that both reflect and promote change in language habits are explored etically as well as emically. Learners' perceptions, motivations and attitudes are relevant to the investigation of informal language learning. They participate in the construction of settings for SLA (Benson 2021) and affect the awareness of the learning potential of media input. Finally, the linguistic affordances available for informal language learning

are examined in descriptive and corpus-based analysis of English-language media input (e.g., audiovisual dialogue, web-based genres, videogames) with the aim of identifying the features that are likely to impact out-of-class SLA.

The structure of the volume ideally mirrors the sequential organisation of the research phases of the national project – commencing with the large-scale survey, and continuing with the qualitative interviews and the longitudinal case studies.

The three phases together amply cover the main, chronologically-oriented stages of research on informal language learning discussed in Kusyk *et al.* (2023). According to the authors, initial research on informal language learning has focused on the inventory of the informal activities which learner-users are involved in, with the aim of identifying the main patterns in frequency of use, accompanying reasons for access, the role of receptive over productive activities, as well as learner-users' profiles. Across the globe, this initial, quantitative research has alerted "researchers to the arrival of a genuine phenomenon rather than just the favored pastimes of certain enthusiastic outliers" (Kusyk *et al.* 2023, p. 5). Banking on these initial results, research has moved on to analysing social and personal variables and exploring individual perspectives and pathways. This second stage can be posited to include a focus on learners' own voices and L2 development over time. The third stage of research focuses on the "hows" of informal language learning and the complex way in which "users avail themselves of the affordances provided by authentic L2 materials outside the classroom" (Kusyk 2023, p.5).

Since several studies suggest that prolonged naturalistic contact with traditional and digital registers has an impact on target language development and L2 digital multimodal literacy, the volume includes studies on the salient features of relevant media discourses. These represent an additional, important component of research on the informal language learning of English, in that they help to characterize the multimodal input to which L2 speakers are mostly exposed to in the wild.

3. The contributions

In her opening article, **Maria Pavesi** presents the main developing lines of the research project. Moving from a definition of informal language learning as "involving free, self-directed contact with the L2 in non-educational locations for entertainment, social and information-seeking purposes" (Pavesi, Ghia 2020, p. 40), the paper offers a critical review of the theoretical and methodological foundations of the project. A spatial perspective on SLA is adopted as the investigation at large "shine[s] the spotlight on the where of language learning" (Benson 2021, p. 1) and the space of human and human/non-human interactions. Within a spatial perspective, the project centres on language input as the most

crucial construct, the *sine qua non* of SLA (VanPatten *et al.* 2020), which is here assessed with reference to the learning potential of media affordances that learner-users encounter in the wild. After outlining the complementary empirical phases of the project, Maria Pavesi carries out a factor analysis on the 2,500 fully completed questionnaires collected at the four participating universities. Through the analysis, Pavesi identified the informal practices that were undertaken together by individual learner-users and were shared across individual learner-users. This profiling operation represents an important first step in the mapping of the informal learning spaces that Italian university students create for themselves through L2 English practices such as watching audiovisual products, playing video games and becoming involved in person-to-person interactions online and offline.

Attention then moves to the results of the IECoL questionnaire, which was administered to comparable samples of students in the four universities of the project. **Elisa Ghia** investigates Pavia university students' informal experience with various multimodal genres – films, TV series and other programmes, and YouTube videos – focusing on the reasons behind their habits. The data were collected through IECoL, administered to 1,091 students from three main disciplinary areas as well as language specialists. The findings show an extensive and intensive experience with English-language audiovisuals among the Italian university students surveyed, with users preferring TV-series and YouTube contents over other audiovisual material and accessing them on a regular basis. English subtitles are preferred across all AV genres, even when they are automatically generated, as is often the case on YouTube. The emerging picture complexifies when reasons for informal experience with audiovisuals are explored. Most respondents report several motivations as guiding their involvement with multimodal English input. A positive orientation to the L2 – comprising both the desire to learn the L2 and the liking of English and English-language original voices and products – generally coexists with pure entertainment and other hedonic reasons. Also, a degree of differentiation emerges when looking at genres and students' areas of studies. Similarly, **Serena Ghiselli** illustrates and discusses the results of the IECoL questionnaire with reference to the students of the University of Salento, with a special focus on the Internet and YouTube. The data clearly show that students predominantly engage in receptive rather than productive activities in English, with a preference for audiovisual content. Among the respondents, YouTube proved to be a widely used platform for viewing videos in English, usually for short durations but on a regular basis and frequently with subtitles. On the other hand, the data showed much lower access to other Internet genres in English, with the reading of social media posts dominating among Internet-based activities. Ghiselli also explores the participants' background and observed that the self-assessed level of English and the student attitude towards this language influences the use of media in

English, while factors such as the secondary school attended and language attitudes significantly impact media exposure.

Two contributions follow, both targeting motivation as a broad and multifaceted construct involved in the informal learning of English. **Paola Leone and Emanuela Paone** offer a qualitative case study that explores the dynamic nature of motivation. By adopting Dörnyei and Ushioda's (2008) L2 motivational self-system theory, their work starts from the assumption that behavioral phenomena cannot be explained by a single factor, with complexity arising from patterns rather than simple cause-and-effect relationships. Motivation is discussed by considering its three dimensions: L2 Learning Experience, which looks at the immediate learning environment and its influence on language learning; Ought-to L2 Self, reflecting other people's expectations; Ideal L2 Self, representing the learner's expectations as a successful second language user. Drawing on open-ended interviews, the paper investigates how 26 university students at the University of Salento perceive the factors driving their informal engagement with English across various media platforms. The findings highlight the complex and multifaceted interactions among external and internal factors and learning experiences. The data uncover the pivotal role of teachers in fostering or diminishing motivation to learn a second language during childhood and adolescence, while indicating students' release from parents' and teachers' external pressures over time. As they begin to take ownership of their language learning, some students proactively seek alternative means, such as non-formal instruction, to enhance their language proficiency. Importantly, media emerged as playing a role in evoking positive emotions and sustaining motivation in language learning. The other contribution on motivation is by **Cristina Mariotti** who analysed semi-structured interviews with 21 students of the University of Pavia within the framework of self-determination theory. Her study reveals the roles of intrinsic (e.g., immersion, entertainment, and personal interests) and extrinsic (e.g., improving English proficiency) motivations and helps understand the complex texture subtending university students' self-directed exposure to English input, and to telecinematic products in particular, during their leisure time. In particular, it was noticed that the decision to access input in English based on the inherent enjoyment and satisfaction derived from the process was associated with very high levels of exposure frequency and intensity, as well as highly positive attitudes towards the language, but not necessarily to advanced proficiency. At the same time, L2 exposure practices appeared to be considerably influenced by the desire to enhance one's English proficiency. Finally, this study suggests that greater exposure and more positive attitudes enhance agency and lead to more meaningful engagements with English.

Next, a more detailed view of informal language learning profiles is offered by **Camilla De Riso**, who reports on her longitudinal research exploring the links between exposure to English registers and learning trajectories. De Riso discusses the output of four longitudinal case studies based on monologic storytelling tasks,

investigating possible relationships between L2 development and frequency, intensity and type of exposure to informal English input. The paper also explores the thoughts and feelings of the four L2 learner-users by investigating their L2 awareness in relation to themselves, others, and the informal activities they access habitually. Complexity, accuracy and fluency trajectories presented strong variability across participants and within individual L2 development. However, some specific elements can be hypothesized to have an impact, such as register and immersion: the written register of novels probably supports lexical and grammatical complexity, while the spoken register of comics and manga does not; immersion in turn seems to enhance lexical complexity. Furthermore, accuracy and fluency appeared to be particularly sensitive to each other.

As learners of English may experience non-formal and formal learning at university as well, the two papers that follow De Riso's study shift the focus from informal learning *strictu sensu*, to the exploration of social media and how they provide affordances and opportunities for language instruction outside traditional educational settings. **Gloria Cappelli and Nicoletta Simi** delve into the dynamic and interactive spaces created by social media platforms. Moving from Wagner's (2021) definition of Instagram as a "multimodal microblogging genre", they critically survey the main assets of Instagram as a privileged site for informal and non-formal English language learning before focusing on Instagram accounts specifically devoted to teaching English as a Foreign Language (EFL). An innovative pedagogically-motivated analytical framework is put forward to critically assess a sample of six institutional and twenty-one non-institutional English language teaching accounts, hence comparing non-formal Instagram teaching to traditional, formal EFL education. The pedagogical categories that make up the framework include contents and topics, teachers and learners, and methodology and modes of presentation. As a main outcome of the critical analysis, the authors show how learning English via Instagram differs from learning the language in traditional settings as the former provides opportunity for individuals to creatively explore content based on personal interests beyond temporal and spatial restrictions and free from rigidly structured syllabi. This study opens up innovative avenues of research on the budding context of non-formal language learning via social media. **Elena Manca** investigates the role of social media in English language learning from the perspective of a special group of users, namely students attending foreign language curricula at the University of Salento. Manca administered a questionnaire that combines some items from the IECoL with new items on social media usage, to ascertain students' preferences and attitudes, but also perceptions regarding the advantages of using social media in English within their L2 learning process. The students taking part in this survey showed a steadfast belief in the language-enhancing capabilities of social media, which suggests exploring ways for integrating social media into formal and non-formal language learning settings as they could contribute to

enhance the students' sensitivity to linguistic nuances and general language proficiency, as well as to promoting cultural exchange and understanding.

The volume continues with three contributions on the nature of the input to which learner-users are typically exposed outside institutional settings, and on specific language features of media registers that are believed to affect L2 acquisitional trajectories. **Francesca Bianchi** investigates the lines of non-player characters in video games and identifies their lexical coverage, i.e. the vocabulary size readers need to understand dialogue between characters without external support. This study confirms previous hypotheses that the lexical profile of in-game dialogue is somehow similar to that of British and American films and that lexical coverage differences largely depend on individual games rather than genre, although genre might play a minor role. Furthermore, the data suggest that setting probably contributes to determining the words used in the dialogues. The results are discussed in relation to language learning. Overall, this paper offers new insights into a type of pop product that is still under-investigated from a linguistic perspective despite its extensive use by young generations. Two contributions follow on audiovisual dialogue, a type of input that has been extensively shown to impact language learning both inside and outside the classroom. In line with the recent attention paid to complexity in language registers and SLA, they tackle complementary aspects of grammatical complexity in films and TV series. **Maicol Formentelli and Raffaele Zago** adopt Biber's (1988) register-functional approach to assess the frequency and functions of clausal and phrasal complexity features in the Sydney Corpus of Television Dialogue, with a focus on finite and non-finite subordinate clauses as well as noun phrase premodification. The distribution of complexity features in TV series is also evaluated with reference to spontaneous spoken language to gauge the degree of proximity between the two registers and ultimately account for the frequency and type of grammatical complexity features in TV dialogue. These are key language elements that are expected to play a significant role in L2 development. This study shows that TV dialogue mimics spontaneous conversation to a considerable extent, but characterises itself for register-specific features that are related to the audience-oriented, narrative functions and scripted nature of television dialogue. The authors conclude the paper by discussing the acquisitional implications of the uncovered complexity patterns. Finally, **Liviana Galliano**'s contribution zooms in on phrasal complexity in film dialogue by adopting the same register-functional approach to complexity (Biber *et al.* 2022) as in Formentelli and Zago's paper. Galliano analyses the latest version of the Pavia Corpus of Film Dialogue to unveil the defining features of nominal pre- and post-modification in the register, while assessing how close in frequency and use these complexity patterns come to analogous features in natural conversation. The results suggest that film dialogue most often relies on adjectival pre-modification, which is followed by nominal pre-modification and post-modification via prepositional phrases. The higher degree of phrasal complexity

in film dialogue is discussed with reference to the functional specificities of the register, which, however, do not detract from its general similarity to natural conversation.

Overall, this volume contributes to depicting the varied landscape of English language learning in the wild and provides new etic and emic insights into the relation that links youths to the language in online (social) media and registers, and pop cultural artefacts. New opportunities have opened up to L2 learner-users while informal language learning is increasingly becoming a largely disembodied experience mainly carried out via new and traditional media, the Internet, satellite communication and wireless technologies. This new scenario cannot be disregarded and should be capitalised on. We hope the insights offered by the results of these studies may also inform new approaches to the teaching of English at tertiary level.

References

Aiello J. 2018, *Negotiating Englishes and English-speaking Identities*, Routledge, London/New York.

Arnbjörnsdóttir B. and Ingvarsdóttir H. 2018, *Language development across the life span. English in Iceland: From input to output*, in Arnbjörnsdóttir B. and Ingvarsdóttir H. (eds), *Language Development across the Life Span. The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham, pp. 1-18.

Aronin L., Fishman J.A., Singleton D. and Ó Laoire M. 2013, *Introduction. Current multilingualism: A new linguistic dispensation*, in Singleton D., Fishman JA., Aronin, L. and Ó Laoire M. (eds), *Current Multilingualism: A New Linguistic Dispensation*, De Gruyter, Berlin, pp. 3-24.

Berns M., de Bot K. and Hasebrink U. 2007, *In the Presence of English: Media and European Youth*, Springer, New York.

Biber D. 1988, *Variation across Speech and Writing*, Cambridge University Press, Cambridge.

Biber D., Gray B., Staples S. and Egbert J. 2022, *The Register-functional Approach to Grammatical Complexity: Theoretical Foundation, Descriptive Research Findings, Application*, Routledge, London.

Cole J. and Vanderplank R. 2016, *Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning*, in "System" 61, pp. 31-42.

Dörnyei Z., Ushioda E. (eds) 2009, *Motivation, Language Identity and the L2 Self*, Multilingual Matters, Bristol.

Dressman M. and Sadler R.W. (eds) 2020, *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Pavesi M. and Ghia E. 2020, *Informal Contact with English. A Case Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English: An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Reinhardt J. 2022, T*he History of learning and teaching beyond the classroom*, in Reinders H., Lai C. and Sundqvist P. (eds), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, London/New York, pp. 9-23.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave MacMillan, London.

Toffoli D., Sockett G. and Kusyk M. (eds) 2023, *Language Learning and Leisure. Informal Language Learning in the Digital Age*, Mouton de Gruyter, Berlin/Boston.

Vandergriff I. 2016, *Second Language Discourse in the Digital World. Linguistic and Social Practices in and beyond the Networked Classroom*, John Benjamins, Amsterdam.

VanPatten B, Williams J., Keating G.D. 2020, *Introduction: The nature of theories*, in VanPatten B., Keating G.D. and Wulff S. (eds), *Theories in SLA. An Introduction*, Routledge, London/New York.

Vasquez-Calvo B. 2021. *Guerrilla fan translation, language learning, and metalinguistic discussion in a Catalan-speaking community of gamers*, in "ReCall" 33 [3], pp. 296-313.

Wagner J. 2015, *Designing for language learning in the wild: Creating social infrastructures for second language learning*, in Cadierno T. and Eskildsen S.W. (eds), *Usage-based Perspectives in Second Language Learning*, Mouton de Gruyter, Berlin, pp. 75-101.

Wagner K. 2021, *Using Instagram for language learning*, in "Second language research & practice" 2 [1], pp. 154-164.

THE INFORMALISATION
OF ENGLISH LANGUAGE LEARNING
A first national research project

MARIA PAVESI
UNIVERSITÀ DI PAVIA

Abstract – The availability of English through traditional and new media has recently increased exponentially, thus multiplying the opportunities for contact with the language worldwide. Concomitantly, extensive L2 experience outside institutional settings is likely to lead to the informalisation of both language learning and use across Europe and elsewhere. However, little is known about the acquisitional and sociolinguistic impact of media-induced contact with English in Italy at present. To address this gap, the Italian national PRIN project that is presented in this article investigates students' private worlds and undetected processes of L2 learning and appropriation at four universities situated in different regions in Italy. The national investigation is informed by a plurality of theoretical premises that are integrated into a multidimensional and interdisciplinary research framework that builds on (i) the spatial dimensions of informal contact with English, (ii) the crucial role of language input in language acquisition, (iii) media register variation giving way to individualised language input, and (iv) the evolving attitudes and stances towards English as an additional language. The stages of the research project will be presented. These involved the creation and administration of a questionnaire about university students' experiences with English and a receptive lexical test, ethnographic investigations into L2 media users' behavioural patterns, motivations and attitudes, and longitudinal studies of untutored high-exposure respondents. Finally, a preliminary statistical overview of the participants' profiles and preferred modalities of informal contact with English is provided as a first illustration of the generalisations that can be obtained via the data collections within the project.

Keywords: Informal language learning; L2 English; Media input; Italian university students; L2 users' exposure profiles.

1. Introduction

Due to contemporary globalisation and media saturation, the availability of English through traditional and new media has increased at an unprecedented rate. Concurrently, extensive contact with the language outside educational settings is leading to the increasing informalisation of L2[1] learning and use

---

[1] In the present article, 'L2' is used to refer to a second, foreign or an additional language.

(Arnbjörnsdóttir, Ingvarsdóttir 2018; Reinhardt 2022; Toffoli *et al.* 2023, among others). These dramatic changes in linguistic circumstances in Europe and elsewhere have recently attracted scholarly attention, as they have been shown to deeply affect the language behaviour and L2 proficiency of young L2 users (Azzolini *et al.* 2022; Berns *et al.* 2007; Dressman, Sadler 2020; Sundqvist, Sylvén 2014; Sockett 2014; Toffoli *et al.* 2023, among others). Of note, while multilingualism and multiculturalism have expanded worldwide, English is the additional language that is mainly involved in present-day informalisation processes due to its unrivalled dominance in the media and popular culture and the increasing role of the Internet worldwide (Werner 2022). As a result, the divide between foreign languages other than English and English itself is widening, with English increasingly moving from the status of a foreign language to that of a second language particularly in northern and central European countries, where "contact with English is almost unavoidable" (de Bot 2014, p. 416; see also De Wilde *et al.* 2021; Arnbjörnsdóttir, Ingvarsdóttir 2018; Modiano 2024). Despite this linguistic scenario, little is known about the acquisitional and sociolinguistic impact of media-induced, informal contact with English in Italy.[2]

To fill this gap, the national PRIN[3] project "The informalisation of English language learning through media: Language input, learning outcomes and sociolinguistic attitudes from an Italian perspective" was conceived to explore what is happening 'out there' in the changing linguistic ecology of the country by investigating Italian university students' private worlds and their undetected processes of L2 learning and appropriation. The project is informed by a multidimensional research framework that is based on a series of descriptive and theoretical premises; it moves from the definition of informal language learning as mainly occurring incidentally and "involving free, self-directed contact with the L2 in non-educational locations for entertainment, social and information-seeking purposes" (Pavesi, Ghia 2020, p. 40). According to Dressman (2023), informal language learning can also be conceived of as a cultural practice, with participants across the world sharing some common features, such as their young age and a specific learning setting – the Internet.

---

[2] A few exploratory studies have recently been conducted in Italy (Aiello 2018; De Riso 2023; Pavesi, Ghia 2020). However, their small size and limited geographical scope make it difficult to make generalisations across the country.

[3] PRIN stands for 'Progetto di Interesse Nazionale' (Project of National Interest). PRIN projects are selected on a competitive basis and financed by the Italian Ministry of University and Research.

The various dimensions and phases of the project will be discussed in the remainder of this article. Section 2 addresses the theoretical premises of the study by focusing on the spatial perspective adopted herein, the role and type of language input that is available in informal learning, and the learner-users' attitudes and motivations. Section 3 outlines the three phases of the investigation, while Section 4 presents the outcomes of a factor analysis that identified the prototypical exposure profiles among the university students who were surveyed. In the concluding section, the main tenets of the project are summarised with a view to providing a backdrop to the individual research strands that are currently being developed in the project.

2. Premises and theoretical background

2.1. Informal language learning: A spatial perspective

In the project, we adopt a spatial perspective on second language acquisition (SLA), as the study "shine[s] the spotlight on the *where* of language learning" (Benson 2021, p. 1) and the space of human and human/non-human interactions. The *where* and the *how* of SLA are strongly intertwined, since learning choices and learning processes are constrained by the spatial factors "governing the availability and accessibility of resources" (Benson 2021, p. 97). In line with this spatial perspective, a series of space-related constructs are foregrounded, including *landscape*, *space*, *place*, *setting*, *environment*, *trajectory*, *layer* and *context*. In general, the space of contemporary informal language contact and learning is extremely fluid. Not only may learner-users access a wide variety of L2 affordances[4] 'in the wild' (Wagner 2015), but they can also easily shift from one language learning resource to another as their trajectories move across several modalities over time.

Overall, applied linguists agree that informal language learning takes place outside of and beyond the boundaries of traditional language instruction (Dressman, Sadler 2020; Reinders *et al.* 2022; Toffoli *et al.* 2023), where boundaries are not to be interpreted as demarcating physical locations but rather as delimiting concrete or abstract learning spaces (Leone 2023). As mediated artefacts and mediated interactions are progressively populating the space of contemporary informal language experiences, the globalisation of information, popular culture and media technology has created a new, unconstrained and disembodied "environmental layer" for language learning (Benson 2022, p. 30; see also Werner 2022). This out-of-class layer combines

---

[4] In L2 learning, affordances are the resources that are available in the environment and what learners do with them (Benson 2021, p. 95).

with both formal and informal settings dedicated to embodied encounters on the ground, such as traditional language instruction, English medium instruction (EMI) and living abroad experiences. As a result, L2 users will not only share language learning environments, but will also "trace unique spatial trajectories across increasingly varied maps of settings and resources" (Benson 2022, p. 31). Similarly, Sockett (2014, 2023, p. 117) observed that L2 input is socially determined and is the result of the unique combination of online and face-to-face encounters that constitute "personal learning environments" – that is, complex ecosystems in which language users experience the various affordances as a whole.

Within this spatial framework, there is a need to account for the unprecedented linguistic milieu inhabited by contemporary Italian learner-users and to explore the way(s) in which they create their own language environments while acquiring new L2 competencies and developing novel stances towards the target language through contact with English-language media.

2.2. Language input and informal language learning

The PRIN project centres on language input as the most important construct, the *sine qua non* of SLA (VanPatten *et al.* 2020). Whereas the term 'input' was initially used to refer to all the linguistic manifestations to which learners were exposed in the target language, recent theories have defined it more restrictively "[…] as language the learner attempts to comprehend during communication events", hence excluding language that the learner does not access for meaning (VanPatten *et al.* 2020, p. 10). This definition highlights the relevance of the L2 input that learner-users encounter and interact with spontaneously through online and offline media affordances.

Accordingly, the aim of the project is to investigate which types of language and texts are accessed in informal media settings in line with SLA models and approaches that place language input at the core of language development. Extensive comprehensible input has long been claimed to initiate and sustain SLA. Krashen's (1985, 2008) comprehension hypothesis posits that L2 input is not only necessary but is also sufficient for SLA provided it is comprehensible for learners (via contextual information) and is just above their current level of language knowledge. Although Krashen's model has been criticised repeatedly, the validity of its basic tenets has been widely supported in terms of the paramount role of exposure to L2 input and the primacy of input comprehension over production (VanPatten *et al.* 2020). In a recent reassessment of Krashen's comprehension hypothesis, comprehensible input has been relabelled as "communicatively embedded input" to underline that learner are not passive receptors but are active

interpreters of meaning when they process the target language for comprehension (Lichtman, VanPatten 2021, p. 297).

In accordance with the centrality of input, input frequency has been the focus in usage-based approaches to SLA. In this regard, learning is construed as being input driven, with frequency of usage being viewed as the key to acquisition, both in terms of token frequency − how often a given item occurs in the input − and type frequency − the incidence of the wider category to which a given pattern belongs (Ellis 2003, 2015; Ellis, Collins 2009; Muñoz, Cadierno 2021; Sockett, Kusyk 2015). A direct link between experience and learning is posited, with learners subconsciously deriving statistical abstractions of form-function mappings from the language to which they are exposed. Consequently, "the more often constructions are experienced and understood together, the more entrenched they become" (Pérez-Paredes *et al*. 2020, p. 4). Learners will proceed by matching recurring forms in English to the appropriate functions that they extract from context across several utterances, while abstract patterns will be created by drawing on processes of schematisation and analogy (Tomasello 2005). Furthermore, although the "experience of language is unique, learners are statistical and attentive to forms and functions across the input" (Pérez-Paredes *et al*. 2020, p. 7). This leads to the prediction that different access via various media will still result in a high degree of commonality in SLA because learners cognitively experience the same core language constructions, irrespective of the uniqueness of their personal learning environments.

The learning potential of media affordances has been further supported by cognitive models and theories of multimodal input learning, such as Paivio's (1986) dual coding hypothesis and Mayer's (2009) cognitive theory of multimedia learning. These models emphasise the processing component in the acquisition of data by focusing on the redundancy of multimodal input, as expressed via multiple channels, and its positive impact on processing capability and memorisation. For example, inter- and intralingual subtitles increase the comprehensibility and redundancy of L2 input (Ghia 2012), while multimodality generally triggers media immersion and "vicarious participation" in onscreen conversations, which are believed to sustain SLA (Lantolf *et al*. 2020, p. 238).

Focusing on the essential role of interaction in SLA, the interaction hypothesis highlights additional dimensions of L2 input that are expected to lead to situated language learning through media. According to the interaction hypothesis (Gass 2013; Long 1996), interactionally modified input is crucial for both the comprehension of the target language and for L2 development. When communicating with expert users, learners not only access positive input – information about what is grammatical and acceptable in the L2, but also receive negative input – information about what is not

allowed in the target language. Moreover, sequences of negotiation of meaning are activated in interactions, during which more understandable input can be provided by the expert speaker while the L2 learner is pushed to produce comprehensible output (Long 1996; Swain 1995). The interaction hypothesis also has a cognitive dimension, as the learner's selected attention to linguistic features is a relevant step between input reception and intake. Input characteristics, such as word frequency, morphological salience, meaning-form transparency and imageability or concreteness, can assist learners to comprehend the input and to focus on specific features of the target language (Gass, Mackey 2020). These same properties are likely to apply to multimodal media text as novel language features become highlighted through onscreen interactions, contextual signs and subtitling (Ghia 2012). Gaming and social media may also provide interactive negotiation sequences, together with a focus on informal opportunities for spontaneous, untutored SLA. However, the actual effects of interactive media on informal learning have only recently begun to be explored and supported empirically (De Wilde *et al*. 2020).

In summary, a complex conceptual framework that integrates cognitive, affective and social parameters, functioning together dynamically, is required to address informal language learning (Sockett 2014; Sockett, Kusyk 2015).

2.3. The media as language input

Freedom of contact with English offers a previously unparalleled abundance of L2 input. As the media are characterised by remarkable linguistic variability and fluidity (Crystal 2006), they provide rich L2 input that varies in part depending on the affordances the individual learner-user accesses regularly.

A brief look at some of the major media can highlight the richness and variability of the language input that is presently available.[5] Beginning with the traditional media, the fictional dialogues in audiovisual, telecinematic[6] products have been found to approximate face-to-face interactions. By adopting Biber's (1988) multi-dimensional (MD) analysis[7] and its dimensions

---

[5] Given their high and constantly increasing number, a thorough review of media registers is beyond the scope of the present article. For purposes of exemplification, we will only mention a few media products that have been found to be popular among contemporary informal learner-users of English (Dressman, Sadler 2020; Krüger 2023; Kusyk 2020; Pavesi, Ghia 2020).

[6] The adjective is commonly used to refer to the language of both film and television series (e.g. Piazza *et al*. 2011; Werner 2021).

[7] MD analysis is a corpus linguistics approach developed by Douglas Biber (1988) to account for variation in written and spoken registers in English (e.g. Berber Sardinha, Veirano Pinto 2019).

of register variation, Forchini (2012) compared the language in a corpus of American films, the *American Movie Corpus*, to the *Longman Spoken American Corpus*. She found that Biber's (1988) first and main dimension of *involved* versus *informational production* (i.e. Dimension 1) accounted for the striking similarity between film language and the language of conversation. Personal pronouns and deictic features, contractions, present tense markers and emphatics abound in films, together with other features of contextualised, impromptu speech. The colloquiality and naturalness of audiovisual dialogue was confirmed by Quaglio's (2009) MD study of *Friends* and by Zago (2016) using the same model to investigate a diachronic corpus of original American films and their remakes,[8] while Werner (2021) documented an overall increase in emotionality and informality as expressed by several colloquial lexical and lexico-grammatical features in two large diachronic corpora of English telecinematic dialogue. Audiovisual dialogue also foregrounds multimodal semiotics, which assists in both the contextualisation of communication practices and learner-viewers' immersion.

Other contemporary television registers are more orientated towards other dimensions of Biber's (1988) MD model. Berber Sardinha and Veirano Pinto (2017) investigated 31 different television registers, such as drama series, sitcoms, soap operas, live politics, reality shows and news, and provided evidence of similarities and differences both among the television registers and with reference to offscreen registers. Whereas all the television registers were situated towards the involvement end of Biber's (1988) Dimension 1, soap operas, for example, had a far higher score compared to live sports broadcasts. Moreover, political broadcasts and spontaneous speech were very similar in terms of Dimension 3, *explicit* versus *situation-dependent reference*, as they both draw on context through language. A later analysis (Berber Sardinha, Veirano Pinto 2021) revealed that television programmes can be grouped into clusters that share similar dimensional profiles. The resulting typology comprises nine different types of texts,

---

MD studies have employed the statistical procedure of factor analysis (see also Section 4) to identify the salient linguistic co-occurrence patterns that underlie registers and link them to situational, social and cognitive functions. The resulting dimensions correspond to communicative functions in text. Biber (1988) originally identified five major dimensions of register variation: *involved* versus *informational production*, *narrative* versus *non-narrative concerns*, *explicit* versus *situation-dependent reference*, *overt expression of persuasion* and *abstract* versus *non-abstract information*. These dimensions have enabled registers to be explored and degrees of similarity and differences across registers to emerge. The original MD model has been variously integrated and modified in subsequent MD analyses of media registers (see the references in the present Section).

[8] See Forchini (2021) for additional evidence pertaining to various film genres.
.

including texts that mainly present information, convey opinions or employ playful discourse. This brief overview shows how realistic, rich and varied the language of film and television programmes is. These characteristics make it a good model for the acquisition of spoken language in its various realisations.

Using personal computers, mobile telephones and iPads, L2 speakers can access English web products, the availability of which has now increased due to the spread of streaming and Internet-based video platforms, as well as the ever-increasing portability of smart devices. These technological innovations have extended the range of media genres, language varieties and registers with which L2 users can engage online. Overall, web-based registers are hybrid, multi-functional, adjustable and changeable (Page *et al.* 2022). As they vary in terms of audience design, users' participation framework, types of interactivity, and levels of accessibility and complexity, they present a variety of linguistic, multimodal and rhetorical features that lead to highly variable input. As a result, Internet registers make up broader and more diverse categories than do traditional registers (Biber, Egbert 2018). Moreover, while some web registers share many linguistic features with their non-web counterparts (e.g. editorials, news reports, songs and encyclopaedia pages), other web-specific genres, such as online news, travel blogs, discussion forums and question-answer forums, show specific linguistic and generic features (Biber, Egbert 2016; Lewis 2003). For example, discussion forums are characterised by specialised lexical items including *imo*, *FAQ*, *re*, *posts*, *threads* and *clicking* (Laippala *et al.* 2021, p. 782). Furthermore, some web registers, such as song lyrics and sports reports, can be identified more easily due to their lexical and grammatical consistency within and across texts; others, including opinion blogs and travel blogs, show greater variation, with lexical and grammatical choices changing according to the varying situational parameters (Laippala *et al.* 2021).

Among the most popular and potentially influential web registers, social media posts are characterised by dialogic and interactional quality, and are orientated towards the expression of stance and interpersonal exchanges of information (Berber Sardinha 2018). Since they activate an increased range of multimodal resources – including emojis and emoticons – and are realised as short texts, social media may be easier for L2 users to access compared to longer, monosemiotic texts. However, comprehension issues may arise due to the widespread use of non-standard spelling, punctuation, abbreviations and hashtags (Berber Sardinha 2022, p. 657). In an MD study of Instagram, Facebook and Twitter, Berber Sardinha (2022) further emphasised the relevance of individual variations across the different user platforms and categories (e.g. politicians and celebrities). Individual users tended to opt for either one of the two dimensions of variation identified,

namely *formal, informational, argumentative discourse* versus *informal, interactive, speaker-oriented discourse*, while simultaneously exploiting the rhetorical richness of the media register – a key feature of optimal input for SLA (Long 2020). In summary, Internet registers offer extremely rich and highly diversified language input that can be exploited for informal language learning in various ways in the digital wild.

From a varilectal and multicultural perspective, it is important to note that social media, social platforms and new media readily allow for the combination of semiotic resources, including different languages and language varieties, polylanguaging (Jørgensen *et al.* 2011) and translingualism (Canagarajah 2012). The implications for informal language learning of the widespread presence of non-native Englishes in the media deserve in-depth investigation.

2.3.1. Media input and second language acquisition

Systematic exposure to online and offline registers is likely to contribute to SLA in several ways, in line with the theoretical premises outlined above. Different media registers will collectively add to the amount of representative language data that is necessary for language learning to occur, and will allow L2 speakers to draw generalisations implicitly, to activate statistical learning and to extract form-function mappings from the target language. As instances of specialised discourses, these registers can also offer evidence of distinctive type and token frequencies while providing instances of contextualised language and text variation. When their linguistic composition is similar to unplanned, unmediated language, media registers will also contribute to the development of general L2 proficiency, as well as to the acquisition of registers that are less readily accessible in English as a foreign language (EFL) settings. In turn, interactional exchanges in social media and onscreen can assist L2 users to unpack complex input while adding salience to specific L2 features.[9]

Some studies have shown that informal access to media has a positive effect on L2 proficiency (Berns *et al.* 2007; Cole, Vanderplank 2016; Verspoor *et al.* 2011). The most robust evidence at present is Azzolini *et al.*'s (2022) investigation of a representative sample of approximately 20,000 school learners from 14 European Union countries.[10] The participants were tested on three language skills, while their informal experience with English was investigated and correlated with other relevant variables. The findings

---

[9] The same acquisitional advantages will apply to multiplayer video games (see Bianchi this volume).

[10] Italy was not included in the study.

showed that the association between exposure to English-language media and English-language competence was particularly sizeable. Informal experience with English – particularly via cultural artefacts and the media – was shown to play a key role in students' L2 proficiency, noticeably exceeding that of formal instruction and other significant variables such as language distance (Azzolini *et al*. 2022, p. 167).

Large-scale empirical research on informal language learning has also suggested that some media access produces better linguistic outcomes with regard to specific language skills (e.g. De Wilde *et al*. 2020; Krüger 2023; Muñoz 2020; Winke, Gass 2018). A national project involving the whole of Iceland (Arnbjörnsdóttir, Ingvarsdóttir 2018) showed that L2 learner-users benefitted strongly from accessing English-language television and the Internet outside of formal settings, as they became highly proficient conversationalists in English. However, they lagged behind considerably in formal academic reading and writing skills, as their out-of-class exposure to these language skills was limited. Careful mappings of informal exposure can provide the necessary information to predict what learner-users will acquire spontaneously through the media.

Studies of informal learners' uptake of specific English-language features have also shown that web-based and audiovisual input affected the acquisition of register-specific features. Sockett and Kusyk (2015) reported that frequent L2 viewers of television series produced more idiomatic texts in English and knew the most pervasive 4-grams in an *ad hoc* compiled corpus of popular television series. From the perspective of interlanguage pragmatics, Lyrigkou's (2023) investigation revealed that English discourse markers, which were found to be absent from classroom instruction, were successfully acquired by Greek school learners who accessed L2 media informally. Overall, however, fine-grained research on SLA-oriented analyses of media registers is limited, and studies that correlate specific media engagement in English with learner-users' competence development are scarce (but see De Riso this volume; Kusyk 2020). As a result, the PRIN project aims to describe the salient features of relevant media discourses to contribute to characterising the multifaceted multimodal and web-embedded English input to which Italian users are informally exposed (see works by Bianchi, Formentelli and Zago, and Galiano this volume). Moreover, the findings pertaining to media registers will inform the choice of features to be investigated in the longitudinal studies of individual high-exposure participants (see Section 3.2 for further details).

2.4. Learners' attitudes, perceptions and motivations

The ubiquity of English and the changing language learning landscapes have strong implications for the ways in which L2 users conceptualise English and English language learning (Aiello 2018; Ferguson 2015; Gnutzmann *et al.* 2015). In their study involving several European countries, Berns *et al.* (2007) showed that not only was there a positive correlation between school learners' exposure to English-language media and their language performance, but their attitudes towards English were also affected positively by out-of-school contact with the target language. Accordingly, the PRIN project started from the assumption that attitudes towards the language and its speakers would both predict and result from language use and language learning.

Language attitudes can generally be defined as evaluative reactions to languages and language varieties (Dragojevic 2017); more precisely, they identify affective, cognitive and behavioural stances towards the speakers of those languages and language varieties (Ryan *et al.* 1982, p. 7). Consequently, the three components of attitudes that need to be examined are feelings, beliefs and activities (Kircher, Zipp 2022). Language attitudes also need to be considered in conjunction with learners' perceptions of their L2 selves, their goals and reasons for accessing English, and their motivations for acquiring the language (Dörnyei, Ushioda 2009; Lee, Xie 2023). For example, Schneider (2020, p. 225) pointed out that, whereas extensive evidence testifies to the role of instrumental motivation in the spread of English worldwide, less research evidence supports a similar role of integrative motivation.

Since attitudes and motivation develop based on experience (Kircher, Zipp 2022), we expect informal contexts to influence the ways in which L2 speakers view themselves as learners or users of the language (Aiello 2018). In a case study by Pavesi and Ghia (2020), many of the 305 Italian university students who were surveyed stated that they watched audiovisual products in the L2 because they wanted to improve their English, hence revealing a language learning motivation for access. However, as the respondents added that they enjoyed the aesthetic, narrative and sociable dimensions of the audiovisual products, their identity as language users − rather than as language learners − also emerged quite strongly. In fact, changes in language practices over time and space may determine identity shifts (Eskildsen, Theodórsdóttir 2017; Ishikawa 2017; Sockett 2023). According to Mauranen (2018), traditional learners of English progressively become ELF users, who do not identify with models of English as a native language (ENL), presumably as a result of their greater experience with the language outside of the classroom.

L2 learner-users' orientations towards English will be influenced by several factors, such as EMI and living abroad experiences, intrinsic liking of the language, as well as frequency and intensity of contact with native and non-native varieties of English (De Wilde 2020; Modiano 2024). Audiovisual texts such as television series, films and stand-up comedies will lean more towards the use of native, mainly northern American, varieties, while blogs, forums and YouTube videos may contain more instances of ELF and texts in second language varieties. As Schneider (2020, p. 235) highlighted, the Internet plays a considerable role in spreading aspects of World Englishes outside of their regions of origin.

These initial considerations call for a wider-scoped, deeper research investigation to account for all the subtleties in the evolving attitudes and stances towards English and informal English language learning.

3. The investigation

3.1. Research questions

Starting from these premises, the project appraises how declared and more hidden engagements with English media are currently affecting the ways in which Italian university students approach the learning and use of the global language. For maximal substantive representativeness, the investigation focuses on four middle-sized/large universities in the North, Centre, South and a major island of Italy. The Universities of Pavia, Pisa, Salento and Catania are long-standing, multidisciplinary institutions of higher education that reflect diverse student populations.

A series of research questions is addressed in the project. The six main ones are:

1. To what extent do the sampled Italian students engage with English for leisure through the media? Which modalities of access (e.g. receptive or interactive) and affordances (e.g. subtitles and online communities) are favoured?

2. What correlations are there among the frequency, intensity and type of informal contact with English and major individual and social variables, such as self-assessed L2 competence, disciplinary area, attendance at secondary school and gender?

3. Does one single correlation and interpretation model apply to the four sample universities? If not, according to which factors do the groups differ?

4. What are the attitudes towards the English language, informal language learning and various modalities of access to English? What are the relationships between English media exposure and orientations towards

ENL/EFL/ELF? Can associated L2-identity issues be identified?

5. What are the results of informal media experience for high-exposure learner-users of English over time? Can both individual learning trajectories and common traits be identified?

6. How can L2 higher education be reframed to bridge formal and informal learning and to foster students' self-regulated learning?

3.2. Methodology and research stages

To address these research questions, we adopted a mixed-methods research design that combines quantitative and qualitative research methodologies, together with collections of cross-sectional and longitudinal data. The research design includes empirical tools that balance etic and emic perspectives of data analysis. Such a complex approach will decrease the limitations of each individual procedure and allow for the triangulation of the results.

The project is structured into three major phases. In Stage 1, a fine-grained questionnaire about university students' contact with English has been constructed, piloted, validated and administered online to students in intact classes to avoid self-selection issues (Pavesi *et al.* 2023). Questionnaires have been widely used in research on informal language learning to collect behavioural and attitudinal data from large samples of L2 learner-users (e.g. Krüger 2023; Kusyk 2020; Muñoz 2020; Pavesi, Ghia 2020; Toffoli, Sockett 2010). The Informal English Contact and Learning (IECoL) questionnaire contains 72 questions and is divided into three macrosections. The first section gathers general information about the participants' language backgrounds and previous linguistic experiences; the second section focuses on the frequency and intensity of exposure to various input types, with subsections pertaining to audiovisual products, YouTube videos, video games, songs, other Internet resources (e.g. blogs, forums, web pages, podcasts and apps), reading practices (e.g. newspapers, books and magazines) and additional input sources such as online video calls and face-to-face interactions. The third macrosection comprises queries about attitudes towards English, self-perceived proficiency levels and social backgrounds. IECoL has been linked to a vocabulary test to obtain first-hand learning data (Pavesi *et al.* 2023). The final sample consisted of approximately 4,000 valid questionnaires and tests that were distributed equally at the four universities to students specialising in the humanities, social sciences and science and technology in comparable proportions. The vast majority of the participants were non-specialist learners of English who were expected to access the L2 in informal contexts independently of language instruction. However, a small number of language students was also included to provide a view of learners who were still learning English in the classroom and had an "intense and

proactive approach to the learning of the target language" (Sockett, Kusyk 2015, p. 163). The identification of disciplinary areas ensured better comparability among the target universities, while also enabling the evaluation of the behaviour of students with different academic backgrounds.

Following the first stage, each unit is pursuing its own research agenda characterised by individual objectives that integrate with each other to provide a broader picture of the different aspects that are involved in informal contact with English (Figure 1). In Stage 2, semi-structured interviews are being conducted with selected (high-, medium- and low-exposure) students who are being asked to reflect on various aspects of their relationship with L2 English and their informal engagement with English-language media (Dressman 2023; Kusyk 2020). Due to the private nature of most informal learning, ethnographic studies using semi-structured interviews were deemed necessary to obtain in-depth opinions and long-term accounts of behavioural patterns, language beliefs, affiliations with communities of practice and affinity groups, together with motivations for informal exposure to English (Kusyk *et al.* 2023). A general difficulty has been noticed when measuring L2 viewers' "orientations towards learning or leisure-based motivations" in survey-based studies if follow-up qualitative research on participants' attitudes and their processing of verbal input is not conducted (Sockett 2014, p. 96). Therefore, Stage 2 in the present project is a response to Kusyk *et al.*'s (2023, p. 5) call for research that focuses on individual acquisitional pathways, as well as the personal and social variables that are associated with individual learner-users.

In Stage 3, longitudinal case-studies of individuals' development in selected language areas are carried out to provide access to the temporal dimension of media-triggered informal learning and to allow for more thorough analyses of individual learning trajectories. As SLA is a process that requires prolonged and extensive engagement with the L2, investigating SLA is best accomplished through repeated samplings, observations and measurements at regular intervals across extended periods (Ross, Masters 2023). The aim here is not the generalisability of the results, but rather the specificity of detail pertaining to individual learning trajectories (Ross, Masters 2023). In this stage, the acquisition of morpho-syntactic, pragmatic and lexical features, such as speech acts and specialised vocabulary, is investigated. Centre stage is occupied by the three constructs of complexity, accuracy and fluency (CAF) to be analysed individually and in interaction to provide details regarding the multifacetedness and dynamicity of L2 development over time (Hasnain, Halder 2024; Michel 2017; Pallotti 2021). For example, CAF measures allow researchers to compare L2 learners' productions in different tasks (e.g. written compositions versus oral narrations), thus capturing different aspects of learners' language (Skehan, Foster 2008).

The PAVIA RU will coordinate the development and implementation of the research design and will:

- Discuss and contextualise the main issues in informal learning of L2 English

- Provide corpus-based descriptions of spoken language in audiovisual (AV) dialogue

- In the interviews, focus on comprehension of (subtitled) AV input and other cognitive processes: e.g., attentional processes, interaction; negotiation of meaning; participants' attitudes towards English – ENL, EFL, ELF, their identities as learners/users and perception of L2-self/selves [interviews]

- In the longitudinal studies, investigate Complexity, Accuracy, Fluency (CAF) trajectories with a focus on Complexity and Accuracy from a grammatical and lexical perspective in spoken and written production [longitudinal studies]

The CATANIA RU will:

- Review the literature on the development of phonology and speakers' attitudes towards English sociolinguistic varieties in the media

- Provide descriptions of media accents and corpus-based analyses of spoken language in AV dialogue

- Focus on participants' stance on English varieties, social and regional accents; participants' responses to English AV input to gauge attitudes towards features of spoken English [interviews]

- Investigate CAF with a focus on Accuracy in pronunciation and Fluency in spoken and written production [longitudinal studies]

Stage 1: Questionnaire + test
Stage 2: Semi-structured interviews
Stage 3: Longitudinal studies + Input description

The PISA RU will:

- Review the literature on incidental vocabulary learning

- Focus on linguistic and multimodal features of web-based genres

- Focus on sources of media exposure, vocabulary size and depth, awareness of vocabulary learning strategies, multimodal input in vocabulary learning [interviews]

- Evaluate lexical knowledge in subsets of participants highly exposed to web-based genres in terms of CAF. The data will be analysed within the Dynamic Systems Theory approach to second language acquisition. Attention will be paid to slang and ESP lexis [longitudinal studies]

The SALENTO RU will:

- Review of the literature on social media interaction by investigating linguistic and multimodal features of specific media genres and gaming

- Examine informal and autonomous learning, cognitive, affective and social parameters, self-efficacy, (meta)cognitive strategies, motivation, learner's beliefs; highlight the interaction between formal and informal learning [in the interviews]

- Assess the acquisition of pragmatic competencies by focussing on speech acts and testing high-exposure participants' pragmatic performance, awareness and error recognition [longitudinal studies]

- Propose a methodological framework for integrating informal English contact with formal teaching in higher education

Figure 1
Outline of the specific goals of the four research units.

Overall, moving from cross-sectional surveys to interviews and longitudinal investigations of individual participants is deemed to be crucial to fully depict the ecology of informal approaches to English at present. "Every digital learner of English has a story to tell about 'how they did it'. Those stories are often impressive and share both commonalities and peculiarities" (Dressman 2023, p. 29). At the core of the research output, the emic accounts will integrate the quantitative results and provide insights into learner-users' language behaviours, language attitudes and evolving L2 competencies.

4. L2 users' profiles: Patterns of informal contact with English

An illustration of the research potential of the PRIN project is provided in the final part of this contribution. An overview of the ways in which practices of informal contact with English typically co-occur in the surveyed learner-users' experiences is presented; that is, informal practices that tend to be carried out by the same participants and are shared across many learner-users can be identified.[11] This profiling operation can also be seen as a first step in the mapping of the informal learning spaces that learner-users create for themselves in the wild, in line with the spatial perspective that is adopted in the project.

To conduct the investigation, the individual replies from the IECoL questionnaire were extracted and subsequently grouped together using a factor analysis, which is a statistical procedure that brings together variables that are strongly correlated while setting apart those that are not, hence identifying underlying traits. In this process, a factor analysis clusters several variables into more general, higher-level factors that help to identify the latent structure in the data (Dörnyei 2007, pp. 233-236). In the present investigation, a factor analysis was used to extract the informal learner-users' profiles. In other words, a factor analysis was functional to combining different input sources and identifying which ones tend to be accessed by the same learners when they engage with English outside of traditional learning settings.

To conduct the factor analysis, an index of exposure was calculated for each individual respondent[12] who had submitted a completed questionnaire, resulting in 2,525 questionnaires in total. The selection of a factor analysis with five factors revealed that distinctive clusters of types of English-language input could be identified which distinguish among English-based practices (Table 1).[13] These draw on the Internet or more traditional media, involve a greater or lesser degree of receptive use, feature multimodality in texts, activate interactivity and so forth.

---

[11] By correlating favourite activities, Berns *et al*. (2007) isolated different learners' profiles among informal learners of English across several European countries. More recently, Lee and Xie (2023) discussed types of recurrent informal exposure using a large sample of Korean students.

[12] Indexes of exposure were obtained following Pavesi and Ghia's (2020, pp. 81-83) calculation procedure.

[13] The applied method was maximum-likelihood factor analysis. The rotation method was promax with power 4. The Kaiser-Meyer-Olkin factor adequacy was meritorious (overall value = 0.86). The factors showed good to excellent levels of internal consistency, with values of alpha varying from 0.71 to 0.92. The analysis was performed using R 4.2.0.

Factor 1 – Receptive Internet access		Factor 2 – Multimodal input		Factor 3 – Proactive Internet use		Factor 4 – Personal contact		Factor 5 – Video games	
Internet_I read socials	*1.05*	TV-series	*0.94*	Internet_I write blogs	*1.04*	Chats	*0.76*	Gaming alone	*0.89*
Internet_I read webpages	*1.02*	Films	*0.94*	Internet_radio	0.73	Online face-to-face	*0.75*	Group video games	*0.83*
Internet_app	*0.99*	Other TV programmes	0.58	Internet_I write on socials	0.66	Email	0.65		
Internet_I read blogs	0.74	YouTube	0.36	Internet_podcast	0.34	In presence	0.60		
Internet_shopping	0.71	Songs	0.31	Internet_other	.030				
Internet_podcast	0.42								
Internet_other	0.44								

Table 1
Five factor clusters for informal English input.[14]

The first factor (Factor 1) identifies online, web-based practices through which learner-users mainly access English-language input receptively. They read content on social networks and webpages, read blogs, use apps (sometimes to make purchases online) and occasionally listen to podcasts. This factor can be labelled 'Receptive Internet access' because users experience English online but mainly in one-way communication, occasionally becoming involved in human-non-human interactions. Importantly, the receptive online affordances in this factor mainly provide written language input, as listening practices in English are correlated less strongly with the other language affordances in the factor. A second major group of correlated activities in English includes watching original language films and television series, which are the two types of input that contribute most to this factor. Significantly, this second factor also includes access to other audiovisual programmes, as well as to YouTube resources and songs, although their weight is lower than that of telecinematic input. Affordances in this factor provide spoken input and are markedly multimodal as they are characterised by the co-deployment of several semiotic resources beyond verbal language, such as images, gestures, movements, gazes and music. Therefore, the second factor can be labelled 'Multimodal input'. Telecinemantic input represents a separate category regardless of the support on which it relies in that it consists of full-length audiovisual narratives in contrast to generally shorter YouTube video clips and equally short songs. Factor 2 also mainly corresponds to receptive contact with English, namely

---

[14] The five factors are shown separately and are ranked by loading, with highest values, or weight, per cluster in italics.

one-way access to language practices onscreen or song lyrics, but differs noticeably from Factor 1 because exposure here is predominantly to spoken language. Due to their strengths, the first two factors capture the respondents' tendency to experience English via screens, often online and with very limited, if any, two-way communication.

Active engagement emerges via Factor 3, which identifies practices whereby learner-users play a more dynamic role on the web, such as when they write blogs or write on social media. This factor can be labelled 'Proactive Internet uses' because it concerns students who like to become directly involved online often through writing practices. We also found access to English through radio stations and podcasts in this factor; radio stations and podcasts are posited as being active listenership given the strong correlations among the variables. Moreover, as comprehension in these modalities relies on linguistic and paralinguistic components, greater effort and involvement is required on the part of the L2 listeners. Factor 4 is also characterised by strong involvement because it combines various types of interactions, such as online chats, online video calls or video conferences, 'person-to-person interactions', as well as emails; that is, language experiences that encompass two-way communication with other people. These practices cut across online and offline distinctions by identifying L2 users who like talking, chatting and generally having exchanges through 'personal contact', hence the name assigned to this factor (see Verspoor *et al.* 2011). Finally, all gaming activities – comprising both single and group video games – are included in Factor 5. These activities have the potential for considerable interactivity, and involve speakers of many native and non-native varieties of English (Bianchi this volume). As no other practices were loaded in this factor, 'Video games' in English appeared to stand alone and to have a weak correlation with other types of engagement with the L2 outside of the classroom.

The five factors that emerged from the analysis can be used to identify groups of prototypical informal users of English: those who like to access the web in English, to read content and to surf the net (Factor 1); viewers of English-language audiovisual programmes and YouTube videos (Factor 2); proactive Internet users who enjoy being involved in writing texts and listening to specific content online in English (Factor 3); and students who use the language in personal interactions, both face-to-face and online (Factor 4). Video gamers stand out due to focusing on that activity, both individually or in groups, but to the exclusion of other informal practices in English (Factor 5). The last trend was evidenced by a further analysis of the correlations among the five factors (Table 2).

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1.00				
Factor 2	0.29	1.00			
Factor 3	0.21	0.33	1.00		
Factor 4	0.69	0.21	0.11	1.00	
Factor 5	- 0.44	- 0.50	- 0.29	- 0.40	1.00

Table 2
Correlation coefficients among factors.

Noticeably, Factor 5 correlated negatively with all the other factors, a statistical trend that once again suggests that students who like to play video games in English tend to become absorbed in that activity while neglecting other opportunities for engaging informally with English. In this regard, students who mainly navigate the net, access YouTube and social networks, or watch films and television series in English – but who also engage in face-to-face interactions – tend not to play video games in English

On the whole, the positive, although often relatively weak correlations among and across the other four factors indicated that students access a variety of input types in English. However, they do so to different extents. One interesting, strong correlation between 'Receptive Internet access' (Factor 1) and 'Personal contact' (Factor 4) was noted, suggesting that students who mainly like to surf the net to read and, less often, to listen to online texts, reserve their interactional practices for more direct, personal encounters, both in person and online. Overall, based on the analysis, five prototypical profiles of informal learner-users of English have emerged, who favour but do not restrict themselves to the preferred type of contact with the L2, as summarised in Figure 2.
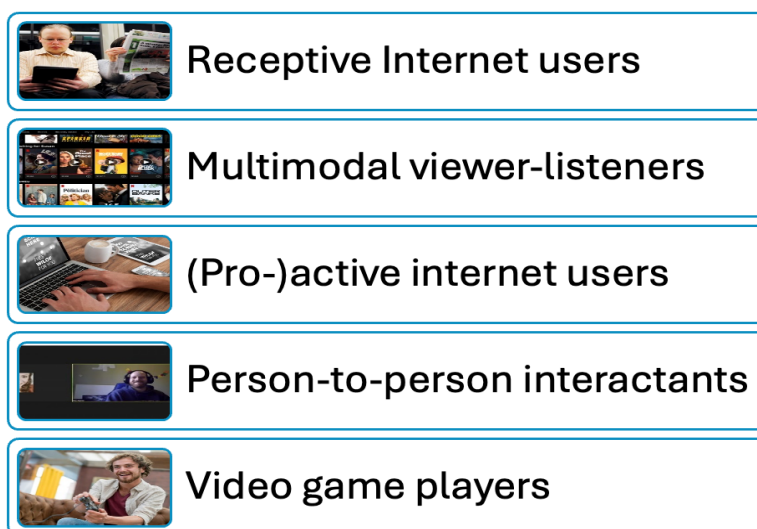
Figure 2
Access profiles of informal L2 English users.

Depending on their access profile, learner-users are expected to engage with diverse discoursal registers in English that are collocated differently on the written-spoken language continua and involve varying degrees of active production and interaction. These profiles are expected to be related to the ways in which learner-users' develop their attitudes towards English and their competence in the L2, as is being investigated in the current stages of the research project.

5. Conclusions

We are witnessing a revolution in the learning of English across the globe. English-language affordances are multiplying and are becoming easily accessible through a wealth of media for the benefit of the receptive, curious and adventurous L2 user. The increasing phenomenon of informal access to English first pertains to individuals, but clearly extends to groups, communities and society at large. As a result, learner-users' media-triggered self-directed contact with English exhibits both private and shared dimensions, all of which are bound to have a dramatic impact on L2 speakers' developing competence at present and in the future. Starting from these considerations and in the wake of recent research in Europe, the PRIN project regarding the informalisation of English language learning through the media is being developed with the overarching aim of mapping Italian university students' informal experiences with the L2 at a time at which they have mainly completed their formal education in English. In the previous sections, the descriptive and theoretical background to the national investigation was outlined, together with the three main phases that were envisaged in the data collection and analysis. To obtain a preliminary overview of the participants' behavioural patterns, a factor analysis was conducted on the approximately 2,500 fully completed questionnaires. Five prototypical profiles of informal learner-users emerged, corresponding to groups of language affordances and activities that tend to occur together in individuals' experiences with English. The detailed patterns of students' behaviour, together with the personal, social and language variables that are associated with these patterns – including attitudes and motivations, disciplinary areas, and the language input provided by each media register – have been explored in the various phases of the research project and are discussed in the contributions to the present volume. Many challenges have been encountered in the initial stages of the study and are being addressed as they emerge in the day-to-day management of the complexities that are intrinsic to research involving real people. Some changes to the initial plan needed to be made to accommodate the shifting contexts and the different

settings at the four universities. Despite the adjustments required by the socially engaging and participatory nature of the investigation, we trust that, by the end of the project, we will have obtained a broader and more insightful perspective on the rapidly changing relationships that university students experience with English in Italy at present.

Bionote: Maria Pavesi is professor of English Language and Linguistics at the Department of Humanities of the University of Pavia. Her research has focused on second language acquisition, the English of science, English language learning through the media, film language and corpus-based audiovisual translation studies. In all these fields she has published widely, both nationally and internationally. She recently co-authored the volume *Informal Contact with English. A Case Study of Italian Postgraduate Students* (ETS, 2020, open access), and "The pragmatic dimensions of swearing in films: Searching for coherence in dubbing strategies", *Journal of Pragmatics* (2023). Maria Pavesi has taken part in several government-funded (MURST/PRIN), EU and international research projects on language corpora, multimodality, language learning and teaching, linguistic analysis and translation. She is currently the Principal Investigator of the national PRIN project titled "The informalisation of English language learning through the media: Language input, learning outcomes and sociolinguistic attitudes from an Italian perspective", https://prin.unipv.it/index.php/en/.

Author's address: maria.pavesi@unipv.it

References

Aiello J. 2018, *Negotiating Englishes and English-speaking Identities*, Routledge, London/New York.

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development Across the Life Span: The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham.

Azzolini D., Campregher S. and Madia J.E. 2022, *Formal instruction vs informal exposure. What matters more for teenagers' acquisition of English as a second language?*, in "Research Papers in Education" 37 [2], pp. 153-181.

Benson P. 2021, *Language Learning Environments: Spatial Perspectives in SLA*, Multilingual Matters, Bristol.

Benson P. 2022, *Mapping language learning environments*, in Reinders H., Lai C. and Sundqvist P. (Eds.), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, London/New York, pp.24-51.

Berber Sardinha T. 2018, *Dimensions of variation across Internet registers*, in "International Journal of Corpus Linguistics" 23 [2], pp. 125-157.

Berber Sardinha T. 2022, *Corpus linguistics and the study of social media: A case study using multi-dimensional analysis*, in O'Keeffe A. and McCarthy M. (Eds.), *Routledge Handbook of Corpus Linguistics*, 2nd ed., Routledge, London/New York, pp. 656-674.

Berber Sardinha T. and Veirano Pinto M. 2017, *American television and off-screen registers: a corpus-based comparison*, in "Corpora" 12 [1], pp. 5-114.

Berber Sardinha T. and Veirano Pinto M. 2019, *Introduction*, in Berber Sardinha T. and Veirano Pinto M. (Eds.), *Multi-dimensional Analysis. Research Methods and Current Issues*, Bloomsbury, London/New York, pp.1-8.

Berber Sardinha T. and Veirano Pinto M. 2021, *A linguistic typology of American television*, in "International Journal of Corpus Linguistics" 26 [1], pp. 127-160.

Berns M., de Bot K. and Hasebrink U. 2007, *In the Presence of English: Media and European Youth*, Springer, New York.

Bianchi F. this volume, *Video game dialogue as input for language learning. An analysis of lexical coverage*.

Biber D. 1988, *Variation across Speech and Writing*, Cambridge University Press, Cambridge.

Biber D. and Egbert J. 2016, *Register variation on the searchable web: A multi-dimensional analysis*, in "Journal of English Linguistics" 44 [2], pp. 95-137.

Biber D. and Egbert J. 2018, *Register Variation Online*, Cambridge University Press, Cambridge.

Canagarajah S. 2012, *Translingual Practice: Global Englishes and Cosmopolitan Relations*, Routledge, London/New York.

Cole J. and Vanderplank R. 2016, *Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning*, in "System" 61, pp. 31-42.

Crystal D. 2006, *Language and the Internet*, 2nd ed., Cambridge University Press, Cambridge.

de Bot K. 2014, *The effectiveness of early foreign language in the Netherlands*, in "Studies in Second Language Learning and Teaching" 4 [3], pp. 409-418.

De Riso C. 2023, *Out-of-the-classroom English input and informal second language acquisition: A longitudinal study of Italian university students*, Unpublished doctoral thesis, University of Pavia.

De Riso C. this volume, *Complexity accuracy and fluency in informal L2 learning among university students. Four longitudinal case studies*.

De Wilde V., Brysbaert M. and Eyckmans J. 2020, *Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important?*, in "Bilingualism: Language and Cognition" 23 [1], pp. 171-185.

De Wilde V., Brysbaert M. and Eyckmans J. 2021, *Young learners' L2 English after the onset of instruction: Longitudinal development of L2 proficiency and the role of individual differences*, in "Bilingualism: Language and Cognition" 24 [3], pp. 439-453.

Dörnyei Z. 2007, *Research Methods in Applied Linguistics*, Oxford University Press, Oxford.

Dörnyei Z. and Ushioda E. (Eds.) 2009, *Motivation, Language Identity and the L2 Self*. Multilingual Matters, Bristol.

Dragojevic M. 2017, *Language attitudes*, in Gile H. and Maass A. (Eds.), *Oxford Research Encyclopaedia of Intergroup Communication*, vol.2, Oxford University Press, Oxford, pp.179-192.

Dressman M. 2023, *Towards an anthropology of Informal Digital Learning of English (IDLE)*, in Toffoli D. Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure. Informal Language Learning in the Digital Age*, Mouton de Gruyter, Berlin/Boston, pp. 21-42.

Dressman M. and Sadler R.W. (Eds.) 2020, *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Ellis N.C. 2003, *Constructions, chunking and connectionism*, in Doughty C. and Long M. (Eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, Oxford, pp. 63-103.

Ellis N.C. 2015, *Cognitive and social aspects of learning from usage*, in Cadierno T. and Eskildsen S.W. (Eds.), *Usage-based Perspectives in Second Language Learning*, Mouton de Gruyter, Berlin, pp. 49-73.

Ellis N.C. and Collins L. 2009, *Input and second language acquisition: The roles of frequency, form and function*, in "The Modern Language Journal" 93 [3], pp. 329-335.

Eskildsen W. and Theodórsdóttir G. 2017, *Constructing L2 learning spaces: Ways to achieve learning inside and outside the classroom*, in "Applied Linguistics" 38 [2], pp. 143-164.

Ferguson G. 2015, *Introduction: Attitudes to English*, in Linn A., Bermel N. and Ferguson G. (Eds.), *Attitudes Towards English in Europe*, Mouton de Gruyter, Berlin, pp. 3-24.

Forchini P. 2012, *Movie Language Revisited. Evidence from Multi-dimensional Analysis and Corpora*, Peter Lang, Bern.

Forchini P. 2021, *Movie discourse. Marvel and DC Studios compared*, in Friginal E. and Hardy J. (Eds.), *The Routledge Handbook of Corpus Approaches to Discourse Analysis*, Routledge, London/New York, pp.185-201.

Formentelli M. and Zago R. this volume, *Complexity matters in TV dialogue as language input. A corpus-based description of clausal and phrasal patterns*.

Galiano L. this volume, *Noun phrase complexity in film dialogue as input for second-language learning. A corpus-based study from a register-functional perspective*.

Gass S.M. 2013, *Input, Interaction, and the Second Language Learner*, Routledge, London.

Gass S.M. and Mackey A. 2020, *Input, interaction, and output in L2 acquisition*, in VanPatten B., Keating G.D. and Wulff S. (Eds.), *Theories in SLA. An Introduction*, Routledge, London/New York, pp. 192-222.

Ghia E. 2012, *Subtitling Matters. New Perspectives on Subtitling and Foreign Language Learning*, Peter Lang, Oxford.

Gnutzmann C., Jakis J. and Rabe F. 2015, *Communicating across Europe. What German students think about multilingualism, language norms and English as a lingua franca*, in Linn A.R., Bermel N. and Ferguson G. (Eds.), *Attitudes Towards English in Europe*, Mouton de Gruyter, Berlin, pp. 165-191.

Hasnain S. and Halder S. 2024, *Intricacies of multifaceted triad – complexity, accuracy, and fluency: A review of studies on measures of oral production*, in "Journal of Education" 204 [1], pp. 145-158.

Ishikawa T. 2017, *Japanese university students' attitudes towards their English and the possibility of ELF awareness*, in "Journal of English as a Lingua Franca" 6 [2], pp. 237-263.

Jørgensen J.N., Karrebæk M.S., Madsen L.M. and Møller J.S. 2011, *Polylanguaging in superdiversity*, in "Diversities" 13 [2], pp. 23-37.

Kircher R. and Zipp L. 2022, *An introduction to language attitudes research*, in Kircher R. and Zipp L. (Eds.), *Research Methods in Language Attitudes*, Cambridge University Press, Cambridge, pp. 1-16.

Krashen S. 1985, *The Input Hypothesis: Issues and Implications*, Longman, London.

Krashen S. 2008, *The Comprehension Hypothesis extended*, in Piske T. and Young-Scholten M. (Eds.), *Input Matters in SLA*, Multilingual Matters, Bristol, pp. 81-94.

Krüger M. 2023, *Media-Related Out-of-School Contact with English in Germany and Switzerland. Frequency, Forms and the Effect on Language Learning*, Springer, New York.

Kusyk M. 2020, *Informal English learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 333-348.

Kusyk M., Sockett G. and Toffoli D. 2023, *Introduction to language learning and leisure*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure. Informal Language Learning in the Digital Age*, Mouton de Gruyter, Berlin/Boston, pp. 1-18.

Laippala V., Egbert J., Biber D. and Kyröläinen A. 2021, *Exploring the role of lexis and grammar for the stable identification of register in an unrestricted corpus of web documents*, in "Language Resources and Evaluation" 55, pp.757-788.

Lantolf J.P., Poehner M.E. and Thorne S.L. 2020, *Sociocultural theory and L2 development*, in VanPatten B., Keating G.D. and Wulff S. (Eds.), *Theories in SLA. An Introduction*, Routledge, London/New York, pp. 223-247.

Lee J.S. and Xie Q. 2023, *Profiling the affective characteristics of EFL learners' digital informal learning: A person-centered approach*, in "Innovation in Language Learning and Teaching" 17 [3], pp. 552-566.

Leone P. 2023, *Teletandem. Apprendere le lingue in telecollaborazione*, Cassia Italia, Cesena/Bologna.

Lewis D.M. 2003, *Online news: A new genre?*, in Aitchison J. and Lewis D.M. (Eds.), *New Media Language*, Routledge, London/New York, pp. 95-104.

Lichtman K. and VanPatten B. 2021, *Was Krashen right? Forty years later*, in "Foreign Language Annals" 54, pp. 283-305.

Long M. 1996, *The Role of the linguistic Environment in second language acquisition*, in Ritchie W.C. and Bhatia T.K. (Eds.), *The Handbook of Second Language Acquisition*, Academic Press, San Diego, pp. 413-468.

Long M. 2020, *Optimal input for language learning: Genuine, simplified, elaborated, or modified elaborated?*, in "Language Teaching" 53 [2], pp.169-182.

Lyrigkou C. 2023*, Informal second language learning and EFL learners' spoken use of discourse markers*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure. Informal Language Learning in the Digital Age*, Mouton de Gruyter, Berlin/Boston, pp. 91-119.

Mauranen A. 2018, *Second language acquisition, world Englishes, and ELF*, in "World Englishes" 37 [1], pp. 106-119.

Mayer R.E. 2009, *Multimedia Learning*, 2nd ed., Cambridge University Press, Cambridge.

Michel M. 2017, *Complexity, accuracy and fluency (CAF)*, in Loewen S. and Sato M. (eds.), *The Routledge Handbook of Instructed Second Language Acquisition*, Routledge, London/New York, pp. 50-68.

Modiano M. 2024, *Identity and standards for English as a European Union lingua franca*, in "World Englishes" *43* [2], pp. 210-227.

Muñoz C. 2020, *Boys like games and girls like movies. Age and gender differences in out-of-school contact with English*, in "RESLA" 33 [1], pp. 171-201.

Muñoz C. and Cadierno T. 2021, How do *differences in exposure affect English language learning? A comparison of teenagers in two learning environments*, in "Studies in Second Language Learning and Teaching" 11 [2], pp. 185-212.

Page R., Barton D., Lee C., Unger J.W. and Zappavigna M. 2022, *Researching Language and Social Media. A Student Guide*, 2nd ed., Routledge, London/New York.

Paivio A. 1986, *Mental Representations: A Dual Coding Approach*, Oxford University Press, Oxford.

Pallotti G. 2021, *Measuring complexity, accuracy, and fluency (CAF)*, in Winke P. and Brunfaut T. (Eds.), *The Routledge Handbook of Second Language Acquisition and Language Testing*, Routledge, London/New York, pp. 201-210.

Pavesi M. and Ghia E. 2020, *Informal Contact with English. A Case Study of Italian Postgraduate Students*, ETS, Pisa.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English: An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Pérez-Paredes P., Mark G. and O'Keeffe A. 2020, *The impact of usage-based approaches on second language learning and teaching,* Cambridge Education Research Reports. https://www.cambridge.org/partnership/research/impact-usage-based-approaches-second-language-learning-and-teaching (11.03.2024).

Piazza R., Bednarek M. and Rossi F. (Eds.), 2011, *Telecinematic Discourse. Approaches to the Language of Films and Television Series*, John Benjamins, Amsterdam.

Quaglio P. 2009, *Television Dialogue. The Sitcom* Friends *vs. Natural Conversation*, John Benjamins, Amsterdam.

Reinders H., Lai C. and Sundqvist P. (Eds.), 2022, *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, London/New York.

Reinhardt J. 2022, *The History of learning and teaching beyond the classroom*, in Reinders H., Lai C. and Sundqvist P. (Eds.), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, London/New York, pp. 9-23.

Ross S.J. and Masters M.C. 2023, *Introduction: Longitudinal design and analyses in second language research*, in Ross S.J. and Masters M.C. (Eds.), *Longitudinal Studies of Second Language Learning. Quantitative methods and outcomes*, Routledge, London/New York, pp. 1-9.

Ryan E.B., Giles H. and Sebastian R.H. 1982, *An integrative perspective for the study of attitudes toward language variation*, in Ryan E. B. and Giles H. (Eds.), *Attitudes towards Language Variation. Social and Applied Contexts*, Edward Arnold, London, pp. 1-19.

Schneider E. 2020, *English around the World. An Introduction*, 2nd ed., Cambridge University Press, Cambridge.

Skehan P. and Foster P. 2008, *Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research*, in Van Daele S., Housen A., Kuiken F., Pierrard M. and Vedder I. (Eds.). *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*, Brussels, University of Brussel Press, pp. 207-226.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave Macmillan, London.

Sockett G. 2023. *Input in the digital wild: online informal and non-formal learning and their interactions with study abroad*, in "Second Language Research" 39 [1], pp. 115-132.

Sockett G. and Kusyk M. 2015, *Online informal learning of English: Frequency effects in the uptake of chunks of language from participation in web-based activities*, in Cadierno T. and Eskildsen S.W. (Eds.), *Usage-based Perspectives in Second Language Learning*, Mouton de Gruyter, Berlin, pp. 153-177.

Sundqvist P. and Sylvén L.K. 2014, *Language-related computer use: Focus on young L2 English learners in Sweden*, in "Re-CALL" 26 [1], pp. 3-20.

Swain M. 1995, *Three functions of output in second language learning*, in Cook G. and Seidlhofer B. (Eds.), *Principle and Practice in Applied Linguistics: Studies in Honour of H. G. Widdowson*, Oxford University Press, Oxford, pp. 125-144.

Toffoli D. and Sockett G. 2010, *How non-specialist students of English practice informal learning using Web 2.0 tools*, in "Asp. La revue du GERAS" 58, pp. 125-144.

Toffoli D., Sockett G. and Kusyk M. (Eds.) 2023, *Language Learning and Leisure. Informal Language Learning in the Digital Age*, Mouton de Gruyter, Berlin/Boston.

Tomasello M. 2005, *Constructing a Language*, Harvard University Press, Cambridge, MA.

VanPatten B, Williams J., Keating, G.D. 2020, *Introduction: The nature of theories*, in VanPatten B., Keating G.D. and Wulff S. (Eds.), *Theories in SLA. An Introduction*, Routledge, London/New York.

Verspoor M.H., de Bot K., and van Rein E. 2011, *English as a foreign language. The role of out-of-school language input*, in De Houwer A. and Wilton A. (Eds.), *English in Europe Today. Sociocultural and Educational Perspectives*, John Benjamins, Amsterdam, pp. 147-166.

Wagner J. 2015, *Designing for language learning in the wild: Creating social infrastructures for second language learning*, in Cadierno T. and Eskildsen S.W. (Eds.), *Usage-based Perspectives in Second Language Learning*, Mouton de Gruyter, Berlin, pp. 75-101.

Werner V. 2021, *A diachronic perspective on telecinematic language*, in "International Journal of Corpus Linguistics" 26 [1], pp. 38-70.

Werner V. 2022, *Pop cultural linguistics*, in Aranoff M. (Ed.), *Oxford Research Encyclopedia of Linguistics*, Oxford University Press, online. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-999 (8.7.2024).

Winke P. and Gass S.M. 2018, *Individual differences in advanced proficiency*, in Malovrh, P.A. and Benati A.G. (Eds.), *The Handbook of Advanced Proficiency in Second Language Acquisition*, Wiley Blackwell, Hoboken, USA, pp.157-178.

Zago R. 2016, *From Originals to Remakes. Colloquiality in English Film Dialogue Over Time*, Bonanno, Acireale.

WHAT DRIVES L2 VIEWERS' INFORMAL ACCESS TO ENGLISH?
A case study on audiovisual input

ELISA GHIA
UNIVERSITÀ DI PAVIA

Abstract – In recent years, several dubbing countries have experienced a change in the way audiovisual (AV) products and other media are accessed in terms of preferred languages and supports. A similar trend appears to be underway in Italy, where younger audiences increasingly access subtitled telecinematic materials and a variety of online media in English. While much research to date has focused on informal second language (L2) practices and the potential of media input for language learning, only a few studies have extensively examined participants' underlying reasons for exposure. To address this under-researched area, the present study investigates Italian university students' informal experience with various AV genres and the reasons behind their habits. The data form part of a larger-scale investigation (the PRIN project "The Informalisation of English Language Learning Through the Media") and were collected through a questionnaire administered to 1,091 students from various disciplines at a medium-sized Italian university. Viewers' reasons for accessing English-language audiovisuals are examined from both an etic and an emic perspective, based on multiple-choice and open-ended questions, and are complemented by data on the frequency and intensity of access, preferred genres, and viewing modalities. The findings reveal a desire to improve one's L2 skills and seek knowledge, but also a desire for entertainment and a growing appreciation of English, which enhances viewers' engagement with AV products. Additionally, accessibility affects viewers' choices, which differ across AV genres. The study prompts further reflection on the growing media affordances and evolving L2 learning environments in contemporary Italy.

Keywords: AV input; informal learning; reasons for exposure; etic and emic perspective; questionnaire-based research.

1. Introduction: Research and recent trends in informal second language practices

In recent years, informal language learning has become a prominent topic in applied linguistics and second language acquisition (SLA) research. The term typically refers to spontaneous, unguided and largely incidental learning (Toffoli, Sockett 2010; Sockett 2014; De Wilde *et al.* 2020; Kusyk 2020; Pavesi, Ghia 2020; Azzolini *et al.* 2022; Krüger 2023). This renewed interest reflects the media saturation and proliferation of today's world and aligns

with a spatial perspective on SLA, which "shine[s] the spotlight on the where of language learning" (Benson 2021, p. 1) – a crucial aspect in a world where globalization and mobility have expanded language affordances and the variety of settings in which languages can be accessed and experienced. Even in a dubbing country like Italy, notable changes appear to be occurring in young people's language behaviours (Cravidi 2016; Casiraghi 2016; Pavesi, Ghia 2020; De Riso 2021, 2023; Special Eurobarometer 2024). This trend calls for further investigation, especially as Italian youth increasingly access a wide range of media in foreign languages – primarily English – predominantly in the form of audiovisual (AV) input from films and television series on large subscription video-on-demand (SVOD) platforms (Pavesi, Ghia 2020). Despite Italy's generally low proficiency in L2 English (ISTAT 2018; EF EPI 2023[1]), younger generations are shifting their attitudes towards English and Englishes (Aiello 2018), with exposure to the language driven by hedonic and social factors rather than purely instrumental ones (Pavesi, Ghia 2020; Ghia, Pavesi 2021).

Recent studies across Europe have explored informal exposure to English in relation to language learning and have focused on the strategies users employ to decode and retain L2 vocabulary and structures (Sockett 2014; Kusyk 2020; Muñoz 2020). While various factors may influence learners' behaviours, most research has concentrated on general informal L2 practices and English as a foreign language (EFL) proficiency, with less attention given to the specific reasons behind informal access to English (Sundqvist, Sylvén 2016; Trinder 2017; Kusyk 2020; Pavesi, Ghia 2020).

Building on a series of surveys conducted at the University of Pavia, the current study explores data from a more recent investigation carried out in 2022 as part of the national PRIN project "The Informalisation of English Language Learning Through the Media", which focuses on informal second language practices (ISLPs) among Italian university students. The study zooms in on AV input and examines viewers' main reasons for exposure from both an etic and an emic perspective, while framing the data within broader patterns of exposure, such as frequency and intensity of contact with English-language AV media, preferred genres, and viewing modalities. The analysis draws on data collected at the University of Pavia, a historic medium-sized institution in northern Italy, through a questionnaire administered to 1,091 undergraduate and postgraduate students (the Informal English Contact and Learning questionnaire, IECoL; Pavesi *et al.* 2023).

In detail, Section 2 introduces key terminology and illustrates typical exposure mechanisms, hinting at how various drivers of informal access come into play; subsection 2.1 provides an overview of prior research on

---

[1] *EF English Proficiency Index*. www.ef.com/epi, https://www. ef.com.tr/epi/regions/europe/

participants' motivations for engaging in ISLPs. Section 3 outlines the study's structure, methodology and findings, presenting them from both etic and emic perspectives based on multiple-choice and open-ended questionnaire items, respectively. Finally, Sections 4 and 5 contextualise the results within the changing L2 learning landscape and the increasing media affordances in contemporary Italy.

2. Reasons for informal contact with English and cognitive processes

Studies on ISLPs focus primarily on Internet-based input, as the web is now one of the primary environments where non-native speakers encounter English and other foreign languages outside traditional instructional settings. This has led to the adoption of terms like online informal learning of English (OILE, Sockett 2014) and more recently informal digital learning of English (IDLE, Lee, Dressman 2018). Online settings offer users opportunities for both receptive and productive language use, with the latter often involving mediated interaction in synchronous and asynchronous modes. The typically multimodal and multisemiotic nature of online input stimulates various forms of contact with English, driven by a wide array of motives. These motives may be intrinsic – originating from users' desires and interests – or extrinsic, arising from such factors as professional or educational requirements, or the influence of instructors or peers. Research on the drivers of informal contact with English has identified several categories of reasons (see Section 2.1; Pavesi, Ghia 2020; Kusyk 2020). Users may engage with English input informally due to hedonic factors, which refer to the pursuit of immediate gratification and sensory pleasure through the activity (Mariotti 2024). Hedonic motives are typically intrinsic, and may range from an appreciation of the language in itself to an affinity for a narrative, original-language dialogue and its authenticity. This is closely related to entertainment, another important reason behind informal access to English, which drives users to engage with English-language media across various contexts because they enjoy the activity and derive long-term fulfillment from it (Vorderer *et al.* 2006). Interpersonal factors may also come into play, stemming from social motivations such as the desire to share experiences and information with others. An orientation to language learning is another trigger to informal exposure, and involves the desire to learn the L2 or to improve one's knowledge of the language at various levels. These motivations can stem from both intrinsic and extrinsic factors. An additional relevant factor is knowledge-seeking, i.e. access to English-language resources to acquire information or learn other content, such as by using informational websites or

wikis on specific topics or disciplinary content. Finally, availability can play a significant role in a media context, where certain affordances such as audiovisual products and video games are exclusively accessible in English.

Given the diversity of motives, English assumes multiple functions in users' personal universes. As Hasebrink *et al.* (2007) put it,

> [d]ifferent opportunities to have contact with English are linked with different sociolinguistic functions. With each additional opportunity to use English its functionality increases: the language develops from a rather uni-dimensional tool – for example, to take part in international pop music or to solve classical tasks in English lessons at school – to a multi-dimensional means of expression and communication which is linked to the professional sphere as well as the private sphere, to globalized mass media entertainment as well as private communication (Hasebrink *et al.* 2007, p. 114).

These exposure scenarios serve as contexts for incidental L2 learning, where potential learning processes may occur as a by-product of a main meaning-oriented activity (Hulstijn 2003), such as searching for information or being entertained, and distance themselves from intentional learning and an explicit focus on the L2. While an orientation to meaning is evident when users typically undertake an activity for enjoyment, hedonic pleasure or knowledge seeking, it may also develop occasionally when access to English input is driven by a desire to learn or improve language skills, especially when the experience is long-lasting and immersive (Pavesi, Ghia 2020). Immersion refers to the deep absorption in a story world and emotional engagement with it and is often facilitated by aesthetic media and multimodal input types, such as videos, video games and virtual reality environments (Green *et al.* 2004; Moreno, Mayer 2007; Wissmath *et al.* 2009; Wilken, Kruger 2016). It is particularly likely to occur when users have sufficient proficiency in the L2 to prevent language barriers, or when some support to comprehension is provided, such as through subtitles in video content. From a neurological perspective, immersion is linked to the activation of mirror neurons, which are stimulated both when an action is performed and when it is observed. Mirror neurons play a role in understanding actions, fostering empathy, and facilitating learning through imitation (Rizzolatti *et al.* 2002). These processes are closely tied to involvement and identification, especially in audiovisual and multimodal settings like video watching and gaming, where attention may shift away from language form and towards the content of the activity.

When engaging informally with L2 input, users' intents may change over time, and a gap often exists between learners' intentions and their actual behaviour (Sheeran, Webb 2016). As Alm (2023, p. 380) notes, informal learners who begin with the intention of improving L2 proficiency by watching TV-series may "get so involved in the plot that they forget to look

up words and binge-watch a few episodes". Typical intentional learning strategies, such as looking up unfamiliar words, may diminish as immersion deepens, as these actions disrupt the flow of the activity (Kusyk 2017). Thus, incidental learning processes can occur regardless of the initial motivation for engaging in ISLPs, as "informal language learning [and exposure] is self-initiated, self-regulated and dynamic, shifting in degrees of intent to learn and in degrees of intensity" (Alm 2023, p. 380).

2.1. Research on reasons for informal second language practices

While empirical studies on ISLPs have flourished in recent years, thorough and systematic research on the factors that drive people's engagement with English outside the classroom remains relatively limited. Most studies tend to focus on the frequency and intensity of exposure to different types of input, proficiency in English, and attitudes towards the language. However, exploring users' motivations for engaging with English can provide valuable insights into how learners interact with L2 input in informal settings and help outline viewers' profiles, by connecting individual exposure patterns with variables such as English proficiency, social and linguistic background, and specific reasons for accessing the language informally.

In their pioneering study on informal contact with English, Berns *et al.* (2007) administered a comprehensive questionnaire to 2,248 school pupils in Belgium, the Netherlands, France and Germany. While the primary focus was on the different contexts in which English was accessed, one section addressed the advantages associated with L2 English proficiency, presenting respondents with a set of predetermined options and allowing for open comments. The preset options covered social benefits (*With English I can make myself better understood abroad; With English I can more comfortably carry on a conversation*), educational and professional advantages (*You need English for further education; With English I have a better chance to get a good job*), the efficacy and appreciation of the language for communicative purposes (*A lot of things sound better in English; For a lot of things there's no equivalent expression*) and the functionality of English as a tool for accessing various media (*With English I can understand English TV programs without subtitles; With English I can understand music texts better*). Overall, students viewed English as a valuable resource across multiple social and personal domains. Yet, while the study touched on issues related to motivation and gathered both etic and emic data, it did not specifically address participants' reasons for engaging with different forms of English-language media; rather, it focused on general attitudes towards English and the broader motivations for using the language, irrespective of privileged settings or input types.

A more targeted exploration of input-specific motivations in ISLPs is found in Trinder (2017), where the author investigates Austrian Economics students' access to English-language media and technologies and collects their views on the impact of these resources on language learning. The study examines various "reasons behind technology choices" (Trinder 2017, p. 403) through open-ended questionnaire items. While most respondents reported accessing English-language media online primarily to improve their L2 skills, entertainment also emerged as a significant factor. Students engaged not only with news websites and journals, but also with TV-series, films and videos, in what the author describes as "dual purpose engagement" (Trinder 2017, p. 407). The respondents highlighted the enjoyment and almost effortless nature of L2 learning while participating in these activities (Trinder 2017, pp. 407-408). Other key factors included personal and peer-group interests; access to social, cultural and pragmatic aspects of the L2; a more accessible input, ensuing from familiarity with both characters and recurrent situations; exposure to idiomatic, colloquial and sociolinguistically varied speech, with visual cues aiding comprehension (Trinder 2017, p. 408).

The impact of social factors and availability is also touched upon in other studies on informal contact with English or on viewers' perception of AV input, including both large-scale research (Arnbjörnsdóttir, Ingvarsdóttir 2018) and case studies focused on specific input environments (Orrego-Carmona 2014). Key interpersonal factors include users' willingness to interact with other speakers in both online and offline settings, such as chats, forums and in-person contact (Arnbjörnsdóttir 2018), as well as recognition from peers within a group or fan community (Orrego-Carmona 2014; Tuominen 2018). Another major factor influencing viewers' habits and decisions to watch English-language AV products in their original version is the earlier availability of subtitled over dubbed TV-series (Orrego-Carmona 2014; Ghia, Pavesi 2021).

Recently, Pavesi and Ghia (2020) explored reasons for informal access to English through various media in their large-scale survey of postgraduate Italian students, marking the first significant study of informal contact with English in Italy. This questionnaire-based study covered a range of media and input types (e.g. films, TV series, YouTube, blogs, forums, video games, web pages, emails, personal contact), with particular attention to AV input. The study also examined the factors influencing viewers' decisions to watch AV content in English. The most common reasons cited by respondents were an orientation to language learning and a hedonic component, which involved an appreciation for original-language telecinematic products and a desire for naturalness and authenticity. These factors also influenced preferences for different subtitling options, where additional variables like accessibility and social alignment (e.g. peer-group preferences) played a role.

In her large-scale study on OILE among French and German university students, Kusyk (2020) used a detailed questionnaire and longitudinal case studies, with the questionnaire also including items on participants' reasons for accessing different media in English. An appreciation of original-language products, entertainment, language improvement or maintenance, and availability were the primary motivations. However, the emphasis shifted towards one factor depending on the type of input and activity involved: while the desire to improve one's L2 skills was only mentioned as a reason for accessing AV content in English, availability and a more diversified choice were what led students to play games and listen to English-language music. The participants engaged instead in online reading activities for academic purposes or to acquire specific knowledge.

Following a similar methodology, De Riso (2023) administered a questionnaire on the type and extent of informal experience with English to Italian university students and complemented it with longitudinal case studies on individual L2 development. Like Pavesi and Ghia (2020), this study paid special attention to AVs and participants' reasons for watching them in English versus their L1. The findings supported previous research, showing a predominant focus on authenticity, followed by an interest in English and an orientation towards L2 learning. Social factors played a relatively minor role in influencing viewers' choice to watch content in English.

Recently, Arndt (2023a, 2023b) explored learner-users' engagement with L2 input in informal settings and developed an instrument to assess it, i.e. the Informal Second Language Engagement (ISLE) questionnaire, which she administered to 506 German adolescents. In educational research, the term engagement refers to learners' involvement, interest and participation in the learning process (Arndt 2023a, p. 2) and is far from a linear and static construct: it encompasses behavioural, cognitive, and affective dimensions. Behavioural engagement refers to the actual activities learners undertake, assessed from an external, etic perspective. Cognitive engagement involves the attentional effort and potential learning strategies learner-users apply during ISLPs. Affective engagement refers to users' emotional states and reactions to informal input sources, such as enjoyment, enthusiasm, frustration, or dislike (Arndt 2023a, p. 3). Individual engagement with different media and input types may change over time and is closely linked to learners' motivation and reasons for participating in ISLPs. In Arndt's studies, the primary motivators for students to engage in informal contact with English were enjoyment and interest in the activity itself, independent of any L2-learning intents. This is reflected in a participant's comment:

> **Jennifer**: I mean, the content is interesting to you and also the people. If you watch a video, then you like the person or what they do. But that it's in English …

> That doesn't matter if it's easy [enough]. It's not like you want to learn English
> and then you watch it, or something like that. (Arndt 2023b, p. 520).

In such contexts, the decision to engage with English input is not linked to learning or to a primary learning orientation, but rather derives from users' general interest in an activity and a desire to be involved in it.[2]

The factors affecting and regulating informal access to English are a valuable area for further research that can provide deeper insights into ISLPs across different participant populations. The current study aims to build on this research by investigating learner-users' reasons for engaging with various media in English, with a particular focus on AV input – one of the most popular *loci* of informal contact with English (see e.g. Muñoz 2020; Dressman, Sadler 2020; Vanderplank 2020; Caruana 2021) and one that, thus far, allows for greater comparability with existing research.

The reasons discussed here are input-specific and transient, bound to vary in relation to not only different media, but also sub-genres and settings. In this, they distance themselves from motivation, which is a more complex construct specifically pertaining to L2 learning (cf. Dörnyei 1998; Mariotti this volume). Given their contextual grounds, reasons can be more aptly described as situated reasons, since they depend on users' interest in a specific activity at a given moment (Arndt 2023b, p. 529), and are not necessarily associated with a learning intent. On the contrary, they often stem from factors like availability, accessibility, social interaction, enjoyment and entertainment.

In what follows, the first part of data analysis examines participants' access to audiovisuals in English to provide a general overview of the sample's behaviour. The second part focuses on the factors driving viewers' choice to watch AVs in English, incorporating both etic and emic perspectives. The study and its findings are detailed in the following sections.

3. The study

3.1. Research questions, sample and questionnaire

This study is part of a broader national project on informal access to English through different input sources and media among undergraduate and postgraduate students at four Italian universities varying in size and geographical location (Pavesi this volume). The present investigation specifically focuses on exposure to English-language audiovisuals at the

---

[2] In the current paper, the term "engagement" is used and interpreted in this more general sense, as access to and involvement with English-language input.

University of Pavia – a medium-sized, historical university in northern Italy – and is based on data collected through a questionnaire. The final sample includes 1,091 undergraduate and postgraduate students from various disciplines. Audiovisual input is the focus of the study, as it has been shown to be the type of input learner-users most frequently engage with, regardless of whether they come from dubbing or subtitling regions and given the spread of SVOD platforms worldwide (Sockett 2014; Arnbjörnsdóttir, Ingvarsdóttir 2018; Ghia, Pavesi 2021). Additionally, AV input allows for greater comparability with similar research conducted in Italy (Pavesi, Ghia 2020; De Riso 2023).

The study addresses the following research questions:

- RQ1: What is the extent and type of informal exposure to AVs in the sample?

- RQ2: Why do participants watch different genres of AVs in English? In other words, what are the reasons for participants' access to English-language AV content?

While RQ1 provides general data on exposure to English-language AVs in the sample and is addressed quantitatively, RQ2 is approached from both quantitative (etic) and qualitative (emic) perspectives. Respondents' answers to multiple-choice questions provide quantitative insights, while an optional open-ended question offers further qualitative depth – with the two dimensions complementing one another.

Data were collected via the IECoL questionnaire, administered to university students from various disciplines at the University of Pavia. Study programmes were grouped into three macroareas: Macroarea A, Arts and Humanities (330 students from Arts, History, Modern Languages, Psychology); Macroarea B, Economics/Law/Political and Social Science (356 students); Macroarea C, Science and Technology (405 students). The questionnaire investigates the frequency and intensity of access to English through several media, along with sociodemographic data, participants' EFL experiences, and attitudes towards English (Pavesi *et al.* 2023). The questionnaire was completed online during lectures in the presence of two researchers, to avoid the self-selection bias ensuing from administration on an exclusively voluntary basis (Dörnyei, Dewaele 2023). On average, it took participants approximately 20 minutes to complete. Given the focus on AVs, the study focuses on sections dealing with exposure to English-language films, TV-series, other programmes, and YouTube videos. Specifically, the items examine overall exposure (frequency and intensity of access to AVs in English; viewing/subtitling modalities; preferred subgenres; viewing supports) and participants' reasons for watching AVs in English – and are detailed in the Appendix.

Lingue e Linguaggi

3.2. Methodology

Both quantitative and qualitative data analyses were conducted. Multiple-choice questions and questionnaire items on the frequency and intensity of exposure to AVs are analyzed quantitatively, and descriptive statistics illustrate exposure trends. In parallel, open-ended questions are explored through qualitative analysis, more specifically thematic analysis (TA; Braun, Clarke 2022).

TA is a qualitative method that involves examining data in stages to identify categories and behavioural patterns. In the present study we apply reflexive TA, which involves prolonged reflection on the dataset through several phases. After reviewing participants' responses to open-ended questions, the author coded and grouped the data into categories that were subsequently merged into broader, recurrent patterns of meaning (Braun, Clarke 2022). These trends were then contextualized by comparing them to the patterns emerging from the quantitative analysis of multiple-choice questions. TA is inductive, semantic and critical (Dawadi 2020). Categories and behavioural patterns are derived from the data themselves (inductive analysis) and reflect the explicit content of participants' responses (semantic TA). In the final phase, the categories and patterns are interpreted critically against the backdrop of viewers' overall responses and the factors driving their behaviours.

3.3. Exposure to audiovisuals in English

3.3.1. RQ1: Overview and contextualisation

This section addresses RQ1 and briefly illustrates general exposure trends, including frequency and intensity of exposure, preferred sub-genres, viewing modalities and supports.

75% of the sample reported accessing AVs in English. Participants indicated that they watch films less frequently than TV-series and other programme types, while TV-series emerged as the most popular telecinematic genre accessed in English (with 45.5% of participants who engage with English-language AVs watching them daily or two to three times a week, Figure 1).
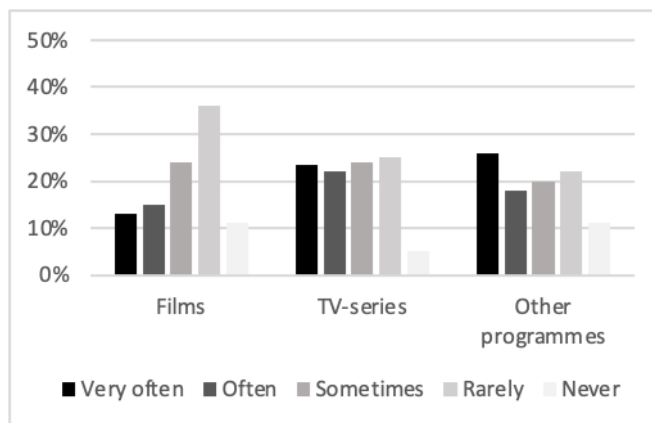
Figure 1
Frequency of exposure to audiovisuals in English.

When participants watch films in English, the activity typically lasts between one and two hours, roughly corresponding to the average duration of a full-length film (Figure 2). Similarly, viewing TV-series often involves longer exposure times, indicating a tendency to binge watch. In contrast, the length of exposure to other programmes varies more widely, as these encompass a range of AV genres.
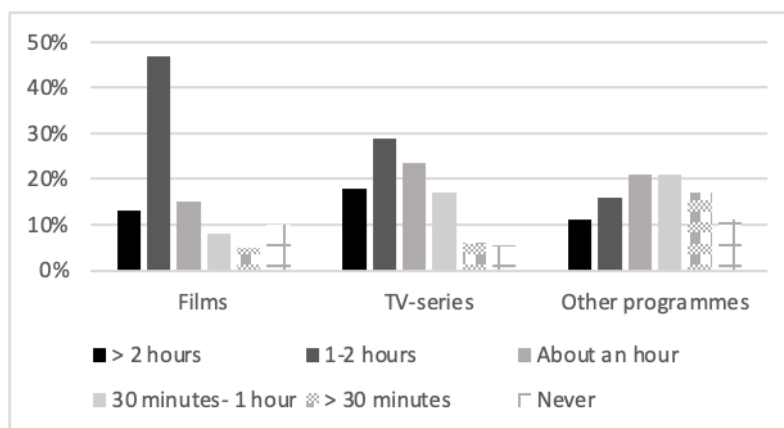
Figure 2
Length of exposure to audiovisuals in English.

Among the additional AV genres included within the category "other programmes" we find products such as documentaries, talk shows and stand-up comedy. Participants expressed a preference for watching talk shows, stand-up comedy, concerts, news programmes and documentaries in English (Figure 3). With a few exceptions (e.g. concerts), the preferred genres primarily consist of highly verbal AV content, offering viewers substantial amounts of oral, fast-paced L2 input.
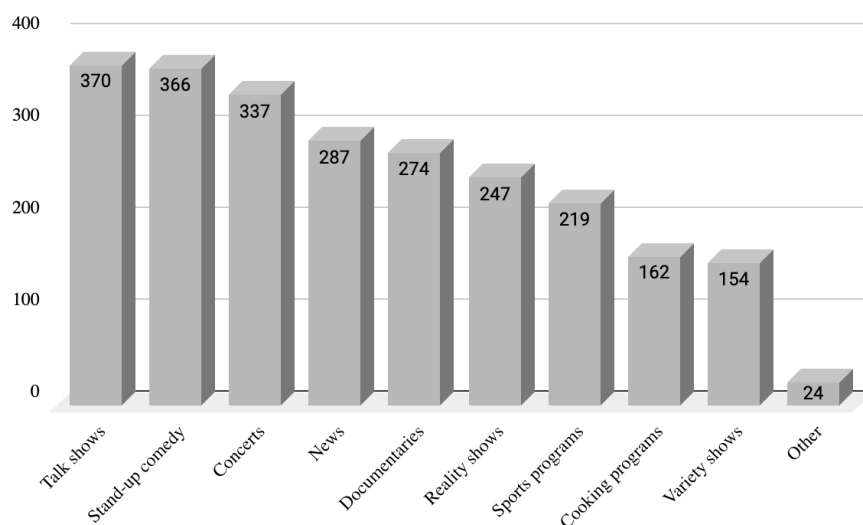
Figure 3
Preferred programmes that participants watch in English (in raw numbers).

Due to the multiplicity of contents it offers, YouTube also triggers different behavioural patterns. The vast majority of students reports accessing YouTube in English (70%), with most doing so on a regular basis (every day or two to three times a week). However, they tend to spend shorter time on the platform, typically between 30 minutes and one hour, or even less than 30 minutes (Figure 4).
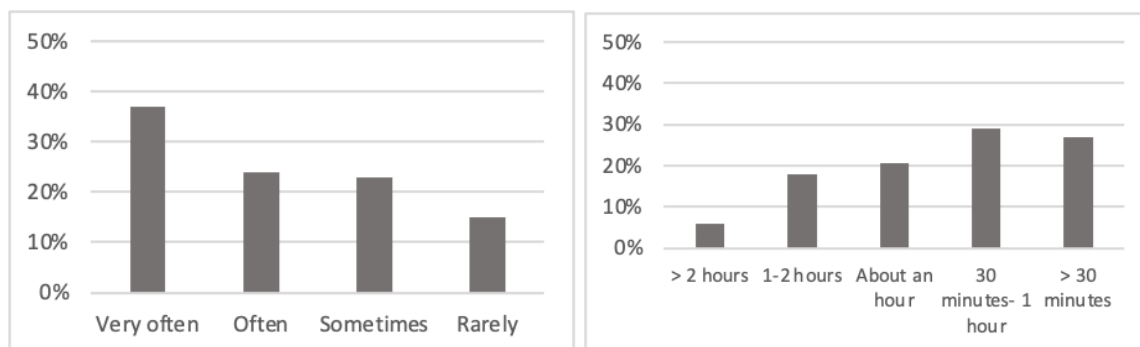
Figure 4
Frequency and length of exposure to YouTube videos in English.

The YouTube video genres that the respondents primarily watch in English include interviews, music videos, tutorials, film trailers, scenes from films and TV-series and YouTubers' videos (Figure 5). These genres typically feature short video content, and lend themselves to quick and flexible exposure sessions.
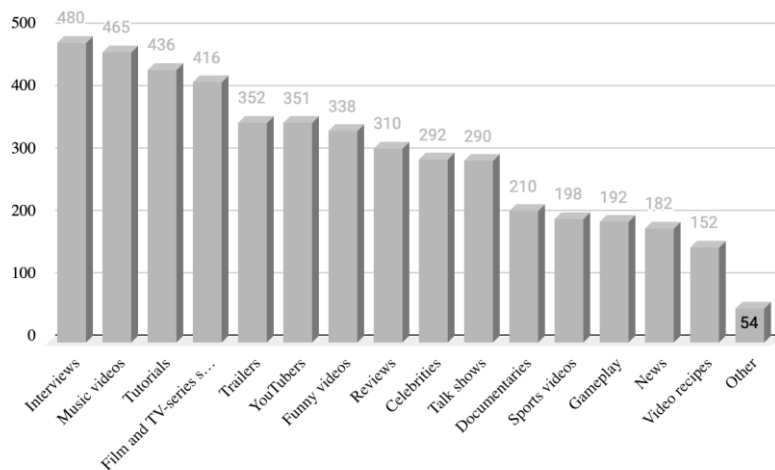
Figure 5
Preferred YouTube genres that participants watch in English (in raw numbers).

When watching films and TV-series in English, most participants report adding English subtitles (respectively 58% and 49%). A lower percentage of students opts for Italian subtitles when watching films and TV-series (28%), while only very few prefer not to use them at all (only 13%). Slightly more respondents watch other programme types without subtitles (23%), a trend that is likely related to their unavailability. Similarly, subtitles are less popular on YouTube, where they are not always an option: among students who access YouTube in English, 45% report using English subtitles, while 42% watch English-language videos without any subtitles, and only 13% add Italian subtitles.

When it comes to supports for watching films and TV-series, about three quarters of the students who overall enjoy telecinematic products in English choose computers (72%), followed by TV and smartphones, while only a minority uses tablets (Figure 6).
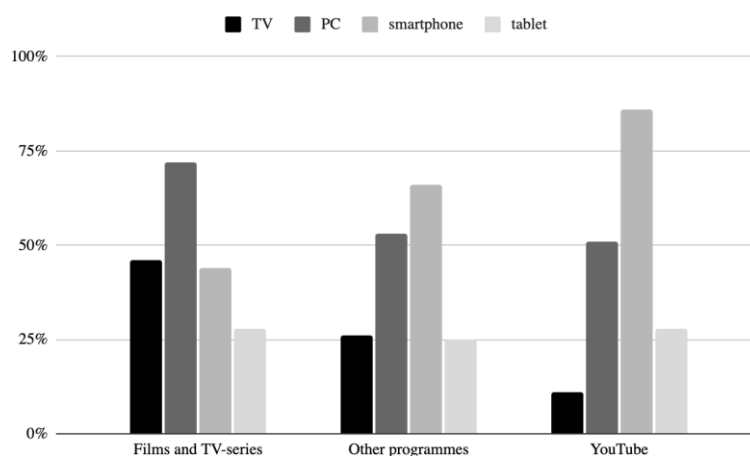
Figure 6
Preferred supports for watching different audiovisuals in English.

In contrast, different trends emerge for other programmes, which are preferably accessed through smaller and portable devices like smartphones (66%), followed by computers (53%), and with only small percentages of participants using tablets and TV. Unsurprisingly, the preference for smartphones is particularly evident when watching YouTube videos (86%). Overall, supports vary by AV genre and consequently video duration, although media convergence allows access to the same contents across multiple digital devices connected to the Internet, including smartphones, laptops, and smart TVs.

3.3.2. RQ2: Reasons for watching English-language audiovisuals

Following the overview of access to AV input in the sample, the analysis will now move onto its main focus: identifying the factors driving informal exposure to audiovisual content in English (RQ2). Initial quantitative data are derived from multiple-choice questionnaire items, which presented the respondents with a set of options (see Appendix) and enabled researchers to collect substantial amounts of behavioural data from an etic perspective. Figure 7 shows the main reasons reported by respondents for watching films, TV-series and other programmes in English, expressed as percentages out of the students who indicated they watch AVs in English.
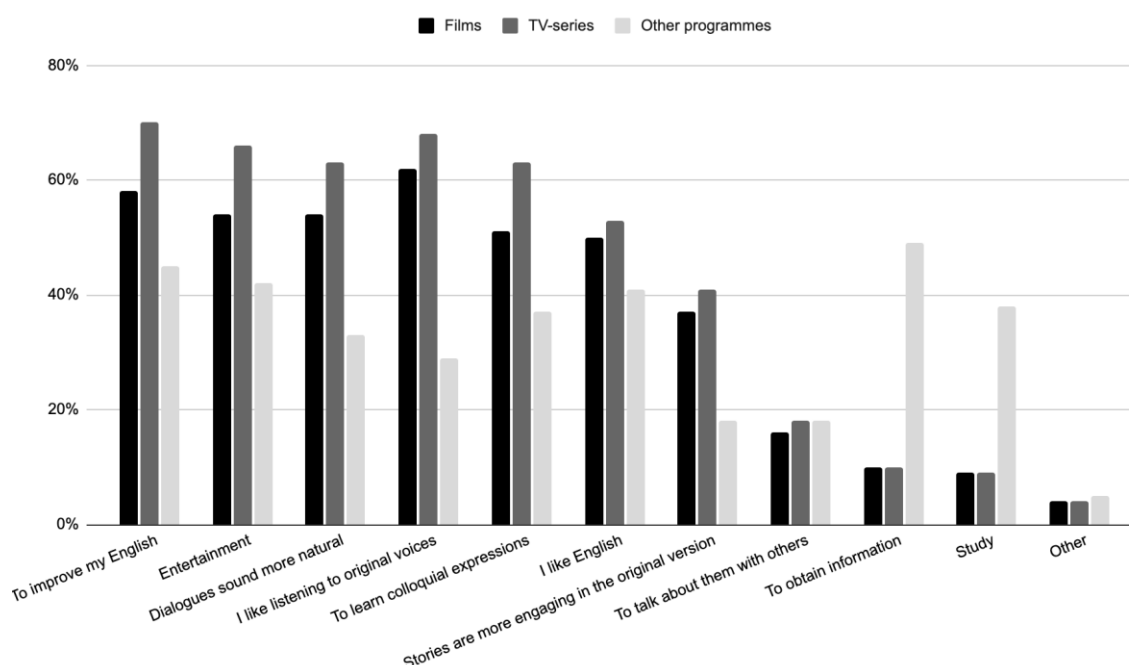
Figure 7
Participants' reasons for watching audiovisuals in English.

The primary motives for watching films in English within the sample include hedonic reasons, such as a search for naturalness in dialogue and actors' voices and an appreciation of the language; a desire to improve L2 skills – encompassing both a general focus on the language and a willingness to learn colloquial expressions; entertainment. Users' expressed intents are more pronounced for TV-series and show a slightly different distribution. An orientation to L2 learning is the primary reason why viewers say they watch series in their original English-language version, as they also serve as tools to develop competence in colloquial registers. The search for naturalness and language appreciation remain highly significant, while entertainment becomes a noticeably stronger factor compared to films. Involvement in the story is cited as an important reason, whereas sociality plays a minor role when users decide to watch TV-series in English.

Conversely, respondents indicate accessing other AV programmes in English for different reasons, primarily knowledge seeking and study. While interest in L2 improvement still affects viewers' decisions, it is to a lesser extent, alongside entertainment, an appreciation of English and other hedonic factors. Due to the lower narrative intent of this genre, involvement is less relevant compared to films and TV-series; similarly, sociality is recorded but remains a minor issue, showing a consistent distribution across all three genres.

When examining the main reasons learner-users access YouTube content in English, the following trends emerge (Figure 8).
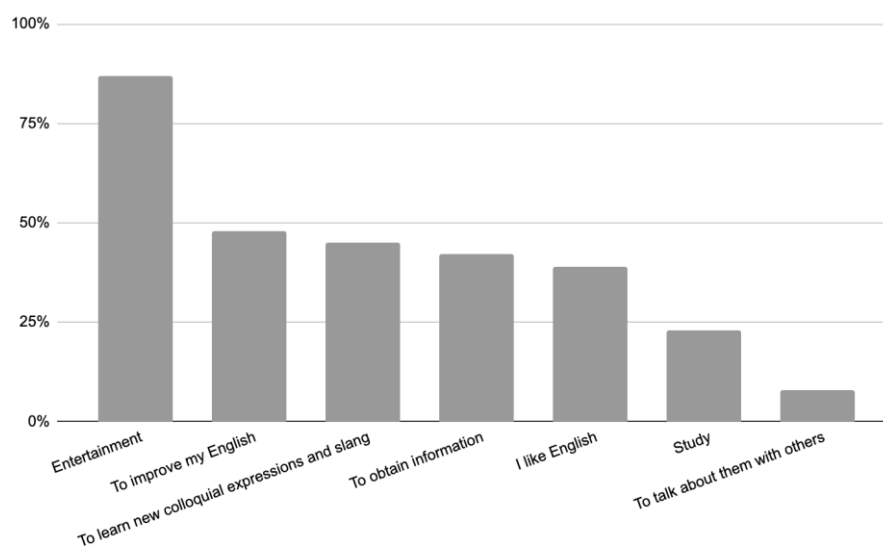
Figure 8
Main reasons for watching YouTube content in English (percentages out of the students who reported watching YouTube content in English).

The participants are drawn to English-language content on YouTube primarily for entertainment (87%). Although respondents' orientation towards L2 development is a relevant factor, it is not the main one; 48% of those who access YouTube in English report doing so to improve their English, and 45% do so specifically to learn colloquial language and slang. Knowledge-seeking also has a considerable impact (42%), as does language appreciation (39%). Only 23% of the respondents are driven by study.

3.3.2.1 Reasons for exposure across different disciplinary areas

In a subsequent phase of analysis, participants were grouped by disciplinary areas. Since Humanities students also include language specialists, a fourth, smaller group consisting of Modern Language students was isolated from the other respondents in the Humanities field. Methodologically, this choice is based on the assumption that language specialists may have different attitudes towards foreign languages and may exhibit different behaviours from their peers. The final groups included 86 Modern Language students (ML), 244 Humanities students (Hum), 356 Social and Political Science, Law, and Economics students (SPLE), and 405 Science and Technology students (SciTech).

 The vast majority of students from all disciplinary areas access English-language AVs. Among the respondents majoring in Modern Languages, 93% report watching films, TV-series and/or other programmes in English. In comparison, 80% of other Humanities students, 72% of social scientists, and 70% of Science and Technology students report watching AVs in English (Figure 9).
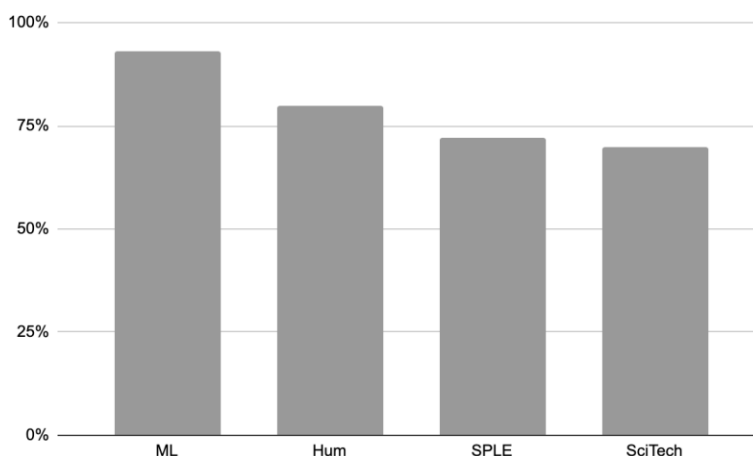
Figure 9
Access to English-language AVs by disciplinary area.

Examining these different disciplinary areas reveals the following patterns (Table 1; percentages are calculated based on the total number of students from each disciplinary area who reported watching AVs in English).

	Films				TV-series				Other programmes			
	Hum	**SPLE**	**SciTech**	**ML**	**Hum**	**SPLE**	**SciTech**	**ML**	**Hum**	**SPLE**	**SciTech**	**ML**
I like listening to original voices	75%	57%	53%	76%	78%	61%	64%	83%	33%	28%	25%	38%
To improve my English	61%	60%	49%	71%	70%	72%	63%	83%	47%	46%	41%	51%
To learn colloquial expressions	52%	50%	46%	73%	62%	60%	60%	84%	34%	39%	34%	50%
Entertainment	58%	48%	52%	71%	68%	60%	66%	81%	44%	37%	45%	46%
Dialogue sounds more natural	61%	52%	49%	58%	69%	59%	60%	66%	34%	32%	32%	43%
I like English	53%	49%	43%	68%	59%	52%	47%	71%	42%	39%	36%	61%
Study	10%	8%	5%	20%	6%	12%	5%	18%	38%	32%	39%	54%
To obtain information	12%	11%	8%	14%	12%	11%	8%	11%	52%	41%	50%	64%
Stories are more engaging in the original version	40%	38%	34%	41%	43%	41%	37%	50%	20%	23%	14%	15%
To talk about them with others	15%	15%	12%	30%	22%	16%	13%	31%	18%	19%	17%	23%
Other	3%	5%	4%	0%	2%	6%	4%	0%	5%	5%	6%	1%

Table 1

Participants' main reasons for watching audiovisuals in English by disciplinary area.

Unsurprisingly, language specialist students engage the most in ISLPs based on AVs. These students demonstrate a stronger orientation towards L2 learning, particularly for films and TV-series (respectively 71% and 83% of students), especially due to a desire to be exposed to colloquial registers (73% and 84%). TV-series are particularly popular, with students often watching them in English also for entertainment (81%), language appreciation (71%) and naturalness (66%). Knowledge seeking (64%) and language appreciation (61%) are the primary motivations for these students when accessing other programmes in English followed by study (54%) and a desire to improve language skills (51%).

Students from other Humanities disciplines access AVs less frequently and intensely than language specialists but more than Social Science and Science and Technology students. Among those who access AVs in English, an appreciation of language and product authenticity prevails when watching films and TV-series in the foreign language, followed by an orientation towards L2 learning. In contrast, knowledge-seeking takes precedence when accessing other programmes, consistent with the overall sample.

Despite their lower engagement with AVs in English, Social Science students show a stronger inclination towards language learning compared to more hedonic and entertainment-related factors when watching all AV genres in the foreign language.

Although they overall access AVs in English to a lesser extent, Science and Technology students align with the main trends observed in the total sample, with entertainment serving as a slightly stronger reason for watching films and TV-series in the L2. Information-seeking remains the primary factor leading these students to access other programmes in English.

87% of language specialists and 73% of other Humanities students declare they access YouTube in English, compared to 67% of social scientists and Science and Technology students (Figure 10). The students' main reasons for watching YouTube content in English are distributed by disciplinary area as illustrated in Table 2.
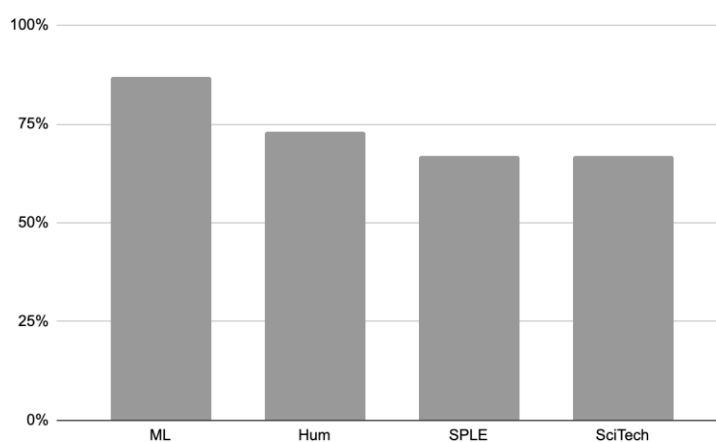
Figure 10
Access to YouTube content in English by disciplinary area.

	Hum.	SPLE	SciTech	ML
Entertainment	89%	82%	88%	96%
To improve my English	50%	45%	45%	64%
To learn new colloquial expressions and slang	47%	43%	40%	67%
To obtain information	44%	36%	46%	35%
I like English	38%	38%	34%	60%
Study	21%	17%	26%	32%
To talk about them with others	7%	9%	7%	11%

Table 2
Participants' main reasons for watching YouTube content in English by macroarea.

Overall, the percentages align with the trends observed in the whole sample. Entertainment ranks first across all disciplinary areas, with Modern Language students still reporting higher engagement than the other groups. An orientation towards L2 learning comes next, serving both as a general motivator and as an opportunity to learn colloquial English. This factor is equally relevant for Science and Technology students as knowledge-seeking.

Language appreciation ranks next; however, with the exception of language specialists, relatively few students in the other areas mention it.

3.3.3. An emic perspective

Further data analysis was conducted on the "Other" option in multiple-choice questionnaire items, where respondents had the opportunity to provide additional feedback beyond the preset answers. This option was included to gain a deeper emic perspective on the data. Although these responses represent only a minority of the total answers (3%, 33 responses), they offer valuable insights into participants' behaviour and enable to observe data from a user-internal dimension.

Through thematic analysis, all responses to the open-ended questions were carefully examined and subsequently coded by similarity, by grouping together answers that addressed similar topics. Based on semantics, different macrocategories were then generated. The reported intents of participants in the free-answer option related to telecinematic texts can be grouped into the following main categories[3]: *availability*, *immersive experience*, *orientation to language learning*, *study*, *habit* and *reference to other AV genres*.

Availability was frequently cited as a relevant factor leading viewers to watch films, TV-series and other programmes in English. Respondents noted that certain AVs are only available in English, leaving viewers with no alternative (example 1). When specified, many of the "English-only" programmes mentioned include sports matches and comedy shows (example 2).

[1] *Alcuni dei programmi o film che guardo non sono mai stati doppiati in italiano* [some of the programmes or films I watch have never been dubbed into Italian]

[2] *Perchè alcuni programmi non sono tradotti, ma sono disponibili solo in lingua originale (ad esempio "Impratical* [sic] *Jokers")* [because some programmes have not been translated, but are only available in the original language (for example, *Impractical Jokers*)]

In other cases, users do have a choice but prefer watching AVs in English because this triggers greater immersion, allowing the product to be fully enjoyed (see 3 and 4, in which L2 competence is also mentioned). This aspect is often linked to a dislike of Italian dubbing (example 4), creating an interplay between entertainment and hedonic gratification.

[3] *Perché guardando film /serie TV in italiano si perde praticamente metà del prodotto che si sta guardando. Sembra di guardare un altro programma* [because when watching films/TV-series in Italian one

---

[3] These factors are not always distinct from the ones provided as preset options, but allowed participants to verbalize their thoughts and views more freely.

L̗ingue e L̗inguaggi

> misses basically half the product they are watching. It's as if you are watching a different programme]
>
> [4] *Generalmente i doppiatori italiani sono bravi, ma ascoltare i film e serie TV o qualsiasi video in lingua originale permette di avere un'immersione maggiore e anche di espandere le proprie conoscenze e abilità linguistiche, importanti più che mai nel mondo odierno* [generally Italian dubbing actors are good, but listening to films and TV-series or any other programme in the original language leads to better immersion and allows to expand one's knowledge and language skills, which is increasingly important in today's world]
>
> [5] *Spesso, le traduzioni e i doppiaggi non mi soddisfano* [oftentimes, I am not fully satisfied with the translation and dubbing]

As mentioned in example 4, an awareness of the language learning potential is further highlighted as a driver to English-language AVs (6), along with study (7) and habit (8):

> [6] *Perché so l'inglese e se non lo faccio lo disimparo* [because I know English and if I don't (watch AVs in English) I unlearn]
>
> [7] *Studio* [study]
>
> [8] *Abitudine* [habit]

In some cases, respondents provided answers related to genres other than films, TV-series and other programmes, most typically YouTube (investigated in the following questionnaire section; example 9), or simply mentioned an AV genre without providing any specific details on their reasons for watching it in English (10):

> [9] *Video in inglese su YouTube per spiegazioni di matematica* [English-language videos on YouTube with math's explanations]
>
> [10] *Talk shows*

Partially different topics emerge regarding YouTube content in English, which include *availability*, *language appreciation*, *variety and appreciation of content* and *sociality*.
The lack of equivalent content in Italian is mentioned in most free responses (11 and 12). A related aspect is the wider range of content offered in English (13):

> [11] *Perché sono interessata a particolari canali fruibili solo in inglese* [because I'm interested in specific channels that are only available in English]
>
> [12] *Perché alcuni contenuti esistono solo in inglese* [because some contents only exist in English]
>
> [13] *Perché la scelta contenutistica è molto più varia di quella italiana* [because the content choice is much more varied than in Italian]

Some viewers reported accessing YouTube in English because of language appreciation and a desire to listen to the language. Example 14 concurrently hints at how practical reasons also affect the respondent's choice:

[14] *Perché la comprendo completamente e mi pare dunque la scelta più immediata* [because I fully understand it [English] and this seems like the most immediate choice]

Lastly, one student mentioned sociality as a trigger for watching YouTube in English, probably in relation to specific video content (15):

[15] *Per rendere contento mio fratello che mi vuole [sic] che guardi questi video* [to make my brother happy as he wants me to watch these videos]

4. Summary and discussion

The quantitative and qualitative analyses of the data gathered in the current study show a considerable consumption of English-language audiovisuals among Italian university students (RQ1). Trends vary by AV genre, with users preferring TV-series and YouTube content and accessing them on a regular basis. According to their reports, respondents often watch TV series and YouTube videos in English daily or two to three times a week and engage in longer (or binge) watching sessions with TV-series. The data also indicate a wide variety of AV content accessed in English through the media, ranging from more "traditional" films and series to talk shows, stand-up comedy, variety shows and interviews. This aligns with recent research showing that, despite its status as a dubbing country on national television and media channels, Italy is experiencing an increase in media affordances in English, with university students and in general younger generations being the key target audience (Ghia, Pavesi 2021). Subtitles are a valuable support when accessing English-language AVs, but the students in the sample predominantly choose bimodal ones, suggesting a desire for a "fully-L2" experience. English subtitles are preferred across all AV genres, even when they may be automatically generated, as is often the case on YouTube. The changing landscape of AV exposure is also evident in the preferred supports for watching, where TV is being replaced by laptops and smartphones, reflecting a tendency to access English-language input on the move. Unsurprisingly, supports vary by genre and video length: the shorter the video, the more likely the viewers are to use smartphones and portable devices.

In response to RQ2, students are drawn to English-language AVs for a variety of reasons, which often intertwine and differ slightly across i) AV genres, ii) participants' areas of study, and iii) students' language-specialist versus non-specialist status. Overall, students report accessing films and TV-series in English as a means to fully appreciate the language and the naturalness of AV products; to improve or maintain their proficiency in English, particularly in colloquial registers; to be entertained; and to feel more involved in the story. The respondents are primarily drawn to L2 TV-

series, where language improvement, the search for greater naturalness and entertainment serve as the strongest motivators, suggesting a multi-purpose intent (cf. Trinder 2017, p. 407). Other AV programmes are accessed in English to a lesser extent and for different reasons, primarily knowledge seeking, although students' awareness of their learning impact still plays a role, along with hedonic factors and the overall enjoyment of the experience.

When examining different user groups, language specialist students engage considerably more with English-language AVs compared to students from other disciplines. Students from other Humanities disciplines also tend to watch AVs in English more frequently and intensely than social scientists and Science and Technology students. Humanities students primarily access AV products to fully appreciate the language and seek an authentic experience, although the orientation to L2 learning is also a relevant factor. Students from the socio-political/economic/legal and science/technology areas watch AVs in English less frequently. When they do, the former show a primary orientation towards learning, while the latter aim for entertainment and enjoyment of the language, although they are aware of the L2 learning potential. For language specialists, the orientation towards L2 learning is stronger, especially when it comes to TV-series. This goes hand-in-hand with entertainment, suggesting that an interest in language improvement coexists with enjoyment, particularly among learner-users who specialize in English.

Entertainment is the primary factor influencing participants when they access YouTube content in English, and this trend is generalised across all groups. While, at a considerable distance, the second most frequent driver is the improvement of L2 skills, it is concurrently L2 development and knowledge seeking for Science and Technology students, again indicating the wider variety of factors that affect informal exposure in this group – and suggesting that language-related aspects have a lower relevance among these students.

The overall results align with trends recorded in recent research in Italy (Pavesi, Ghia 2020; De Riso 2023) and abroad (Trinder 2017). In particular, a search for authenticity and an appreciation of the language emerge as strong motivators in the sample, suggesting that the target audience largely likes English in itself before (or alongside) viewing it as a means to achieve other goals. This is especially true for students who are potentially more "sensitive" to language issues, such as those majoring in modern languages and humanities. More instrumental aims, i.e., accessing AVs in English as a means to improve L2 competence or to reach specific content, characterise students from other disciplinary areas, although in general they seek contact with English-language AVs to a much lower extent.

Even though improving one's skills in the L2 is a pervasive goal, the greater immersion that original-language AVs can provide is explicitly

acknowledged by the participants, who also mention it spontaneously in response to open-ended items. Despite its relevance in other studies (Orrego-Carmona 2014), sociality appears to be a minor factor when users decide to watch AVs in English. However, its variable impact across genres and majors suggests that it does play a role in learner-users' experiences.

Further instrumental factors emerge from the emic insights gained through the questionnaires, primarily availability. Learner-users decide to watch AVs in English, particularly YouTube content and TV-series, because they are mainly interested in accessing products that are only available in that language. When availability is not an issue, hedonic aspects come into play, as some students report choosing English due to a dislike of Italian dubbing.

5. Conclusive remarks

ISLPs are nowadays an important dimension of individuals' L2 learning experiences. The media represent the primary affordance for learner-users, and audiovisuals play a key role due to their wide accessibility, the flourishing of subgenres, and the multiplicity of viewing modes. Access to English-language AVs is also facilitated through different supports, creating a contact experience that is becoming increasingly flexible, dynamic and self-directed (Dressman, Sadler 2020; Benson 2021).

In line with previous research (e.g. Trinder 2017; Pavesi, Ghia 2020; De Riso 2023), the current study shows that when watching AVs in English the target Italian university learner-users do so for a variety of reasons, suggesting a multi-purpose approach to access. While an orientation to L2 learning is relevant, so are language appreciation, entertainment, involvement and knowledge-seeking – which are especially evident when some genres are concerned, e.g. YouTube and other programme types. Reasons for access are therefore dynamic, situated constructs that change across contexts, media types and students' disciplinary interests. Overall, the wide variety of factors mentioned by the participants hints at an extensive orientation to meaning and content during the activities and at a generally immersive experience. Within an environment characterised by high dynamism and growing L2 affordances, the target audience can therefore be described as well aware of the learning potential of such affordances, while at the same time eager to be entertained when accessing the L2 outside the classroom and to enjoy a fully authentic viewing experience. An orientation towards language improvement, hedonic or enjoyment factors and more contingent variables (e.g. availability) often merge and overlap in a mutually-feeding, virtuous circle, and recreate a second language learning environment even in a traditionally foreign language learning context like Italy.

Lingue e Linguaggi

By addressing only a single portion of the total project sample and relying exclusively on questionnaire data, the current study paves the way for larger-scale research and further qualitative, emic investigations. Such research may help shed more light on individual motivators for accessing AV input and combine with a deeper understanding of learner-users' engagement, attitudes and beliefs towards English and Englishes (cf. Aiello 2018) within an ever-changing scenario of L2 use and learning.

Bionote: Elisa Ghia, Ph.D. in Linguistics, is an Associate Professor of English at the University of Pavia, Italy. Her research interests are foreign language acquisition, audiovisual translation, corpus linguistics and the study of spoken English. She has published in the fields of L2 acquisition and audiovisual translation, with a special focus on dubbing and subtitling. With Maria Pavesi, she is co-author of the book *Informal Contact with English. A case study of Italian postgraduate students* (2020).

Author's address: elisa.ghia@unipv.it

References

Aiello J. 2018, *Negotiating Englishes and English-Speaking Identities*, Routledge, London.

Alm A. 2023, *Engaging with L2 Netflix. Two in(tra)formal learning trajectories*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language learning and leisure. Informal language learning in the Digital Age*, De Gruyter, Berlin.

Arnbjörnsdóttir B. 2018, *English exposure, proficiency and use in Iceland*, in Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.), *Language Development across the Life Span. The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham, pp. 35-55.

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development Across the Life Span: The Impact of* English *on Education and Work in Iceland*, Springer International Publishing, Cham.

Arndt H. 2023a, *Construction and Validation of a Questionnaire to Study Engagement in Informal Second Language Learning,* in "Studies in SLA" 45 [5], pp. 1-25.

Arndt H. 2023b, *Behaviour, thoughts, and feelings. Informal second language learning through the lens of task engagement,* in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language learning and leisure. Informal language learning in the Digital Age*, De Gruyter, Berlin, pp. 327-357.

Azzolini D., Campregher S. and Madia J.E. 2022, *Formal instruction vs informal exposure. What matters more for teenagers' acquisition of English as a second language?*, in "Research Papers in Education" 37 [2], pp. 153-181.

Benson P. 2021, *Language Learning Environments*, Multilingual Matters, Bristol.

Berns M., de Bot K., Hasebrink U. 2007, *In the Presence of English: Media and European Youth*, Springer, New York.

Braun V. and Clarke V. 2022, *Thematic analysis: A practical guide*, Sage, London.

Caruana S. 2021, *An Overview of Audiovisual Input as a Means for Foreign Language Acquisition in Different Contexts*, in "Language & Speech" 64 [4], pp. 1018-1036.

Casiraghi L. 2016, *'A Key', 'a Plunge', 'an Anchor': A Study of the Role of English in the Experience of Erasmus Students*, unpublished postgraduate thesis, University of Pavia.

Cravidi G. 2016, *Accessing English Input Outside University: A Qualitative and Quantitative Investigation*, unpublished postgraduate thesis, University of Pavia.

Dawadi S. 2020, *Thematic Analysis Approach: A Step by Step Guide for ELT Research Practitioners*, in "Journal of NELTA" 25 [1-2], pp. 62-71.

De Riso C. 2021, *English Audiovisual Input: A Study of Teenagers' Habits, Attitudes and Responses to Subtitling Versus Dubbing*, in "RILA - Rassegna Italiana di Linguistica Applicata", 53 [1-2], pp. 245-260.

De Riso C. 2023, *Out-of-the-classroom English input and informal Second Language Acquisition: A longitudinal study of Italian university students*, unpublished doctoral thesis, University of Pavia.

De Wilde V., Brysbaert M. and Eyckmans J. 2020, *Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important?*, in "Bilingualism: Language and Cognition" 23 [1], pp. 171-185.

Dörnyei Z. 1998, *Motivation in Second and Foreign Language Learning*, in "Language Teaching" 31, pp. 117-135.

Dörnyei Z., and Dewaele J.-M. 2023, *Questionnaires in Second Language Research: Construction, Administration, and Processing*, Routledge, London.

Dressman M. and Sadler R.W. (Eds.) 2020. *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Ghia E. and Pavesi M. 2021, *Choosing Between Dubbing and Subtitling in a Changing Landscape*, in "Lingue e Linguaggi" 46, pp. 161-77.

Green M.C., Brock T.C. and Kaufman G.F. 2004, *Understanding media enjoyment*, in "Communication Theory" 14 [4], pp. 11-27.

Hasebrink U., Berns M. and de Bot K. 2007, *In the presence of English: A resume after step one of an international study*, in Berns M., de Bot K. and Hasebrink U. (Eds.), *In the presence of English: Media and European youth*, Springer, New York, pp. 111-119.

ISTAT 2018, *Rapporto sulla conoscenza 2018*. https://www.istat.it/storage/rapporti-tematici/conoscenza2018/Rapportoconoscenza2018.pdf (11.12.2023).

Hulstijn J.H. 2003, *Incidental and intentional learning*, in Doughty C. and Long M. (Eds.), *The Handbook of Second Language Acquisition*, Cambridge University Press, Cambridge, pp. 349-381.

Kusyk M. 2017, *Les dynamiques du développement de l'anglais au tarvers d'activités informelles en ligne: une étude exploratoire auprès d'étudiants français et allemands*, Doctoral dissertation, University of Strasbourg/Karlsruhe University of Education.

Kusyk M. 2020, *Informal English Learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 333-348.

Krüger M. 2023, *Media-Related Out-of-School Contact with English in Germany and Switzerland. Frequency, Forms and the Effect on Language Learning*, Springer, New York.

Lee J.S. and Dressman M. 2018, *When IDLE Hands Make an English Workshop: Informal Digital Learning of English and Language Proficiency*, in "TESOL Quarterly" 52 [2], pp. 435-445.

Mariotti C. this volume, *Exploring Italian university students' self-directed exposure to English input in informal settings. Insights into motivation and agency*.

Mariotti C. 2024, *Exploring Motivation and Agency in Informal Contact with English. An emic investigation of Italian university students*, Pavia University Press, Pavia.

Moreno R. and Mayer R. 2007, *Interactive multimodal learning environments*, in "Educational Psychology Review" 19 [3], pp. 309-326.

Muñoz C. 2020, *Boys Like Games and Girls Like Movies. Age and Gender Differences in Out-of-School Contact with English*, in "RESLA" 33, pp. 172-202.

Orrego-Carmona D. 2014, *Subtitling, Video Consumption and Viewers: The Impact of the Young Audience*, in "Translation Spaces" 3, pp. 51-70.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English: An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Pavesi M. this volume, *The informalisation of English language learning: A first national research project*.

Rizzolatti G., Craighero L. and Fadiga L. 2002, *The mirror system in humans*, in Stamenov M.I. and Gallese V. (Eds.), *Mirror Neurons and the Evolution of Brain and Language*, John Benjamins, Amsterdam, pp. 37-59.

Sheeran P. and Webb T.L. 2016, *The intention-behavior gap*, in "Social and Personality Psychology Compass" 10 [9], pp. 503-518.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave Macmillan, London.

Special Eurobarometer 540 2024, *Europeans and their languages*, European Commission. https://europa.eu/eurobarometer/surveys/detail/2979 (11.12.2023).

Sundqvist P. and Sylvén L.K. 2016, *Extramural English in Teaching and Learning: From Theory and Research to Practice*, Palgrave Macmillan, London.

Toffoli D. and Sockett G. 2010, *How non-specialist students of English practice informal learning using Web 2.0 tools*, in "Asp. La revue du GERAS" 58, pp. 125-144.

Trinder R. 2017, *Informal and Deliberate Learning with New Technologies*, in "ELT Journal" 71 [4], pp. 401-412.

Tuominen T. 2018, *Multi-method research: Reception in context*, in Di Giovanni E. and Gambier Y. (Eds.), *Reception studies and Audiovisual Translation*, John Benjamins, Amsterdam, pp. 69-89.

Vanderplank R. 2020, *Video and informal language learning*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 183-201.

Vorderer P., Steen F.F., Chan E. 2006, *Motivation*, in Jennings B. and Vorderer P. (Eds.), *Psychology of Entertainment*, Mahwah, Erlbaum, pp. 2-31.

Wilken N. and Kruger J.L. 2016, *Putting the audience in the picture: Mise-en-shot and psychological immersion in audio described film*, in "Across Languages and Cultures" 17 [2], pp. 251-270.

Wissmath B., Weibel D. and Groner R. 2009, *Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment*, in "Journal of Media Psychology" 21 [3], pp. 114-125.

Annexes

Target questionnaire items:
1. Guardi film, serie TV o altri programmi in lingua inglese? Sì/No
[Do you watch films, TV-series or other programmes in English? Yes/No]

1.1 Quanto spesso guardi in lingua inglese…?
[How often do you watch in English…?]

	Molto spesso (tutti i giorni o quasi) [*very often: every day or almost every day*]	Spesso (due o tre volte la settimana) [*often: two or three times a week*]	Qualche volta (una volta alla settimana) [*sometimes : once a week*]	Raramente (una o due volte al mese) [*rarely: once or twice per month*]	Mai [*never*]
Film [films]					
Serie TV [TV-series]					
Altri programmi (documentari, talk show, news, ecc.) [other programmes (documentaries, talk shows, news, etc.)]					

1.2. Per quanto tempo guardi film, serie TV e/o altri programmi in lingua inglese complessivamente il giorno in cui lo fai?
[For how long do you watch films, TV-series or other programmes in English on the whole on the day you do it?]

	Più di due ore [More than 2 hours]	Da una a due ore [1 to 2 hours]	Circa un'ora [About an hour]	Da 30 minuti a un'ora [30 minutes to one hour]	Meno di 30 minuti [Less than 30 minutes]	Mai [Never]
Film [films]						
Serie TV [TV-series]						
Altri programmi [other programmes]						

1.3. Se guardi film e serie TV in lingua inglese, li preferisci:
[If you watch films and TV-series in English, you prefer them:]

In lingua originale senza sottotitoli [in the original language without subtitles]
Con sottotitoli in inglese [with English subtitles]
Con sottotitoli in italiano [with Italian subtitles]

1.4 Se guardi altri programmi (documentari, talk show, news, ecc.) in lingua inglese, li preferisci:
[If you watch other programmes (documentaries, talk shows, news, etc.) in English, you prefer them:]

In lingua originale senza sottotitoli [in the original language without subtitles]
Con sottotitoli in inglese [with English subtitles]
Con sottotitoli in italiano [with Italian subtitles]

1.5. A parte film e serie TV, quali altri generi audiovisivi/ televisivi ti piace guardare in lingua inglese?
    [Apart from films and TV-series, which other audiovisual/TV genres do you like watching in English?]
|__| News
|__| Sport
|__| Talk shows *[talk shows]*
|__| Reality shows
|__| Concerti *[concerts]*
|__| Varietà *[variety shows]*
|__| Documentari *[documentaries]*
|__| Programmi di cucina *[cooking shows]*
|__| Stand-up comedy
|__| Altro *[other]*

1.6 Quale supporto utilizzi in genere?
[Which support do you usually use?]

	TV	Computer	Smartphone	Tablet
Film [films]				
Serie TV [TV-series]				
Altri programmi [other programmes]				

1.7. Per quali ragioni guardi programmi in lingua inglese originale?
[Why do you watch AV programmes in English?]
È possibile indicare più di una opzione. [You can select more than one option]

	Film [films]	Serie TV [TV-series]	Altri programmi [other programmes]
Per migliorare l'inglese [To learn / improve my English]			
Per svago/intrattenimento [Entertainment]			
Per studio universitario [University study]			

Perché i dialoghi sono più naturali [Dialogue sounds more natural]			
Per accedere a informazioni [To obtain information]			
Perché mi piace ascoltare le voci originali [I like listening to original voices]			
Perché ne posso parlare con altre persone [To talk about them with others]			
Perché mi piace la lingua inglese [I like English]			
Perché nella versione originale le storie sono più avvincenti [Stories are more engaging in the original version]			
Per imparare espressioni colloquiali [To learn colloquial expressions]			
Altro _____ [Other]			

1.7.1 Se hai scelto Altro, puoi specificare:
[If you have selected Other, you can specify:]
_____
_____

2. Guardi YouTube in lingua inglese? Sì/No
[Do you watch YouTube in English? Yes/No]

2.1 Quanto spesso guardi YouTube in lingua inglese?
[How often do you watch YouTube in English?]
Molto spesso (tutti i giorni o quasi) [*very often: every day or almost every day*]
Spesso (due o tre volte la settimana) [*often: two or three times a week*]
Qualche volta (una volta alla settimana) [*sometimes: once a week*]
Raramente (una o due volte al mese) [*rarely: once or twice per month*]

2.2. Per quanto tempo guardi YouTube in lingua inglese complessivamente il giorno in cui lo fai?
[For how long do you watch YouTube in English on the whole on the day you do it?]
Più di due ore [More than 2 hours]
Da una a due ore [1 to 2 hours]
Circa un'ora [About an hour]
Da 30 minuti a un'ora [30 minutes to one hour]
Meno di 30 minuti [Less than 30 minutes]

2.3 Se guardi video su YouTube in lingua inglese, li preferisci:
[If you watch YouTube videos in English, you prefer them:]

In lingua originale senza sottotitoli [in the original language without subtitles]
Con sottotitoli in inglese [with English subtitles]
Con sottotitoli in italiano [with Italian subtitles]

2.4. Quali tipi di video guardi su YouTube in inglese? *[What types of videos do you watch on YouTube in English?]*
|__| Video musicali *[music videos]*
|__| Tutorial *[tutorials]*

|__| Recensioni *[reviews]*
|__| Scene di film e serie TV *[scenes from films and TV-series]*
|__| Sport
|__| Video comici *[funny videos]*
|__| Documentari *[documentaries]*
|__| Ricette di cucina *[recipes]*
|__| Talk show *[talk shows]*
|__| Gameplay
|__| News
|__| Trailer *[trailers]*
|__| Interviste *[interviews]*
|__| Celebrità *[celebrities]*
|__| YouTubers
|__| Altro *[other]*

2.5 Se guardi video su YouTube in inglese, quale supporto utilizzi?
[If you watch YouTube videos in English, which support do you usually use?]
TV / Computer / Smartphone / Tablet

2.6. Per quali ragioni guardi video su YouTube in lingua inglese?
[Why do you watch YouTube videos in English?]
È possibile indicare più di una opzione. [You can select more than one option]

Per svago/intrattenimento [Entertainment]
Per studio [study]
Per migliorare l'inglese [To improve my English]
Per parlarne con altre persone [To talk about them with others]
Per accedere a informazioni [To obtain information]
Per imparare nuove parole ed espressioni di uso corrente e slang [To learn new words and everyday/slang
          expressions]
Perché mi piace la lingua inglese [I like English]
Altro _____ [Other]

WORLD-WIDE ENGLISH
The Internet as a language learning tool

SERENA GHISELLI
UNIVERSITÀ DI BOLOGNA

Abstract –Access to the Internet, both to study and for leisure activities, is ubiquitous today. The Web contains an extensive range of contents, of which about 55% is in English (W3Techs, 2023). In Iceland, a collective seven-year research project was carried out on English language exposure (Arnbjörnsdóttir, Ingvarsdóttir 2018). One of the findings was that students learn more English vocabulary through access to the media in their free time than through focused learning. Studies on the increased informal contact with English were also carried out in other European and extra-European contexts (Berns *et al.* 2007; Kusyk *et al.* 2020; Muñoz, Elke 2020) and Italy is seeing a similar trend (Pavesi, Ghia 2020), but research about naturalistic English learning in the country is still scarce. The ongoing PRIN project "The informalisation of English language learning through media: Language input, learning outcomes and sociolinguistic attitudes from an Italian perspective" (Prot. 2020NNJTW3) aims at filling this research gap. The goal of this presentation is to give a descriptive overview of a set of data collected in the framework of the PRIN project at the University of Salento, Lecce. A survey about English media access and usage was administered to 995 University students. This contribution focuses on the access to the Internet, in particular to YouTube, social media, blogs and forums, web pages, podcasts, radios, apps and e-commerce websites. Data about the frequency and the length of online activities carried out in English will be discussed in relation to students' self-assessed level of English, their attitude towards the language and their socio-cultural background.

Keywords: EFL, the Internet, YouTube, informal learning, incidental learning.

1. Introduction

English is widely used worldwide, in both the physical and digital realms. Today, accessing the Internet for educational and recreational purposes is a widespread habit and it has been estimated that about 55% of online content is in English (W3Techs 2023). Therefore, there are many occasions of incidental English learning on the Web.

Berns *et al.* (2007) highlighted the dominance of English in the European context, where it is the main language used in the media, in schools, in science and technology and in the workplace (where English knowledge is increasingly required of employees). However, despite the extension of informal learning networks, formal learning is not really harnessing the

potential of informal learning opportunities to enhance formal teaching (Czerkawski 2016). It is, therefore, necessary to study informal access to English by students, in order to understand the habits, beliefs and socio-cultural characteristics that favour or hinder informal English acquisition, and take such context into consideration when designing formal English courses.

This paper begins with a brief overview of informal English access, followed by a description of the results of a survey administered at the University of Salento (Lecce, Italy) about informal contact with English. The survey was carried out within the framework of a national project financed by the Italian Ministry of Education, titled 'The informalisation of English language learning through the media'. Subsequently, exposure indexes used to analyze the exposure to YouTube and the Internet in English will be discussed. The final sections examine the interaction between students' socio-cultural backgrounds and attitudes, and media exposure levels, leading to observations on emerging trends. The paper ends with some remarks on the findings and how they contribute to research on informal English acquisition.

2. Theoretical framework

According to Dressman (2020, p.4), informal and formal language learning are defined as follows:

> [...] informal language learning refers to any activities taken consciously or unconsciously by a learner outside of formal instruction that lead to an increase in the learner's ability to communicate in a second (or other, non-native) language. By "formal instruction," I mean learning activities organized by a teacher that are systematic and regularly scheduled.

Informal English learning has been the object of various studies in Europe. Toffoli and Sockett (2010) conducted research to investigate online informal learning practices among students of the University of Strasbourg majoring in non-language subjects. The very same survey was administered in 2012 at a French university, providing a more detailed view of how these students engaged in informal English learning online. In particular, the survey was administered to students who had a limited formal English exposure consisting of a two-hour weekly language course over a 24-week academic year (Sockett 2014). Even if the Internet originated as a source of information based on textual data, Sockett's survey results from 2009 and 2012 show a dominance of listening over reading.

A study by Tan (2013) focused on how students' search and assessment methods within informal learning contexts, particularly on YouTube, influenced their transition to a mixed formal learning environment. Through semi-structured interviews conducted in focus groups, several noteworthy findings emerged. Initially, students continue to depend on educators to pinpoint trustworthy and academically robust information. In situations where information is diverse and abundant, students typically lack the necessary skills to effectively navigate and discern credible content from unreliable sources. Additionally, students highlighted the value of informal networks in their learning journeys, despite their preference and esteem for formal educational settings. Tan emphasises the necessity of aiding students in developing digital literacy skills while maintaining autonomy, enjoyment, and exploration in informal learning environments.

Kusyk (2017) administered a questionnaire about online informal learning of English (OILE) to 953 university students in France and Germany. This was followed by a qualitative study on the development in complexity, accuracy, and fluency of three respondents over five months. In both countries, participants engaged in reception activities more frequently than in production activities. Kusyk's findings show major differences among individuals and within the same person, emphasising the necessity to view second language development as an intricate rather than straightforward process, particularly when examining it in an informal online setting.

In Iceland, a seven-year research project was carried out on English language exposure (Arnbjörnsdóttir, Ingvarsdóttir 2018). In every chapter of the volume the findings of different studies are described; together, these research projects contribute to the joint effort of mapping the role of English in the Icelandic society. One of the findings was that students acquire more English vocabulary through access to the media in their free time than through focused learning.

A recently published volume about informal contact with English in Germany and Switzerland (Krüger 2023) revealed that young learners have extensive exposure to English-language media. Greater exposure correlated with enhanced language skills and variations in media preferences, and media channel choices (e.g. television, books, the Internet) were influenced by socio-economic background. In particular, students of higher socio-economic status have greater exposure to English outside of formal education, primarily through diverse media sources such as books, newspapers, movies, TV series, and online content. On the other hand, students from lower socio-economic backgrounds spend more time on English-language websites but have less overall exposure to English at home. In addition, the majority of students consume media channels in a passive manner, primarily by reading, listening

to, and watching English media content. Conversely, only a small percentage of students actively participate by creating and sharing content themselves.

According to the results of a first study with 305 university students by Pavesi and Ghia (2020), Italy is seeing a similar trend compared to the rest of Europe, but research about naturalistic English learning in the country is still scarce. The project 'The informalisation of English language learning through the media' aims to fill this research gap. The project is based on two instruments constructed and validated specifically to collect data about informal access to English: a questionnaire, the Informal English Contact and Learning questionnaire (IECoL) and a receptive vocabulary test adapted from Nation (1990). Reliability and internal consistency of the instruments are confirmed by Cronbach's Alpha coefficient (Pavesi *et al.* 2023). This is a methodological innovation and a key contribution for a research field where data are mainly represented by the information provided by respondents.

3. Research questions and aims

The focus of this paper is YouTube access and exposure to online content in English among students of the University of Salento measured through a survey. The approach taken is descriptive, with a final data analysis section looking at interactions between media exposure and the socio-cultural background of respondents.

The research questions of this study are the following:

1) What are the respondents' habits in the use of YouTube and the Internet in English?
2) Do the self-assessed level of English and respondents' attitude towards the language influence their exposure to these media in English?
3) Does the respondents' socio-cultural background influence their exposure to these media in English?

In this paper, the term 'media' is used exclusively to refer to the Internet and YouTube. The aim of this paper is to gain a better understanding of the use of the Internet and YouTube in English by the students of the University of Salento. Such knowledge may support researchers and university teachers in formulating ideas about how students' informal contact with English could be harnessed to foster formal English education in the university context.

4. Materials and methods

Information about students' access to media in English in their free time was collected through an online survey, which students compiled during university

lectures using either their computers, smartphones or tablets. The survey was anonymous and anonymity was guaranteed by codes which were given to the students as usernames to access the survey. The data were collected at the University of Salento in November and December 2022. Both undergraduate and graduate students of various degree courses participated in the survey. The data collected were divided by Bachelor and Master degree students. According to the area of study, three groups were identified: science and technology (engineering, biology, mathematics, physics); humanities (languages, Italian studies, primary teacher education, philosophy, art); and social, economic, and legal studies (law, economics, political science).

To analyze the collected data, two media exposure indices were calculated, one for YouTube and the other for the Internet, following the example of Pavesi and Ghia (2020). These authors gave the following definition of high-exposure subjects:

> High-exposure subjects correspond to those participants who report high frequency of access to English input, from often to very often, and a length of exposure exceeding 30 minutes each time (Pavesi, Ghia 2020, p. 87)

The responses were divided into three levels (no exposure, low exposure, and high exposure) based on the amount of time spent on an activity multiplied by its frequency (exposure index= frequency * time).

Numeric values were assigned to Likert scale responses regarding the frequency of access and the length of time devoted to a certain media content in English. For frequency, the values were the following: never = 0, rarely = 1 (once or twice a month), sometimes = 2 (once a week), often = 3 (twice or three times a week), very often = 4 (every day or almost every day). Duration was coded using the following values: never = 0, less than 30 minutes = 1, between 30 minutes and one hour = 2, about one hour = 3, between one and two hours = 4, more than two hours = 5. 'No exposure' is the category corresponding to participants not engaging in a specific activity; 'low exposure' for those scoring below 6 for a particular activity; 'high exposure' includes students scoring 6 or above. The threshold of 6 for high exposure indicates that the person accesses that type of English content from often to very often and with a duration of at least 30 minutes every time they are exposed to English contents (Pavesi, Ghia 2020).

This paper focuses on the questions about the access to YouTube and to Internet contents. To test the access of English contents on the web, the survey included the following questions (Pavesi *et al*. 2023):

15. 2. *Guardi YouTube in lingua inglese?* * [Do you watch videos in English on YouTube?]
    [only one option possible]
    Sì [Yes] [respondent redirected to question 16]
    No [No] [respondent redirected to question 22]

16. 2.1. *Quanto spesso guardi YouTube in lingua inglese?* [How often do you watch videos on YouTube in English?]
[only one option possible]
Molto spesso (tutti i giorni o quasi) [Very often: Every day or almost every day]; *Spesso (due o tre volte la settimana)* [Often: Two or three times a week]; *Qualche volta (una volta alla settimana)* [Sometimes: Once a week]; *Raramente (una o due volte al mese)* [Rarely: Once or twice per month]; *Mai* [Never]

17. 2.2. *Per quanto tempo guardi YouTube in lingua inglese complessivamente il giorno in cui lo fai?* [When on YouTube how long do you watch videos in English overall in a day?]
[only one option possible]
Piu di due ore [More than two hours]; *Da una a due ore* [Between one and two hours]; *Circa un'ora* [About an hour]; *Da 30 minuti a un'ora* [Between 30 minutes and one hour]; *Meno di 30 minuti* [Less than 30 minutes]; *Mai* [Never]

18. 2.3. *Se guardi video su YouTube in lingua inglese, li preferisci:* [If you watch videos on YouTube in English, do you prefer them]
[only one option possible]
In lingua originale senza sottotitoli [In the original language without subtitles]; *Con sottotitoli in inglese* [With English subtitles]; *Con sottotitoli in italiano* [With Italian subtitles]

19. 2.4. *Quali tipi di video guardi su YouTube in inglese?* [What types of videos do you watch on YouTube in English?]
È possibile indicare più di una opzione. [More options possible.]
Video musicali [Music videos]; *Tutorial* [Tutorials]; *Recensioni* [Reviews]; *Scene di film e serie TV* [Scenes from films and TV-series]; Sport; *Video comici* [Funny videos]; *Documentari* [Documentaries]; *Ricette di cucina* [Recipes]; *Talk show* [Talk shows]; *Gameplay* [Gameplay]; *News* [News]; *Trailer* [Trailers]; *Interviste* [Interviews]; *Celebrità* [Celebrities]; *YouTuber* [YouTubers]; *Altro* [Other]

20. 2.5. *Se guardi video su YouTube in inglese, quale supporto utilizzi?* [If you watch videos on YouTube in English, which device do you use?]
È possibile indicare più di una opzione. [More options possible.]
Televisore [TV] / *Computer* [Computer]/ *Tablet* [Tablet] / *Smartphone* [Smartphone]

[…]

34. 5. *Usi Internet in lingua inglese?* * [Do you surf the Internet to engage in activities in English?]
[only one option possible]
Sì [Yes] [respondent redirected to question 35]
No [No] [respondent redirected to question 44]

35. 5.1. *Quanto spesso svolgi le seguenti attività in lingua inglese?* [How often do you engage in the following activities in English?]
[only one option per row possible]
Leggo post e contenuti sui social network [I read posts and contents on social networks]; *Scrivo contenuti sui social network* [I write content on social networks]; *Leggo blog e forum* [I read blogs and forums]; *Scrivo su blog e forum* [I write in blogs and forums]; *Leggo pagine web* [I read web pages]; *Ascolto podcast* [I listen to podcasts]; *Ascolto programmi radio* [I listen to radio programmes]; *Utilizzo app* [I use apps]; *Faccio acquisti su siti* [I shop online]; *Altro* [Other]

36. 5.2. *Per quanto tempo usi Internet in lingua inglese per le seguenti attività complessivamente il giorno in cui lo fai?* [When you surf the Internet how long do you engage in the following activities in English overall in a day?]
[only one option per row possible]
Leggo post e contenuti sui social network [I read posts and contents on social networks]; *Scrivo contenuti sui social network* [I write content on social networks]; *Leggo blog e forum* [I read blogs and forums]; *Scrivo su blog e forum* [I write in blogs and forums]; *Leggo pagine web* [I read web pages]; *Ascolto podcast* [I listen to podcasts]; *Ascolto programmi radio* [I listen to radio programmes]; *Utilizzo app* [I use apps]; *Faccio acquisti su siti* [I shop online]; *Altro* [Other]

37. 5.3. *Se e quando accedi ai social network, qual è la percentuale approssimativa di contenuti in inglese?* [When you access social networks what is the percentage of English contents that you find approximately?]
[only one option possible]
100% / 75% / 50% / 25% / 0%

38. 5.4. *Se accedi a social network in inglese, quali usi?* [If you access social networks in English, which ones do you use?]
È possibile indicare più di una opzione. [More options possible]
Facebook; Instagram; TikTok; Twitter; Pinterest; Tumblr; *Altro* [Other]

39. 5.5. *Se accedi a pagine web in inglese, a quali accedi?* [If you access web pages in English, which ones do you access?]
È possibile indicare più di una opzione. [More options possible.]
Wikipedia; *Altri wiki* [Other wikis]; *Dizionari di inglese* [English dictionaries]; *Siti di notizie e attualità* [News and current affairs web pages]; *Hobby e cucina* [Hobbies and Cooking]; *Altro* [Other]

40. 5.6. *Se accedi a blog e forum in inglese, a quali accedi?* [If you access blogs and forums in English, which ones do you access?]
È possibile indicare più di una opzione. [More options possible.]
Gaming; *Musica* [Music]; *Viaggi* [Travels]; *Estetica e moda* [Beauty and fashion]; *Tecnologia* [Technology]; *Cucina* [Cooking]; *Libri* [Books]; *Grammatica e uso dell'inglese* [Grammar and Use of English]; Cinema; *Auto/moto* [Cars/motorbikes]; Sport; *Altro* [Other]

41. 5.7. *Su Internet interagisci in inglese:* [On the web you interact in English:]
[only one option per row possible]
Con parlanti nativi di inglese [With native English speakers]; *Con parlanti non nativi di inglese* [With non-native English speakers]

For a full description of the questionnaire and the list of all the other questions, refer to Pavesi *et al.* (2023).

First, a description of the data is provided. For each section, graphs were created with Microsoft Excel 2019 to facilitate data evaluation and comparison, specifically:

- bar charts for survey distribution among the three study areas and to represent online interaction in English;

- a pie chart for the self-assessed level of English;

- 100% stacked bar chart to represent access to a certain type of content (frequency) and exposure duration (length);

- sorted bar charts to represent questions allowing multiple answers (e.g. genres of audiovisual products the participant watches). The bars are sorted from the most common answer to the least common.

A second part of the analysis focuses on the results of correlations between a set of predictors and exposure to YouTube and Internet contents in English. The goal of the analysis was to explore whether exposure to YouTube and the Internet in English (outcome) significantly differs depending on students' self-assessed level of English, their attitude towards the language and their socio-cultural background (predictors). Separate group comparisons were carried out between the outcomes and every predictor.

The study employed Pearson's Chi-squared test (χ^2) to conduct comparisons, a statistical method designed for assessing the presence of significant associations between categorical variables (Weiss 2011). This test was selected due to the categorical nature of all variables under examination, aligned with the paper's objective of investigating potential differences in media exposure across distinct groups.

The analysis assessed statistical significance, which evaluates the probability that the null hypothesis is true, relative to the permissible degree of uncertainty about the true result. The amount of uncertainty a researcher is prepared to tolerate, known as the significance level, is set at $p < 0.05$, meaning there is a 5% probability that the study's outcome may be wrong (Tenny, Abdelgawad 2024). Moreover, the analysis extended to incorporate effect size, a quantifiable indicator of the magnitude of observed phenomena relevant to the research object (Kelley, Preacher 2012). Specifically, Cramer's *V* was computed as this effect size measures the strength of association between two categorical variables (Bobbitt 2023). Cramer's *V* is commonly used for chi-squared tests and is the square root of the chi-squared statistic divided by the total number of observations, adjusted for the table dimensions. Cramer's *V* ranges from 0 (indicating no association) to 1 (suggesting perfect association). Results were interpreted based on the following guidelines: ≤ 0.1 for a small effect, between 0.2 and 0.5 for a moderate effect, and ≥ 0.5 for a large effect. The statistical analysis was performed using *Studio 4.2.1* (RStudio 2022), specifically the packages *stats* and *effectsize*.

5. Results

The total number of surveys conducted was 1097. Participants were required to be native Italian speakers; therefore, eight students who declared a different native language were subsequently excluded. Additionally, among these eight non-native Italian speakers, seven had not submitted the vocabulary test, along with another 68 students, despite instructions emphasising the need to submit

both parts by clicking as required. These 75 surveys were excluded from the analysis because it was not possible to compare the questionnaires on English media use with the results of the receptive vocabulary test aimed at assessing foreign language proficiency. The vocabulary test is not considered in this paper. Another 26 surveys were excluded due to incomplete questionnaire responses. Overall, the data presented in this paper are based on 995 surveys that include both questionnaire and vocabulary test data. The results described in this section focus on questionnaire answers about YouTube and the Internet exposure. These results are organised into four subsections: participants' background, exposure to YouTube, exposure to the Internet and the influence of social cultural background on media exposure.

5.1. Participants' background

65% of respondents are female, 34% male and 1% did not declare their gender. The majority of participants come from the province of Lecce (69%), followed by Brindisi (14%) and Taranto (12%); the remaining 5% come from other provinces. This sample can be considered representative of the student population of the University of Salento, since these proportions are in line with the data about students enrolled in AY 2022/2023, when the questionnaires were administered (Università del Salento 2023). Ten students indicated a second mother tongue: dialect (6), English (2), German (1) and Japanese (1). The majority of them (59%) do not use other languages, apart from English, in their free time. The mean age of their first contact with English at school was M=5.96 (SD= 1.27), which corresponds to primary school entry age in Italy, and most of them (61%) had taken extracurricular English courses at some point during their lives. Few respondents had spent time in a foreign country where they had to use English: only 23% had been abroad, but only 16% of them had spent more than three months abroad.

Overall, over half of the students indicated proficiency in more than one language in addition to their mother tongue. Specifically, 14% know one more language, 33% two, 31% three, 14% four and 8% more than four additional languages. The age of first contact with English during leisure activities ranged from 10 to 15 years (57% of respondents).

At the time of the survey most respondents (66%) had an age falling between 20 and 22 years. Almost half of the respondents (48%) come from a humanities or science high school (*liceo classico* and *liceo scientifico*), 20% from a technical high school, 14% from a language high school (*liceo linguistico*), 4% from a professional high school and the remaining 14% from other types of high schools.

Respondents answered a question about their parents' level of education where they had to select the highest qualification held by their father and the

highest qualification held by their mother, choosing among middle school certificate, high school diploma, university degree, PhD/Master and Other. Combining the answers given for both parents, some trends stand out:

27% of the students have parents (both father and mother) who did not gain a high school diploma; 49% have parents who hold a high school diploma (either father or mother); 24% have parents who gained a university degree (either father or mother). Respondents' distribution by curriculum area (science and technology; humanities; and social, economic, and legal studies) and by study level (Bachelor's vs Master's degree) is represented in Figure 1, as the count of the number of answers collected.

As can be seen from the bar charts, the majority of those surveyed come from curricula in the humanities area, where there were larger classes, followed by social, economic, and legal studies and by science and technology. Overall, there are more questionnaires from students in Bachelor's degrees (BA) compared to Master's degrees (MA). The target number of surveys was 140 per area for Bachelor's degrees and 80 per area for Master's degrees. There are more answers from the curriculum of humanities due to larger class sizes, notably in language and educational programs. Enrollment targets were largely met, except for social, economic, and legal Master's degrees, where 66 surveys were collected.
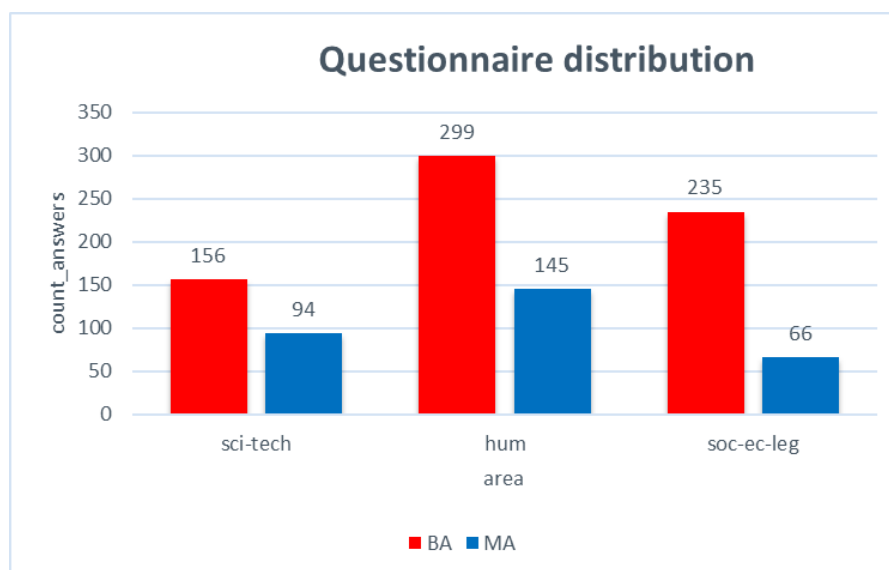
Figure 1
Questionnaire distribution by area and level of studies.

The survey also included questions about respondents' perception of the English language, namely: "*Quanto è importante per te conoscere l'inglese su una scala da 1 a 10?*" [How important is it for you to know English on a scale from 1 to 10?] and "*Quanto ti piace la lingua inglese da 1 a 10*? [How much do you like English on a scale from 1 to 10]. Figures 2 and 3 are stacked bar

charts representing the participants' answers; the students were to indicate values on a Likert scale from 1 to 10, where 1 corresponds to not important or not liked and 10 to very important or very much liked.
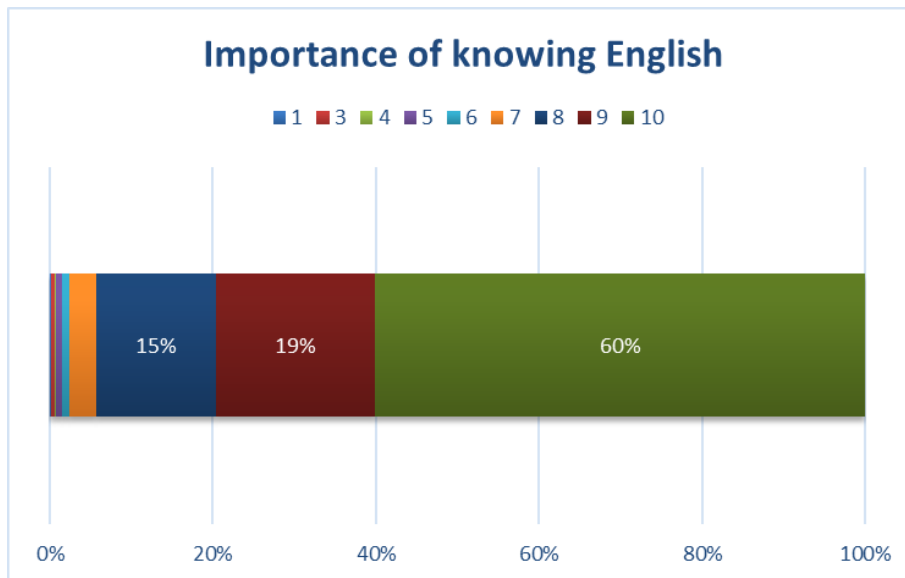
Figure 2
Importance of knowing English
(1= not important at all; 10=very important).
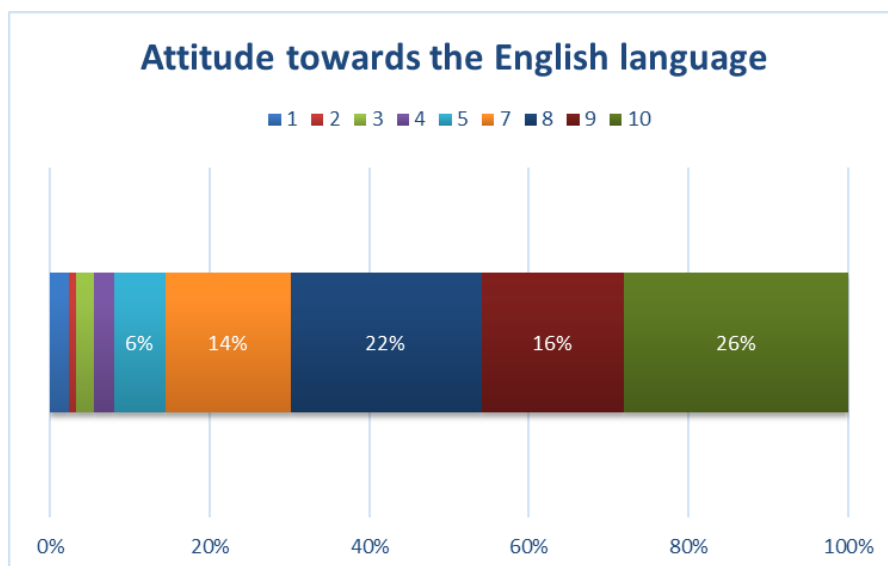
Figure 3
Attitude towards the English language
(1= I do not like it; 10= I like it very much).

From Figure 2 it can be seen that the perception that English is important is widespread (10 out of 10 for 60% of the respondents). The attitude towards the

English language represented in Figure 3 is not as positive as the perception of its importance, and it is more varied. Overall, it remains towards the high-end of the Likert scale, since 64% of answers about the attitude towards English are between 8 and 10, with 10 corresponding to 'I like it [English] very much'. The self-assessed level of English is represented in Figure 4.
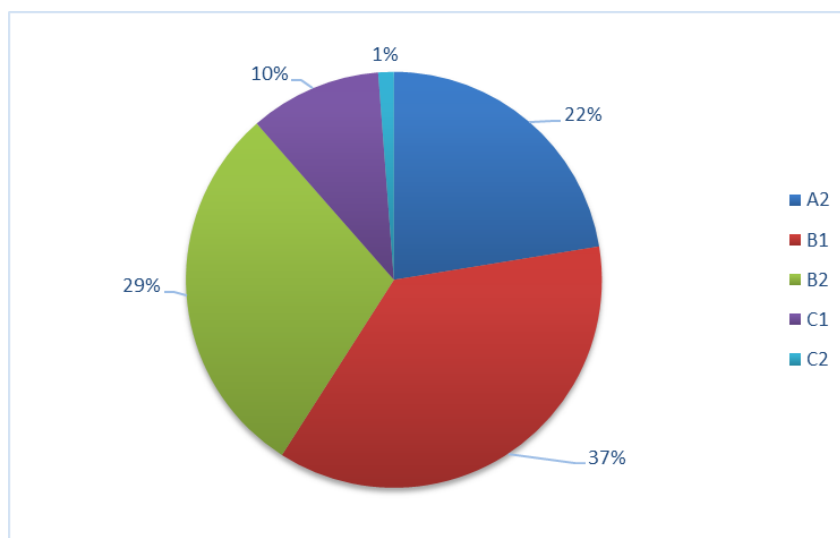
Figure 4
Self-assessed level of English according to the Common European Framework of Reference for Languages (Council of Europe 2020).

The majority of students consider their level of English as an intermediate level, corresponding to B1-B2 (66%), a minority of respondents describe their level as advanced (11%) and 22% of students think they have an elementary level (A2).

5.2. Exposure to YouTube

When students were asked about the use of YouTube in English, 56% of them answered that they use YouTube in English. Figures 5 and 6 illustrate the frequency and length of YouTube use. Figure 5 does not highlight a specific trend in the frequency of use of YouTube, because the answers are evenly distributed among the options 'sometimes' (30%), 'often' (23%) and 'very often' (31%), whereas only a minority of respondents chose 'rarely' (10%). As far as the length of use of YouTube in English is concerned, when on YouTube 57% of respondents watch videos for less than one hour, while 42% for longer time. When asked about the use of subtitles, 46% of students answered that they use English subtitles, 28% that they add Italian subtitles and 26% that they watch English contents without subtitles.
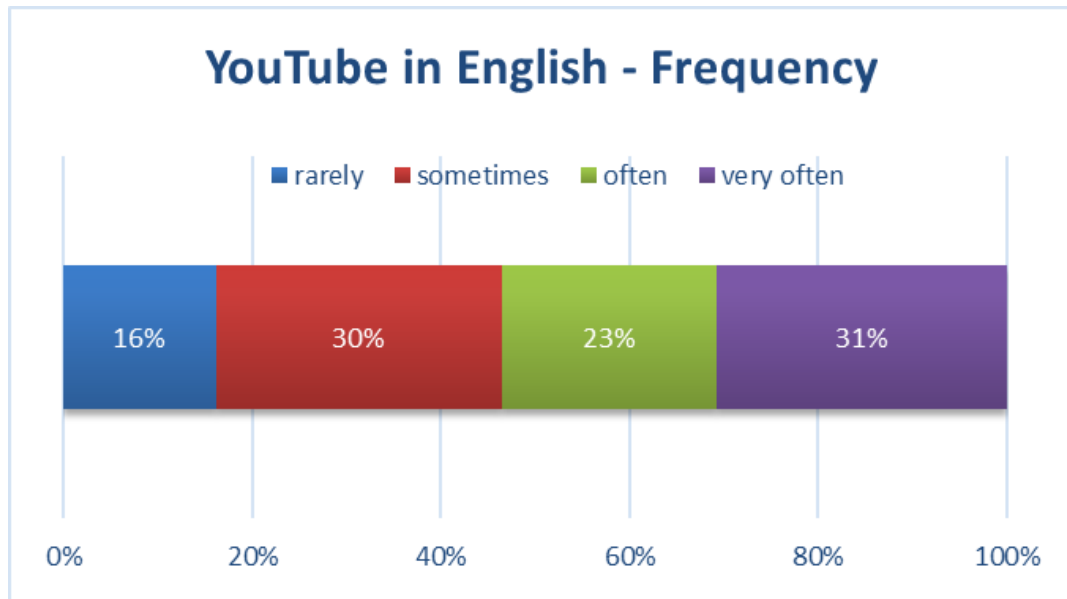
Figure 5
Frequency of use of YouTube in English.
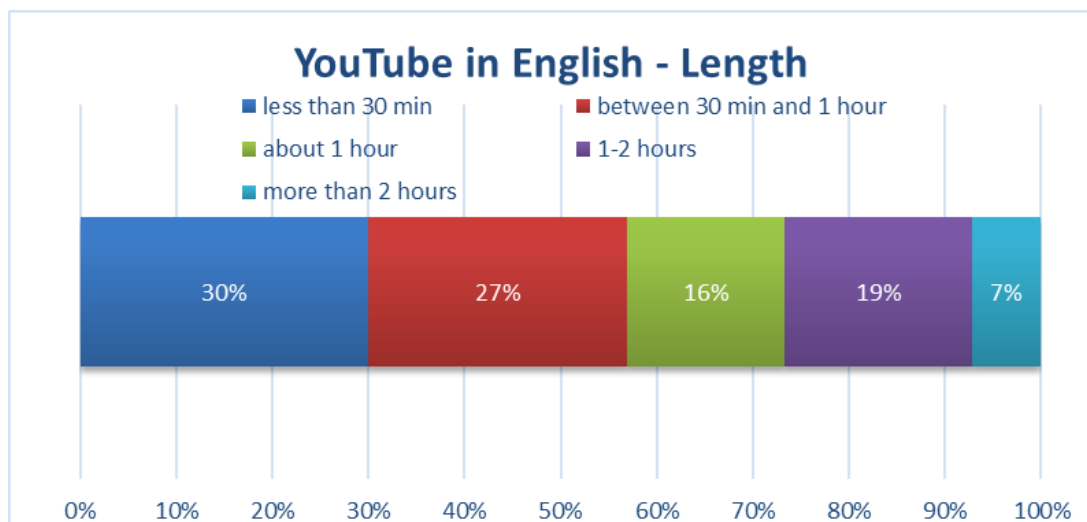
Figure 6
Length of use of YouTube in English.

The sorted bar chart shown in Figure 7 shows the types of contents respondents access on YouTube, represented as the count of answers from the highest to the lowest. Respondents had a pre-defined list of YouTube content types and could choose more than one option. The most popular YouTube contents are music videos (384), tutorials (330) and film/TV scenes (320).
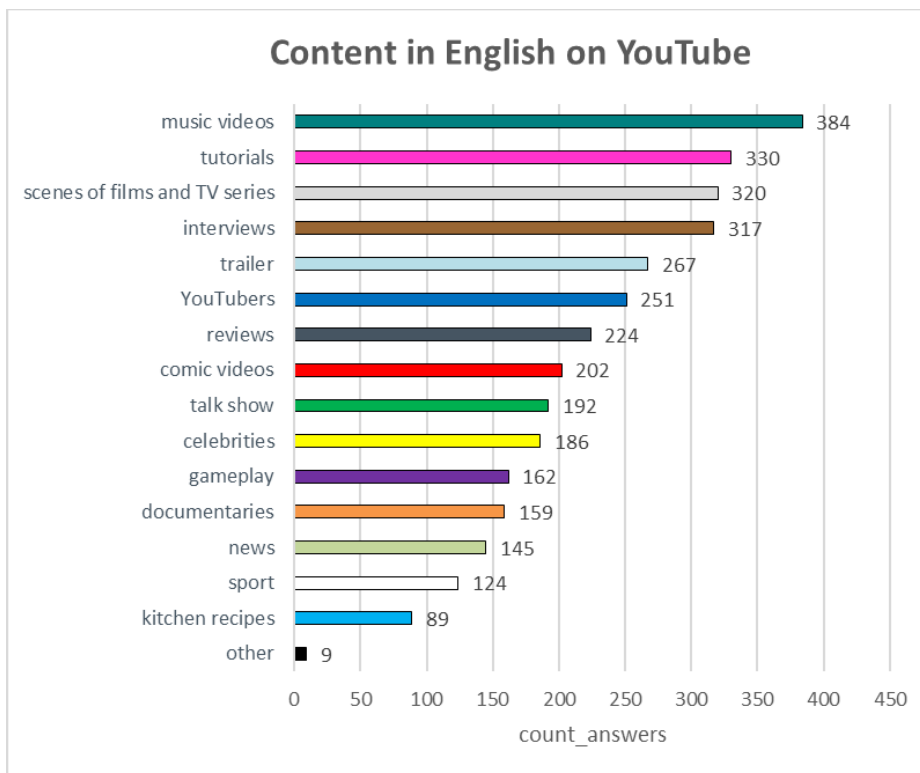
Figure 7
Types of content in English accessed through YouTube videos.

As far as the device with which students access videos is concerned, respondents had to indicate one or more devices among television, computer, tablet and smartphone. Smartphone was the most common answer (464), followed by computer (343), tablet (130) and television (74).

5.3. Exposure to the Internet

Data about Internet English usage differ from those of YouTube. In fact, most students (65%) reported no usage of the Internet in English. The use of the Internet was surveyed with questions about a variety of contents (see Section 4).
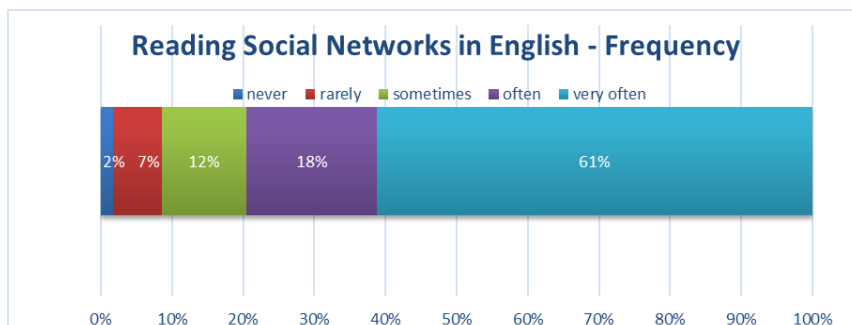
Figure 8
Frequency of reading content in English on social networks.

Reading posts on social networks in English (see Figure 8) is an activity which is part of respondents' daily routine since the majority of answers (61%) were that they do it 'very often' (corresponding to every day or almost every day).

Reading Social Networks in English - Length

■ no time	■ less than 30 min
■ between 30 min and 1 hour	■ about 1 hour
■ 1-2 hours	■ more than 2 hours

2% 26% 13% 19% 20% 20%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
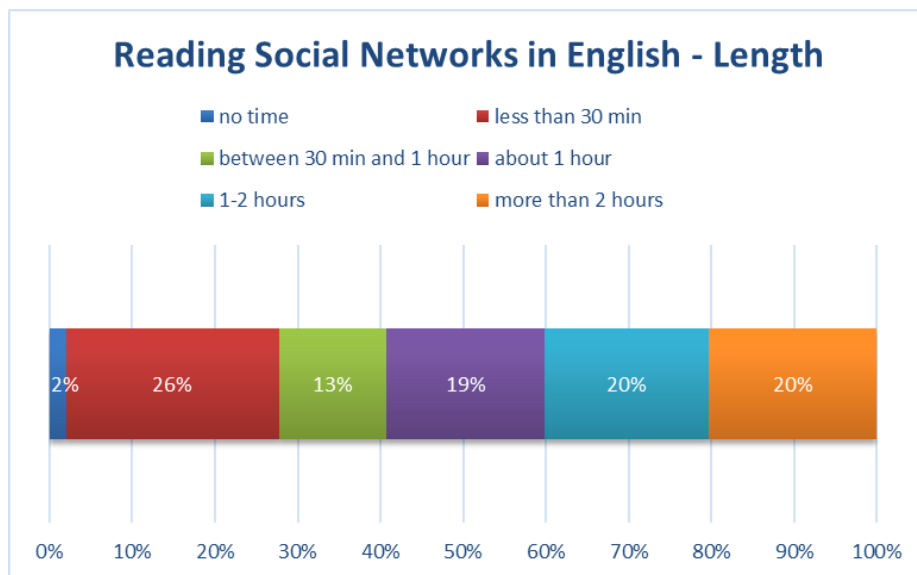
Figure 9
Length of time dedicated to reading content in English on social networks.

The length of time devoted to reading content in English on social networks (see Figure 9) is varied and evenly distributed among the different options.

Writing Social Networks in English Frequency

■ never ■ rarely ■ sometimes ■ often ■ very often

40% 27% 15% 8% 11%

0% 20% 40% 60% 80% 100%
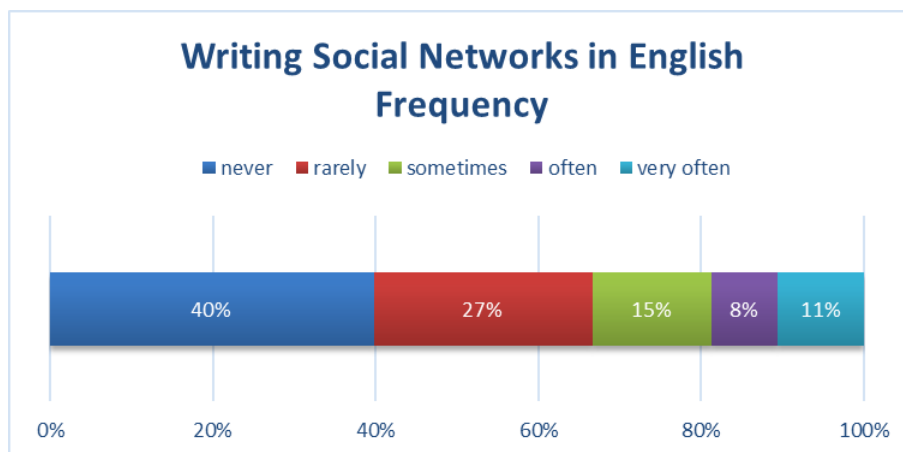
Figure 10
Frequency of writing content in English on social networks.

Writing content in English on social networks (see Figure 10) is not a habit for the majority of respondents, since 40% never do it and 27% do it rarely (which corresponds to once or twice a month).
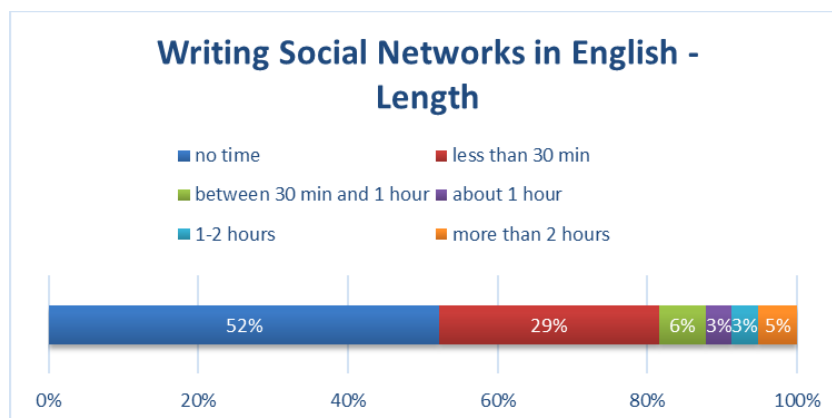
Figure 11
Length of time dedicated to writing content in English on social networks.

In line with the results about the frequency of writing content in English on social media, also the length of time confirms this is not an activity that respondents engage with to any considerable extent. From Figure 11 it can be seen that 52% of those who access the Internet in English devote no time to writing in English on social media.

Respondents had a pre-defined list of social networks and could choose more than one option. Commonly used social media platforms in English include Instagram (307 answers), TikTok (155) and Facebook (120), followed by Twitter (113), Pinterest (104), Tumblr (28) and other social media (20) (see Figure 12). Among other social media, 16 students out of 20 mentioned Reddit.
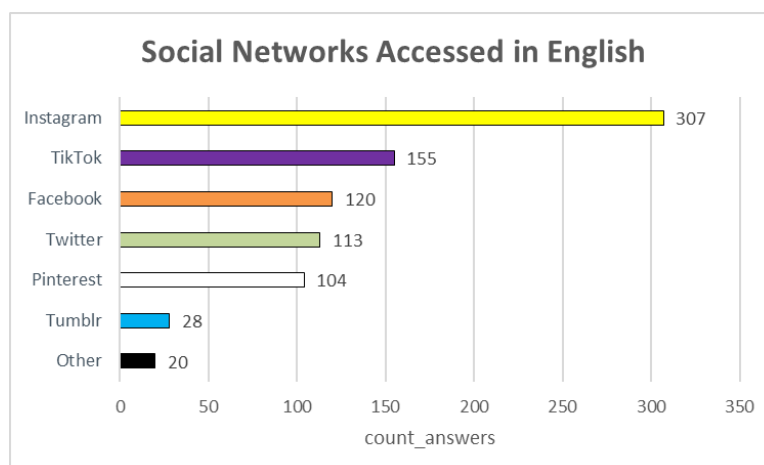
Figure 12
Social networks accessed in English by respondents.

When asked about the proportion of English content they access on social networks, 41% of respondents said 75%, 37% reported half, and 17% indicated a quarter. Furthermore, 2% of the students access no content in English while 4% all the content in English.

Figures 13-17 display the data collected for blogs and forums.

Reading Blogs and Forums in English - Frequency

never ■ rarely ■ sometimes ■ often ■ very often

| 13% | 19% | 22% | 23% | 23% |

0%   20%   40%   60%   80%   100%
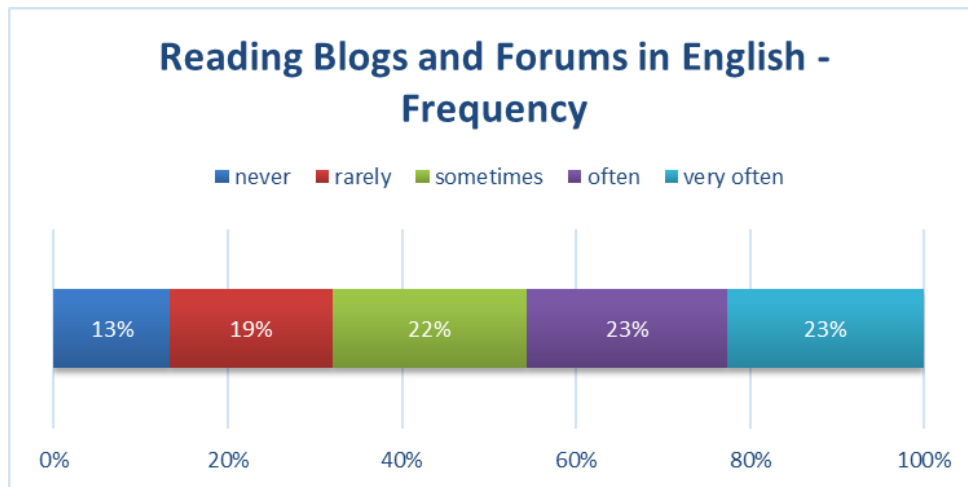
Figure 13
Frequency of reading content in English on blogs and forums.

The frequency of reading content in English on blogs and forums is varied and evenly distributed among the different options.

Reading Blogs and Forums in English - Length

■ no time   ■ less than 30 min
■ between 30 min and 1 hour ■ about 1 hour
■ 1-2 hours   ■ more than 2 hours

| 18% | 33% | 21% | 11% | 8% | 9% |

0%   20%   40%   60%   80%   100%
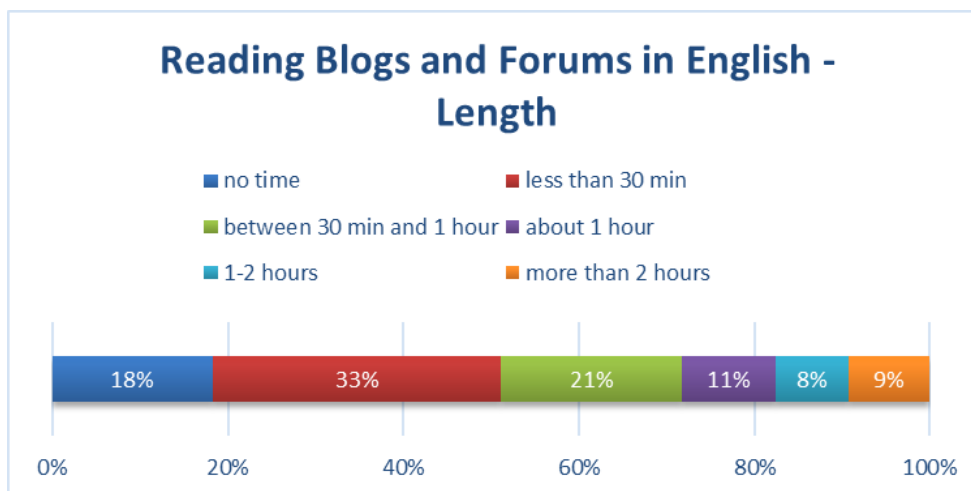
Figure 14
Length of time dedicated to reading blogs and forums in English.

Also for the length of time devoted to reading blogs and forums there is no clear trend, even if for this question there were a few more answers corresponding to a duration shorter than 30 minutes (51%) rather than a longer duration of the activity (49%).
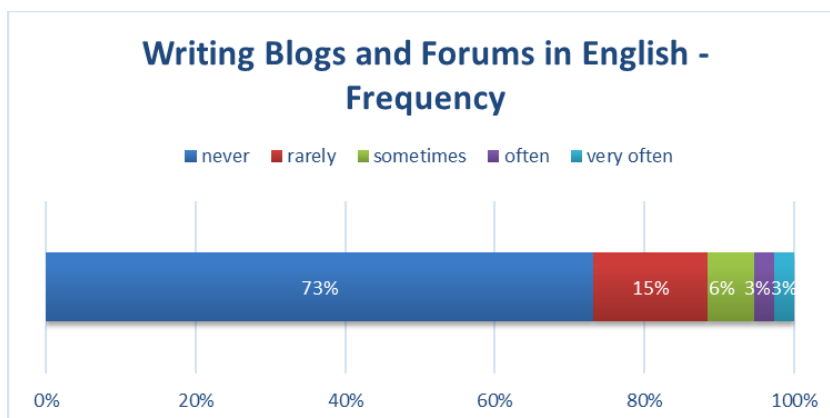
Figure 15
Frequency of writing content in English on blogs and forums.

Writing content in English on blogs and forums is an activity respondents do not generally do, in fact 73% answered that they never write in English on blogs and forums.
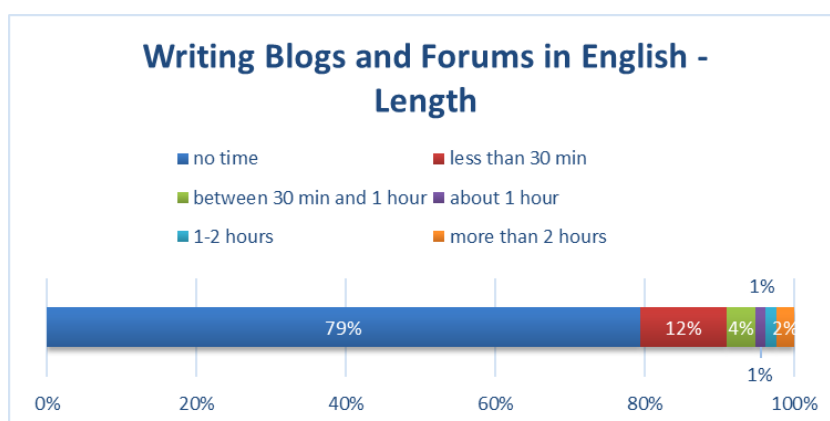
Figure 16
Length of time dedicated to writing content in English on blogs and forums.

In line with the results about the frequency of writing content in English on blogs and forums, also the length of time confirms this is not an activity that respondents generally do at all, as 79% of them devote no time to it. Not writing in English on blogs and forums is in line with what happens for social networks, but for blogs and forums this trend is more marked.

Respondents had a pre-defined list of topics of blogs and forums and could choose more than one option. The main topics respondents read about in blogs and forums in English are music (218 answers), cinema (172) and technology (143), followed by travels (131), books (130), gaming (114), grammar and English usage (112), beauty and fashion (87), sport (75), cooking (51) and cars/motorbikes (31) (see Figure 17).
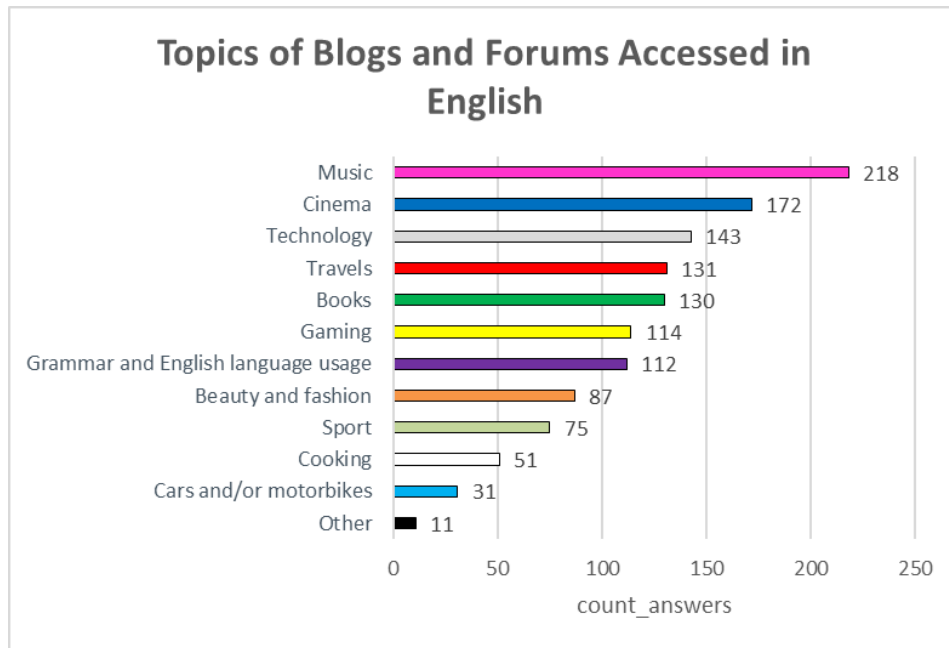
Figure 17
Topics of blogs and forums accessed by respondents.

Figures 18-20 display the data collected for the access to websites in English.
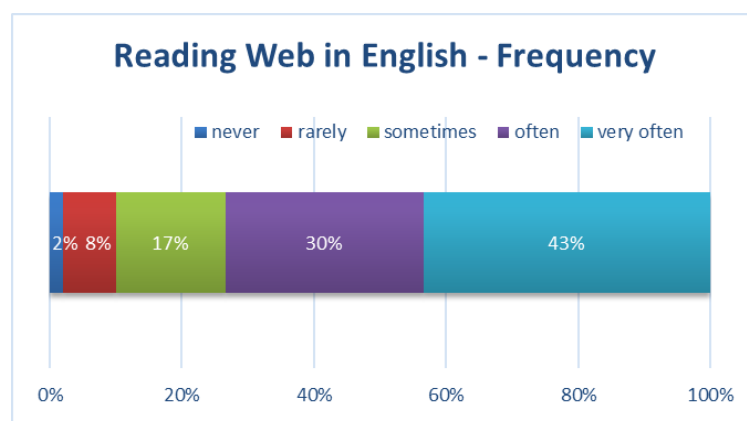
Figure 18
Frequency of reading Web content in English.

Not surprisingly, reading web content in English is something most respondents do either often (30%, corresponding to twice or three times a week) or very often (43%, corresponding to every day or almost every day).

When considering the length of time devoted to reading web content in English, from Figure 19 it can be seen that there were a variety of answers, with overall more respondents (56%) that spend relatively short time (less than one hour) reading online in English compared to those who do it for at least one hour (45%).
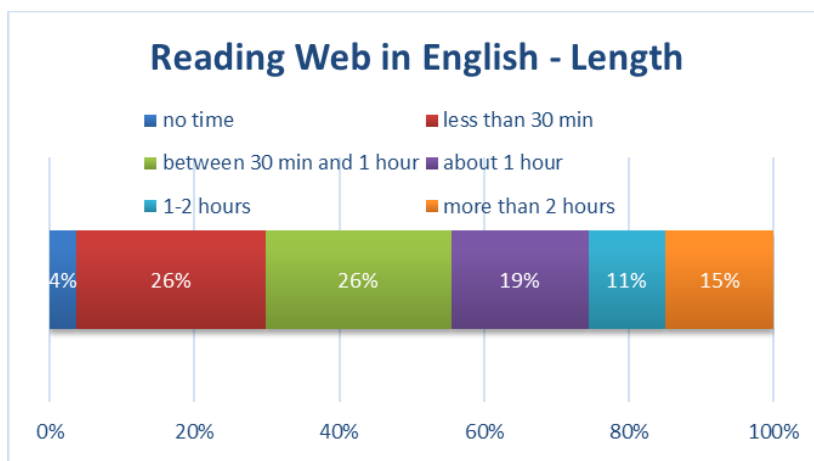
Figure 19
Length of time dedicated to reading Web content in English.

Respondents had a pre-defined list of web pages and could choose more than one option. The main websites respondents access in English are Wikipedia (275 answers), news and current affairs (223), English dictionaries (198), other wikis (118), pages about hobbies and cooking (82), other websites (22) (see Figure 20).
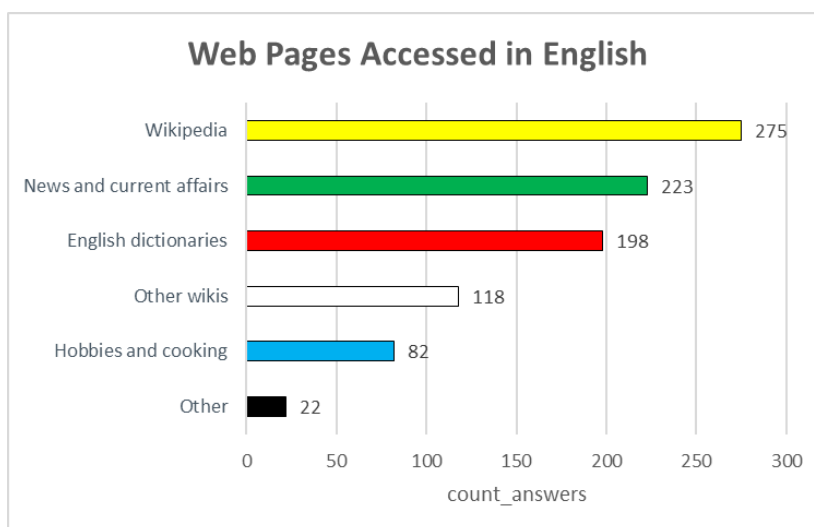
Figure 20
Web pages accessed by respondents.

Figures 21-28 display the data collected for the access to podcasts, radios, apps and e-commerce websites. Listening to podcasts in English is not very common among respondents, only 21% do it often (twice or three times a week) or very often (every day or almost every day). In line with the answers about the frequency of listening to podcasts, also the length of time shows that this is not a common activity. Respondents either do not spend time doing this activity (42%)

or they spend less than one hour (37%). Only a minority of respondents (16%) devote one hour or longer to listening to podcasts in English when they do so.
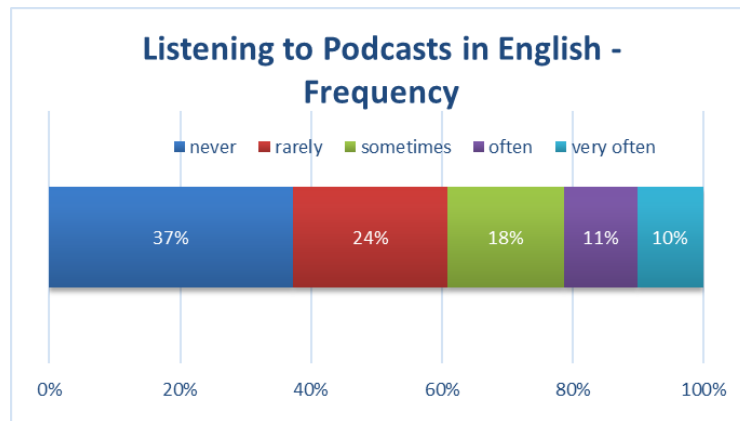
Figure 21
Frequency of listening to podcasts in English.

Figure 22
Length of time dedicated to listening to podcasts in English.
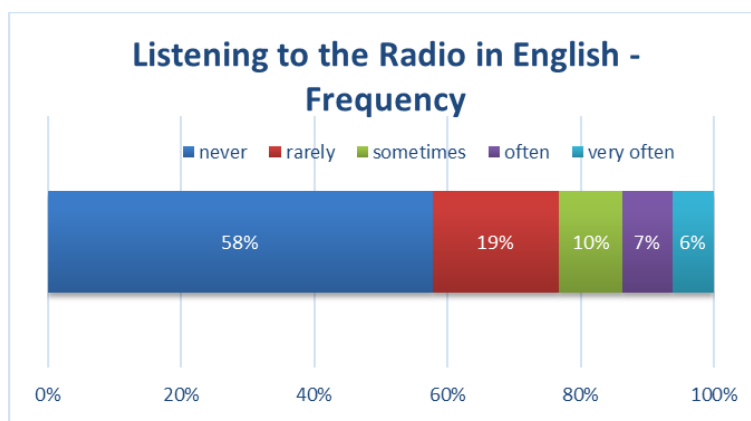
Figure 23
Frequency of listening to the radio in English.

Listening to the radio in English is even less common than listening to podcasts. In fact, 58% of respondents answered that they never do it.
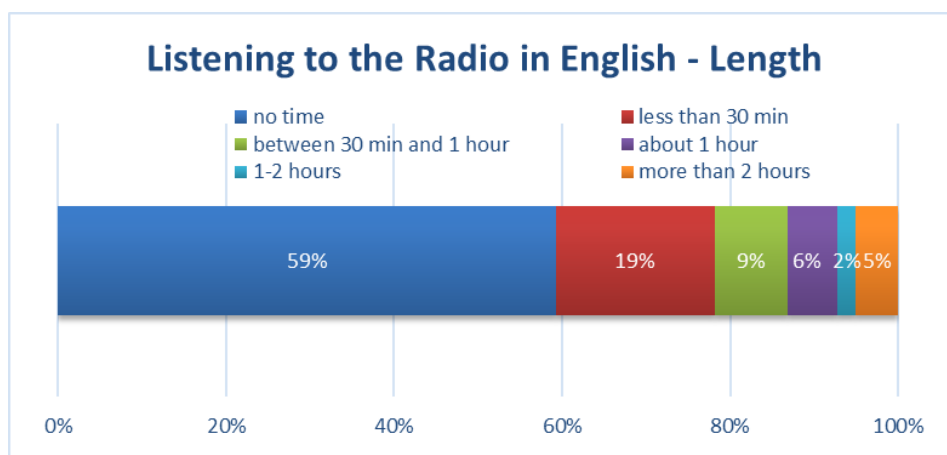
Figure 24
Length of time dedicated to listening to the radio in English.

In line with the results about the frequency of listening to the radio in English, also the length of time confirms this is not an activity that respondents do. The majority of them devote no time (59%) or less than 30 minutes (19%) to it. Overall, respondents do not listen to audio only contents online and those who do it prefer podcasts to the radio.
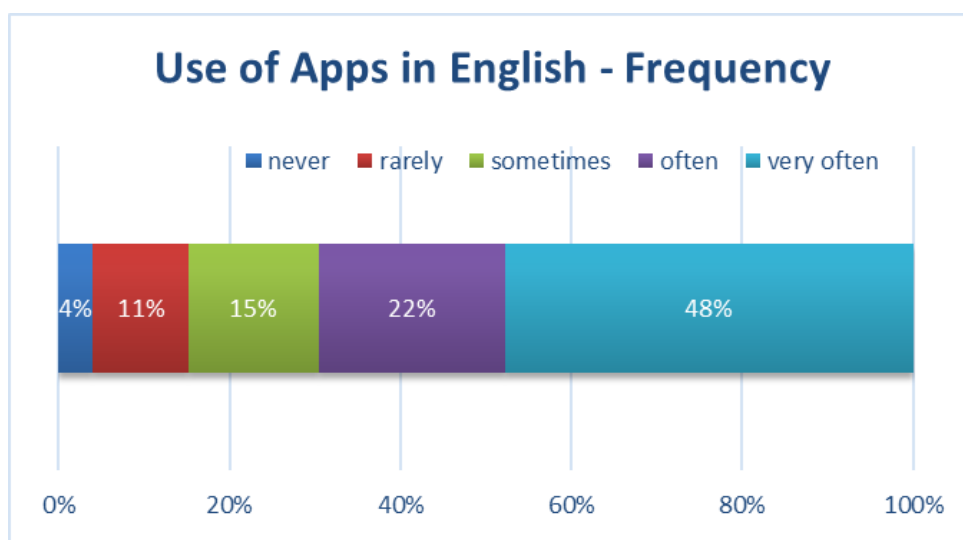
Figure 25
Frequency of use of apps in English.

Most respondents (96%) use apps in English and many (48%) use them very often, that is every day or almost every day. Figure 25 shows that using apps as an activity is clearly a habit, but from Figure 26, concerning the length of

time spent on apps in English, no clear trend stands out about its duration, since answers are scattered.
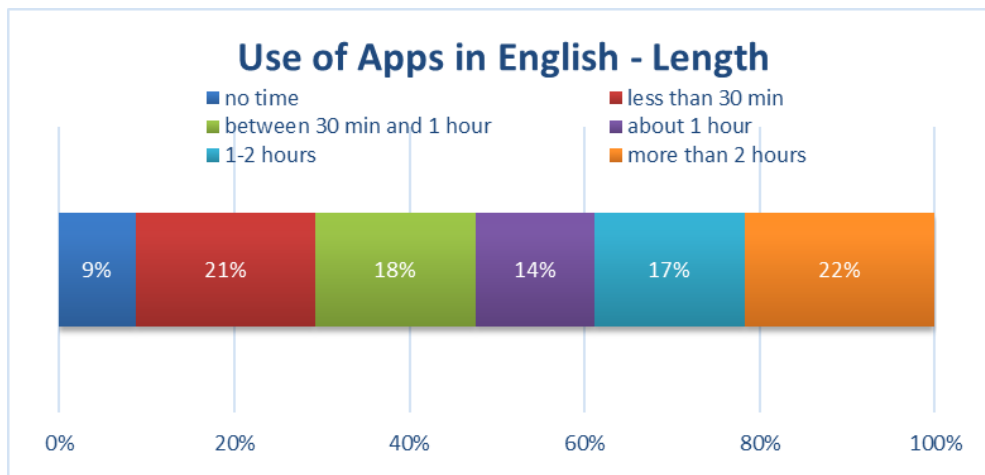
Figure 26
Length of time dedicated to the use of apps in English.
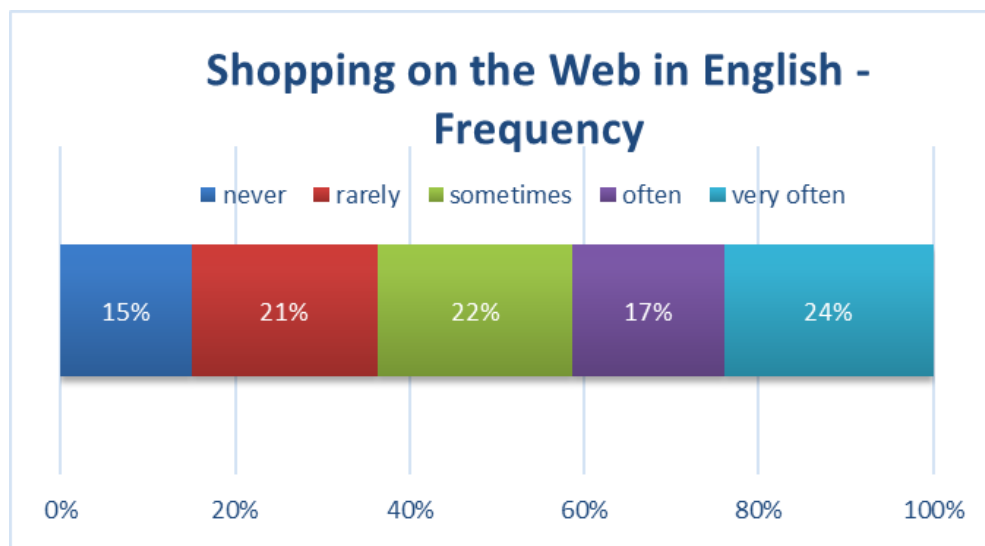
Figure 27
Frequency of use of e-commerce websites in English.

As Figure 27 shows, the frequency of use of e-commerce websites is variable among respondents, without major differences among the options.
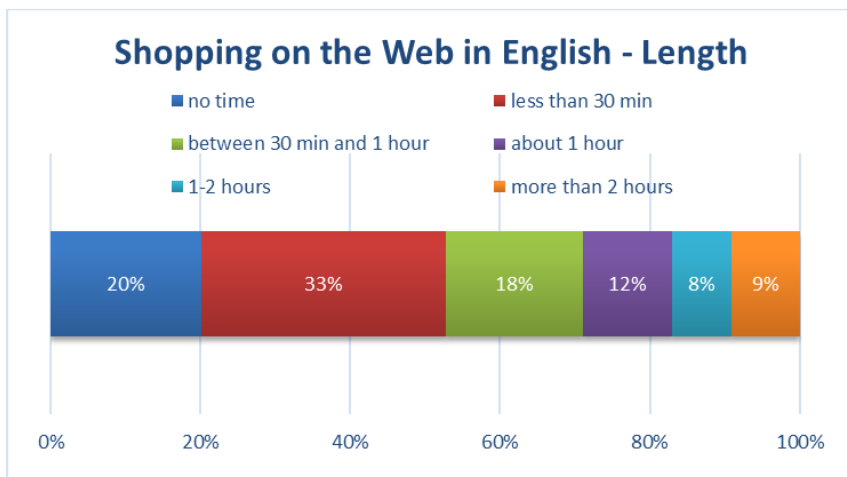
Figure 28
Length of time dedicated to the use of e-commerce websites in English.

Regarding the length of time spent shopping online illustrated in Figure 28, the most common answer (33%) was that respondents spend less than 30 minutes doing this activity, with 20% who do not spend time shopping online and 29% who spend one hour or more on e-commerce websites when they shop online.

Another question about online activity was whether respondents interacted online with native English speakers or non-native English speakers. Interaction with people in English online takes place often (115 answers, 17%) and very often (76 answers, 11%) for a minority of students. And, as can be seen from Figure 29, interactions happen indifferently with native English speakers or non-native English speakers.
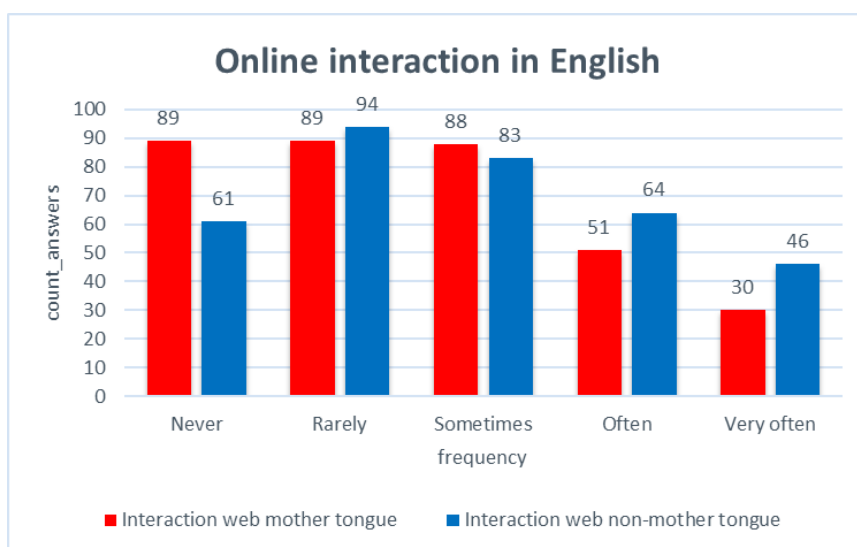
Figure 29
Interaction online in English with mother tongue and non-mother tongue speakers.

In the chart the red bars correspond to online interactions with English mother tongue speakers and the blue bars to interactions with non-native English speakers. They are paired by frequency, going from never to very often (see Section 4).

5.4. Influence of socio-cultural background on media exposure

The influence of social cultural background on media exposure was tested using the chi-squared test and the effect size Cramer's *V*. The levels of exposure to YouTube and to the Internet (outcome) were divided by a set of predictors connected with respondents' socio-cultural background: high school attended, parents' education, area of studies at university, perception of the importance of English, attitude towards English and self-assessed level of English. The results obtained with pair comparisons are shown in Table 1.

Predictor	$\chi 2$		p-value		Cramer's *V*	
	ExpLevelYT	ExpLevelNet	ExpLevelYT	ExpLevelNet	ExpLevelYT	ExpLevelNet
High school	54.31	35.96	6.008e-09*	1.788e-05*	0.15	0.12
Parents' education	8.45	8.44	0.08	0.08	0.05	0.05
Area of studies	10.38	6.28	0.03*	0.18	0.06	0.03
Importance of English	46.34	32.56	2.047e-07*	7.399e-05*	0.14	0.11
Liking of English	153.94	92.11	< 2.2e-16*	6.035e-12*	0.14	0.11
Self-assessed lev. of English	241.53	147.82	< 2.2e-16*	< 2.2e-16*	0.34	0.26

Table 1
Influence of socio-cultural background on media exposure.

The table shows the results of the chi-squared test divided by exposure to the two types of content analyzed: 'ExpLevelYT' is the exposure to YouTube in English and 'ExpLevelNet' is the exposure to the Internet in English. The chi-squared test for all the predictors gave a significant result with a small effect size, except for the self-assessed level of English, which had a moderate effect, and the predictors 'Parents' education' and 'Area of studies', which did not have a correlation with media exposure.

6. Discussion

The first research question was about respondents' habits in the use of YouTube and the Internet. According to the exposure indexes (see Section 4), over half of the respondents (56%) use YouTube in English and, among them, 54% are highly exposed. When considering YouTube frequency, a distribution of results across the various options can be observed. Instead, in the case of length, most students watch YouTube for a short period of time equal to less

than 30 minutes, or between 30 minutes and one hour per access. When accessing YouTube content in English, the majority of students use subtitles, mainly English subtitles (46%), but also Italian subtitles (28%).

The picture that emerges from the data about the use of YouTube is that this website is very popular among respondents for accessing videos in English, especially for a short time but very often. Moreover, students in this sample use subtitles most of the time, which can depend on their need for a written support to decode the audio but also to the availability of automatic English subtitles, provided by the platform.

In the case of the Internet, instead, only 35% of respondents use the Internet in English and, among them, only 22% are highly exposed. Among the various types of web content (see Paragraph 5.3) that students access in English, in this sample social media is very common: students read content on social media in English very often (61%). For length, there is more variability: some students spend a lot of time reading English content on social media (more than two hours) but others do not. Conversely, writing in English on social media is not a habit for respondents, since about half of them answered that they never write in English on social networks. In addition, the answers about the amount of English content on social networks confirm the pervasive presence of English online: 41% of respondents mentioned that 75% of the content they encounter on social media is in English and 37% declared that half of their social media content is in English. These answers are in line with the data about the frequency of reading English content on social media, to which 61% of answers were 'very often'.

Conversely, reading blogs and forums is a less frequent activity in this sample, with diverse usage patterns and, as far as writing is concerned, 79% of students answered that they never write in blogs and forums. Moving on to reading English content on websites, the results indicate that it is common among students, with varying usage durations. Podcasts and online radio are never or rarely used. English apps are widely used by students, with varying durations. E-commerce site usage is varied, with 33% of respondents spending less than 30 minutes shopping online.

Data collected about Internet usage among the students of the University of Salento who took part in this study point out that accessing contents in English on the Web is much less common than watching videos in English on YouTube. In addition, students read contents in English mainly on social media and not on websites and do not engage in active language production, such as writing, either on social media or on blogs and forums. Wikipedia (275 answers) is the website that most respondents access in English and many students also access the websites of English dictionaries (198 answers). Considering these answers, reading English contents on websites seems to be an activity more related to studying than to leisure. Audio-only input, such as radios and podcasts, were

chosen by a minority of respondents (see Paragraph 5.3), who prefer using audiovisual materials, which they can easily find on YouTube.

The second research question was about whether students' attitude towards English and self-assessed level of English influence their use of media in English. The answer is yes for both predictors. The perception that students have of the importance of English has a significant impact and a small effect size on their access to media in English ($p < 0.05$, $V= 0.14$ and $p < 0.05$, $V= 0.11$ for YouTube and the Internet respectively). When considering how much respondents like English, this trend is even stronger, especially for the use of YouTube, because the impact is significant and of moderate size. Higher proficiency and greater liking for the language correlate with increased exposure to both YouTube and Internet contents.

The third research question took socio-cultural background into consideration. In this case there are mixed results. The high school attended by respondents has a significant impact of small size, both in the level of exposure to YouTube and in the level of exposure to the Internet. Specifically, there is a larger share of highly exposed participants among those who attended a humanities, science or language high school (*licei*). Parents' level of education, instead, has no discernible effect on media exposure. In addition, the current area of university studies of respondents has a negligible influence on YouTube usage habits and no significant influence on the use of the Internet. Consistent patterns are observed across both YouTube and the Internet. Perceived importance and affinity for English have a positive correlation with media exposure. Self-assessed English proficiency significantly impacts media exposure. In summary, respondents predominantly engage in receptive rather than productive English activities, with a preference for audiovisual content, use of subtitles, and varied Internet usage patterns. Factors such as high school attended and language attitudes significantly impact media exposure.

7. Conclusion

The main trend observed in the data is that students are involved more in receptive (reading) than productive (writing) activities in English online. These findings are in keeping with what was observed by Kusyk (2017) and Krüger (2023) in France, Germany and Switzerland and are much as one would have expected since receptive activities are more passive and, therefore, simpler compared to productive language activities. As Pavesi and Ghia (2020) pointed out, the replication of the same trend implies that students' habits do not depend on nationality or source culture. This finding could be explained by the ever growing pervasive availability of the Internet connection and of free online contents in English, which gives students plenty of opportunities to engage in receptive activities in English.

$\mathcal{L}$ingue e
$\mathcal{L}$inguaggi

The role of receptive activities is essential to develop the language skills needed also for productive activities. Students access YouTube videos in English with subtitles much more than other web contents, showing a preference for audio-visual material over text-only material, which is in line with Sockett's (2014) findings.

Knowing that students usually read and listen to authentic English contents is useful for trainers to decide which activities they should focus more on in class and which ones can be assigned as self-study. On the basis of these findings, students would be comfortable with looking up contents online or doing reading and listening exercises as self-study. More guidance may be needed for speaking and writing, which could be planned more often as class activities than as self-study.

When looking at the participants' socio-cultural background, results do not show a significant influence of parents' education over English media exposure. This is not in line with Krüger's (2023) study, where there was an impact of participants' socio-cultural status on exposure, since students coming from families having a higher socio-economic status had an overall wider exposure to media in English. Even if they do not completely overlap, a higher socio-economic status is often associated with a higher level of education. The reason for the different patterns observed in Krüger and in this study could be that the respondents of this study are young adults, therefore older than Krüger's participants, who were adolescents. Older participants are likely to be more autonomous in the choice of the contents they access online, irrespective of their family background.

A promising research path to follow for the future is clarifying the connection between media exposure, self-assessed level of English and the attitude towards the English language: are students with a higher language level or liking of English more likely to access English media, or is it by accessing English media that students improve their level and like English more? A mixed-method approach, including both quantitative and qualitative data, is needed to answer this interesting question.

Bionote: Serena Ghiselli is a junior assistant professor at the Department of Interpreting and Translation at the University of Bologna. She worked as post-doctoral fellow at the same department and at the University of Salento, Lecce. Her research interests focus on cognitive interpreting studies, bilingualism, English language learning and corpus linguistics.

Author's address: serena.ghiselli3@unibo.it

Lingue e Linguaggi

References

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development across the Life Span*, Springer International Publishing, Cham.

Berns M., Claes M.-T. de Bot K. Evers R., Hasebrink U., Huibregtse I., Truchot C. and Wijst P. 2007, *English in Europe*, in Berns M., de Bot K. and Hasebrink U. (Eds.), *In the Presence of English: Media and European Youth,* Springer US, Boston, MA. eBook, pp. 15-42.

Bobbitt Z. 2023, *Statology*, in "Statology". https://www.statology.org/ (4.8.2023).

Council of Europe 2020, *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume* (council of Europe Publishing), Strasbourg. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (4.8.2023).

Czerkawski B. 2016, *Blending Formal and Informal Learning Networks for Online Learning*, in "The International Review of Research in Open and Distributed Learning" 17 [3], pp.138-156.

Dressman M. 2020, *Introduction*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp.1-12.

Kelley K. and Preacher K.J. 2012, *On Effect Size*, in "Psychological Methods" 17 [2], pp. 137-152.

Krüger M. 2023, *Media-Related Out-of-School Contact with English in Germany and Switzerland*, Springer, Wiesbaden.

Kusyk M. 2017, *The Development of Complexity, Accuracy and Fluency in L2 Written Production through Informal Participation in Online Activities*, in "CALICO Journal" 34 [1], pp. 75-96.

Kusyk M., Dressman M. and Sadler R.W. 2020, *Informal English Learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning,* John Wiley & Sons, Ltd., Hoboken, pp. 333-348.

Muñoz C. and Elke P. (Eds.) 2020, *Language Learning from Multimodal Input – Studies in Second Language Acquisition* (Vol. 42), Cambridge University Press, Cambridge.

Nation I.S.P. 1990, *Teaching and Learning Vocabulary*, Newbury House, New York.

Pavesi M. and Ghia E. 2020, *Informal Contact with English – A case study of Italian postgraduate students*, Edizioni ETS, Pisa.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling Informal Experience with English. An Instrument for Data Collection*, in "Lingue e Linguaggi" 59, pp. 303-32.

RStudio R.C.T. 2022, *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* https://rstudio.com/

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave Macmillan UK, Houndmills.

Tan E. 2013, *Informal learning on YouTube: Exploring digital literacy in independent online learning*, in "Learning, Media and Technology" 38 [4], pp. 463-477.

Tenny S. and Abdelgawad I. 2024, *Statistical Significance*, in *StatPearls*, StatPearls Publishing, Treasure Island, FL. https://pubmed.ncbi.nlm.nih.gov/29083828/ (6.11.2024).

Toffoli D. and Sockett G. 2010, *How non-specialist students of English practice informal learning using web 2.0 tools*, in "ASp. La Revue Du GERAS" 58, pp. 125-144.

Università del Salento 2023, *Statistiche*. https://www.unisalento.it/statistiche (25.9.2024).

W3Techs 2023, *Usage Statistics and Market Share of Content Languages for Websites*. https://w3techs.com/technologies/overview/content_language (14.6.2023).

Weiss N.A. 2011, *Introductory Statistics*, 9th ed.*,* Pearson, Boston.

Lingue e
Linguaggi

THE DYNAMICS OF MOTIVATION AND INFORMAL ENGLISH EXPOSURE
Evidence from a case study[1]

PAOLA LEONE[1], EMANUELA PAONE[2]
[1]UNIVERSITÀ DEL SALENTO, [2]UNIVERSITÀ DI BOLOGNA

Abstract – This qualitative case study investigates the dynamic nature of motivation in language learning, grounded in Dörnyei and Ushioda's L2 Motivational Self-System theory. This framework emphasises complexity arising from patterns rather than simple cause-and-effect relationships, focusing on motivational thinking rather than viewing motivation as a fixed entity. Data are open-ended interviews. The participants consist of 26 university students attending various undergraduate and graduate courses at a medium-sized university in Southern Italy. The findings highlight the intricate and multifaceted influences of social and family relationships, internal factors, and learning experiences on extramural English exposure, while also revealing emerging patterns in students' motivational thinking. The data underscore the pivotal role of teachers in fostering or diminishing motivation to learn a second language (L2) during childhood and adolescence. A significant similarity in experiences was noted during the transition from adolescence to adulthood, indicating a release from perceived external pressures from parents and teachers. By contrast, a distinct pattern of difference was observed in how teachers affected students' motivation. Effective teachers inspire students, while encounters with pedagogically challenged teachers do not always lead to demotivation. Some learners proactively seek alternative means, such as non-formal instruction, to enhance their language proficiency. This transition marks a significant shift in students' motivational thinking, as many began taking ownership of their language learning, aligning with their Ideal L2 Self. Additionally, a causal pattern emerged regarding the role of media in motivating language learning, with video games and English-language videos cited as catalysts for interest in the language. These media sources evoked positive emotions and sustained motivation. A motivational pattern related to the frequent exposure to video was also observed during the COVID-19 lockdowns, with students increasing their use of video-based content over reading, a shift interpreted as a reaction to the pandemic and its associated restrictions, which decreased their attention levels.

Keywords: motivation; English; second language; informal language exposure; learning practices; media.

---

[1] The authors worked together in analysing data, participated in discussions to address any differing opinions, and ultimately reached a consensus or shared perspective. The paper is the result of a writing partnership. Though, to accomplish some countries' academic requirements, we specify that Paola Leone is responsible for sections and subsections 1, 2, 3 (including 3.1, 3.1.1, 3.1.2); Emanuela Paone is responsible for sections and subsections 2.1., 2.2., 2.3, 4. Both authors are responsible for the abstract.

1. Introduction

English holds a unique position among second languages (L2) due to its status as a global language. It serves as a bridge connecting people from diverse linguistic and cultural backgrounds worldwide. Consequently, there are many contexts and resources in which this language is used. In diverse manners and extents individuals feel motivated to enhance language skills in this language both in and outside the institutional contexts.

This paper investigates how students perceive the factors driving university students to engage with English across various media platforms (e.g. TV series, music, social networks, blogs/vlogs, YouTube, video games, websites, online press, radio, and music, outside the institutional context). Methodological considerations related to qualitative research and findings regarding the factors influencing students' exposure to the English language are presented. Specifically, the study focuses on informal language contacts, which refer to any spontaneous language encounters by students that are not prompted by the institution (Caruana 2021; Pavesi, Ghia 2020; Pavesi *et al.* 2023; Sundqvist, Sylvén 2016). Throughout the paper, in line with Sundqvist and Sylvén (2016), this type of exposure will be referred to as 'extramural'.

The study is rooted in the L2 Motivational Self System theory (Ushioda 2001; Dörnyei, Ushioda 2011; Dörnyei *et al.* 2015). It encompasses three dimensions, all of which will be examined: a) L2 learning experience, concerning the perception of how the immediate learning environment and experience influence language learning (e.g. the teacher's impact; Dörnyei, Ushioda 2011); b) Ought-to L2 Self, reflecting the attributes that individuals believe they should possess to meet others' expectations (e.g. parents, friends) and avoid potential negative outcomes; c) Ideal L2 Self, representing the learner's positive self-image as a successful L2 user.

The approach described in this paper is in line with the complex dynamic systems theory, which asserts that 'no behavioral phenomenon has a single explanation' (Dörnyei, Ushioda 2011, pp. 98-99) and suggests that complexity within a phenomenon emerges through patterns rather than predictable cause-effect relationships. Embracing this view, termed socio-dynamic (Ushioda, Dörnyei 2017), implies moving beyond those cognitive and psychosocial theories (Gardner 1985) that consider the context as a static element and as an external variable influencing the individual. The fundamental objective of studies in this latter perspective is to create generalizable linear models capable of forecasting which types of motivation may influence certain learning behaviors within particular contexts. This understanding would consequently guide appropriate pedagogical interventions aimed at altering unproductive motivation patterns and ultimately enhancing learning behaviors and results. Instead, our approach aligns more closely with a contemporary situative

*L*ingue e
*L*inguaggi

perspective, which regards the individual as an integral part of the context and as an agent capable of influencing and being influenced by the dynamic and changeable nature of that context. As a result, the focus of inquiry moves away from viewing motivation as a static, quantifiable concept to considering motivational thinking. According to Ushioda (1998), this refers to students' evolving perceptions of the factors influencing motivation and their perspective on different facets of their motivational experience over time. In simpler terms, motivational thinking refers to how individuals conceptualise what motivates them and how their ideas about motivation evolve with increasing experience over time. Essentially, it pertains to our understanding and perception of what drives us to action and how this understanding evolves throughout life.

Following Dörnyei and Ushioda's (2011) recommendation, the current investigation employs in-depth interviews as a research instrument capable of capturing the inherent complexity of motivation, in the sense of motivational thinking. Data have been collected from 26 participants attending the University of Salento (Lecce, Italy) who took part in a nationwide quantitative study, carried out in 4 Italian Universities.

Extensive research (see Al-Hoorie, MacIntyre 2020 for a review) has shown that motivation is pivotal for language learning. However, the influence of this internal factor on informal language exposure has so far received little attention in the research literature. The current paper contributes to this field, recognising that language exposure in extramural contexts (Sylvén, Sundqvist 2012; Sundqvist, Sylvén 2016) is becoming increasingly important in individuals' daily lives.

The paper is organised as follows: Section 2 outlines the study's objectives and research design, including details on participants, data collection and analysis procedures. In Section 3, findings are presented. Finally, Section 4 provides insights into the main points and key findings of the study and suggests potential avenues for further research.

2. Researching university students' motivation

Dynamicity, a fundamental concept within Dörnyei and Ushioda's L2 Motivational Self-System theory, embodies the complex, changeable and intricate nature of the phenomenon of motivation. As described by Dörnyei (2009), dynamicity is characterised by the interplay of attracting or propelling forces, which are influenced by various factors. The strength of a particular force can be heightened or diminished by specific combinations of environmental and temporal conditions. Thus, something that seemed insignificant recently may now exert a temporary or lasting influence, depending on the circumstances.

The methodology chosen for the present study aims to capture the complexity of how motivation is perceived, emphasising the need to avoid concentrating solely on a singular deterministic explanation. In this light, the analysis required a broad perspective to identify and highlight patterns in the data that reflect this complexity.

Patterns, in this context, correspond to the elements of Dörnyei's model, where motivation is shaped by the interplay between Learning Experience, the Ought-to L2 Self and the Ideal L2 Self. These patterns (see also Saldaña 2013) can be manifested as follow:

- *Similarity* and *difference patterns* reflect for instance how students' L2 Selves are similarly shaped by their personal aspirations or goals or vary according to external pressures or expectations in distinct contexts.

- *Frequency* shows how often certain motivational factors, like the influence of personal goals or external obligations, arise in learners' experiences.

- *Sequence* refers to how motivational processes unfold over time, such as when the Learning Experience evolves and impacts the Ideal or Ought-to Self.

- *Causation* explores how one aspect of motivation, like self-confidence or the learning environment, can drive other behaviors.

- *Correspondence* links motivational patterns to related factors, like how Learning Experience might reinforce or hinder the development of the Ideal L2 Self.

The following subsections begin by describing the interview, namely research tool used for eliciting data. They go on to present subjects that have been used for the study and the coding method employed for the analysis. Finally, findings on Dynamicity are presented.

2.1. Interview structure and data collection procedure

In line with the methodological guidelines proposed by Ushioda and Dörnyei, we developed a semi-structured interview framework aimed at exploring students' motivational perspectives. The interview was structured into four sections, with each section featuring questions based on a synthesis of insights from two quantitative studies (Dörnyei, Ushioda 2011; Ushioda 2001). We incorporated elements from both the structured questions typically found in quantitative research and suggestions from qualitative interviewing literature. As the conversation progressed, questions related to each section arose organically, allowing topics to unfold naturally in the flow of dialogue.

The initial section (*What English language learning motivation means*, Ushioda 2001) aimed to explore the concept of motivation in English language learning and its underlying factors. It included questions such as *Ti definiresti una persona motivata allo studio della lingua inglese? In che misura? In che modo*

manifesti la tua (scarsa/elevata) motivazione all'esposizione alla lingua inglese fuori dai contesti istituzionali? Ti senti (più o meno) motivato/a rispetto ai tuoi compagni? ('Would you consider yourself a motivated person in the study of the English language? To what extent? How do you express your (low/high) motivation for exposure to the English language outside institutional contexts? Do you feel (more or less) motivated compared to your peers?').

The second section (*Ought-to L2 Self,* Dörnyei, Ushioda 2011; Ushioda 2001) delved into the influence of relationships such as family, parents, teachers, and friends in shaping the students' perception of their 'Ought-to L2 self' and its impact on their motivation. It included questions like: *C'è qualcuno che ti ha incoraggiato a studiare la lingua inglese/I genitori, gli amici*?; *è importante per te non deludere le loro aspettative?* ('Has anyone encouraged you to study the English language? For instance, your parents or friends? Is it important for you not to disappoint their expectations?').

The third section examined students' extramural exposure to English through media, exploring the modalities of access, the affordances, the attitude towards the English language. It included questions like: *In quali occasioni svolgi attività in modo del tutto autonomo – senza sollecitazione da parte dei tuoi professori – in lingua inglese? In generale guardi film in lingua originale anche in lingue diverse dall'inglese?* ('On what occasions do you engage in activities in an entirely autonomous manner – without prompting from your teachers – in the English language? Do you generally watch movies in their original language, even in languages other than English?).

Lastly, the fourth section sought to gain insights into students' 'Ideal L2 self (Dörnyei, Ushioda 2011), investigating how they envision themselves in the future and the role of English in their lives, e.g. *Come ti immagini in futuro? Quello che vuoi fare in futuro richiede la conoscenza della lingua inglese? Ti immagini in futuro a lavorare adoperando esclusivamente la lingua inglese?* ('How do you see yourself in the future? Does what you want to do in the future require knowledge of the English language? Can you imagine yourself working exclusively using the English language in the future?').

The data were collected and recorded by a PhD student who conducted the interviews in Italian. To create a relaxed atmosphere conducive to informal interactions, we intentionally chose a young interviewer. This approach aligns with Lambert's concept of 'wine and conversation,' as highlighted by Lambert (1968 in Spolsky 2000), and Ushioda's emphasis on the reliability of informal interactions in assessing motivation (Ushioda 2020). The interviewer prepared for the task by studying literature on conducting interviews. The interviews underwent a pilot phase, involving individuals not included in the sample. Subsequently, researchers and interviewers discussed the recorded interviews in terms of question quality and opportunities to further develop some issues that emerged during the conversations.

Lingue e Linguaggi

2.2 Participants

Participants were selected from a nationwide quantitative study (Pavesi *et al.* 2023; Pavesi this volume), focusing on individuals enrolled at the University of Salento, a mid-sized institution in Southern Italy. Specifically, participants were those who willingly volunteered to participate in our qualitative investigation. Considering the open-ended nature of the questions, we opted for a limited sample size of 26 students. The number was deemed appropriate as it provided a manageable amount of data, enabling the identification of common patterns and facilitating comparisons among students. The selection criteria comprised participants' level of informal exposure to the English language[2] (low/high) and their academic discipline, which included humanities, social sciences, and STEM fields. Although variations in degree programs could potentially influence student motivation data, this factor was not examined in the current study but may be investigated in future research. The sample is composed of 12 students with low exposure and 14 with high exposure to English. Although the participant group had a higher proportion of female participants (18 female and 6 male students), this was not considered significant to the study for two key reasons. First, gender differences were not included in the focus of the research (for those interested in this topic, see Yashima *et al.* 2017). Second, contemporary L2 acquisition theories view gender as a social identity that intersects with other identities such as ethnicity and social class (Norton, McKinney 2011). This perspective recognises that gender is inherently complex and varies from person to person. Consequently, behaviors cannot be generalised based solely on this aspect (McKay, Wong 1996).

Each participant was offered a voucher worth 25 euros to be redeemed at a bookshop in exchange for the participation in the interview. In accordance with Ushioda (2001), this strategy effectively fostered a collaborative attitude.

Table 1 outlines the composition of disciplinary groups as elicited in the previously mentioned quantitative study, shedding light on their demographic characteristics.

	N. Subjects	Age range
SAA Group (Humanities)	15	21-27
SAB Group (Social sciences)	5	19-21
SAC Group (STEM)	6	21-25

Table 1
Subject demographics.

[2] Exposure indices were created by multiplying the average frequency of activities by their average duration (Pavesi, Ghia 2020; Pavesi *et al.* 2023).

*L*ingue e
*L*inguaggi

Researching motivation for informal English exposure and learning. A case study

111

2.3 Coding and analysis procedure for highlighting dynamicity

To conduct this investigation, we adopted a qualitative research design using content analysis (Neuendorf 2017; Saldaña 2013), using the Transana software.[3] The interviews were automatically transcribed through Speechmatics, a plugin integrated with Transana.

A preliminary content analysis was conducted on a limited dataset to formulate initial codes, guided by Charmaz's (2001) suggestion that, in qualitative data analysis, a code functions as a construct, imbuing specific data pieces with interpreted meaning.

The coding decisions were grounded in the theoretical framework guiding this research, incorporating key elements of Dörnyei and Ushioda's L2 Motivational Self-System, the emergent conceptual framework, and the research objectives. Thus, we developed codes such as 'Ideal L2 Self' and 'Ought-to L2 Self', aligning with the concepts outlined in the model. Moreover, our study goals influenced certain choices in defining these codes. For instance, given that our research focuses on understanding the dynamic nature of students' motivational thinking regarding informal language exposure, we recognised the need to not grant particular prominence to classroom learning experience, as in Dörnyei and Ushioda's model. Instead, we expanded our scope to include the roles of teachers and peers as additional factors influencing motivation, categorised under 'Factors Affecting Motivation' (FAM), with a specific focus on the relationship dimension rather than the classroom learning experience as delineated in the Motivational Self-System model. This category (FAM) also encompasses parents and family as subcodes. Furthermore, we introduced a category labeled 'Media Access' to precisely capture students' extramural English language input. Conversely, the 'practice' code was utilised to highlight behaviors associated with the use of various resources, outside the classroom and unsolicited by teachers.

Following the preliminary coding, all data underwent two cycles of analysis. The first cycle involved applying initially defined codes to the entire dataset, accompanied by a comprehensive review and refinement of the previously established codes. In the second coding cycle, the meanings previously attributed served as a foundation for a more refined task, which included identifying patterns, establishing more in-depth categories, and developing theories.

This analysis utilised Transana's code mining function, which organises data based on their codes (Saldaña 2013), and streamlines analytical procedures, thereby establishing a more stable bridge between data and their significance according to the adopted analytical framework.

---

[3] Woods D. 2024, *Transana v5.2x*. https://www.transana.com. Madison, WI: Spurgeon Woods LLC.

Employed coding methods were mixed and matched (Saldaña 2013). Consequently, we primarily utilised two types of coding:

- Structural or holistic coding: this method was instrumental in generating comprehensive lists of major categories, such as the previously mentioned FAM and the Ideal L2 Self.

- Pattern coding: used to explore explanations within the data, according to the model. A code such as 'dynamicity' was employed, particularly to elucidate fluctuations in language learning motivation.[4] The 'dynamicity' code cuts across the other developed structural codes (i.e. Factors Affecting Motivation, Ought-to L2 self, Ideal L2 Self, Media Access).

To grasp the dynamic essence of motivation within the data, we sought contextual factors (e.g. relationships with teachers, parents, and peers) or internal factors (individual self-concepts: ought-to L2 self and ideal L2 self), or media access and practices that could either enhance or diminish students' opportunities for exposure to English. This task was facilitated when students explicitly used the word 'motivation' or 'demotivation', or related vocabulary (i.e. 'he demotivated') following changes in the broader context (e.g. different teachers) or time-related factors (e.g. before it was X, now Y). Nevertheless, it is crucial to acknowledge the complexity inherent in capturing these oscillations. As Ushioda (2009, p. 215) contends, attempting to delineate a linear causality between contextual factors and individuals' self-perceptions oversimplifies the intricate relationship.

Although some interviewees discussed changes in their learning experiences and language input, these were not coded as motivational dynamicity if they did not influence motivation. For instance, one participant briefly used language-learning apps like Duolingo and Babbel but did not find them effective, leading to disuse. This reflects a shift in resource preference rather than a change in motivation for language exposure.

Similarly, another participant described a shift in her approach to learning unfamiliar words while watching media in a foreign language. While she initially paused to note and search for word meanings, she no longer feels the need to do so, likely due to improved language competence. This change in learning practice was not coded as motivational Dynamicity, as it did not directly affect motivation.

Essentially, the coding and analysis approach primarily aimed to capture the complexity and uniqueness of participants' motivational responses to specific life events and experiences, while avoiding reductionist cause-and-effect interpretations.

---

[4] We did not have the opportunity to employ 'In Vivo coding', an approach that involves using participants' direct language as codes to anchor the analysis in their perspectives.

3. Findings: Insights into the dynamic nature of motivation

The results highlight the dynamic nature of motivation as it emerges in the narratives of interviewees' personal lives. Analysing various key factors discussed in response to interviewer questions reveals patterns in the data, illustrating the complexity and variability of the phenomenon. Additionally, the interaction between environmental and temporal conditions influencing English learning and exposure brings out unique characteristics and behaviors in individuals.

This section is organised as follows: subsection 3.1 delves into school, social and family relationships as factors influencing student exposure to English; subsections 3.2. and 3.3. illustrate how dynamicity emerges in interviewee' motivational thinking, specifically, in terms of ought-to L2 self and ideal L2 self, and media access and practices.

3.1. School, social and family relationships impacting students' motivation

This subsection explores the relationships students believe affect changes in their lives and shape their motivation toward English language exposure. Early in the interviews, students reflect on whether they feel motivated or unmotivated, revealing factors that either boost or hinder motivation. These include the influence of teachers, friends, parents, and experiences such as travel or time spent abroad.

3.1.1. Teacher's role dynamics: Catalysts and barriers to student motivation

Teachers play a crucial role in shaping student motivation, either fostering it or causing demotivation. Students describe their experiences with 'good' and 'bad' teachers, focusing on teaching methods, relationships with students, and overall attitudes. A shift from a good to a bad teacher reduces motivation, while the reverse increases engagement and interest in studying English. When it comes to characteristics, effective teachers are those that inspire students to develop a passion for the subject (SAB0512: '*nelle scuole superiori mi è – mi ha appassionata tanto perché la professoressa era un sacco in gamba e ti faceva insomma amare la materia*'; 'in high school I was so passionate about it because my teacher was really skilled and she/he made you love the subject').

Students also value certain aspects of the 'instructional context' (Dörnyei, Ushioda 2011, p. 26) that are attributable to teachers. During their childhood 'very playful and enjoyable teaching activities' made interviewees love the English language, as stated by SAA0517: '*attività didattiche molto giocose, molto piacevoli anche da quel punto di vista e senza neanche*

accorgermi stavo imparando un sacco di cose quando sono arrivata alle medie' ('very playful and enjoyable teaching activities even from that perspective, and without even realising it, I was learning a lot when I reached middle school).

Conversely, in more advanced schools, attention towards language accuracy is appreciated and the importance of the study of grammar is estimated as relevant for shaping student's English expertise. Indeed, SAA0575 says: *'la professoressa ci teneva molto specie a inquadrarci dal punto di vista grammaticale poi effettivamente mi è servito e durante gli anni'* ('The teacher cared a lot, especially about helping us focus from a grammatical point of view, which actually served me well over the years').

On the other hand, it is positive to be a remarkably innovative teacher, suggested by SAA0575: *'avevo dei docenti che erano veramente all'avanguardia per l'inizio 2000'* ('I had teachers who were truly ahead of their time for the early 2000s'). For instance, a student appreciates one of her teachers who uses more updated teaching methodologies paying attention to communication and not to language accuracy. In fact, she mentions what the teacher used to say, as shown in extract 1.

(1) SAA0556_FAM_TEACHER_DYNAMICITY[5]
SAA0556: *a me non interessa una frase sintatticamente corretta, mi interessa che voi sappiate parlarlo o che questa è la cosa più importante. Mi interessa che voi siate fluidi perché poi la correzione sintattica arriva nel tempo e questo è ovviamente parliamo 2010, quindi i primi del 2010 è un metodo di insegnamento molto avanti rispetto a quello che che che viene sempre insegnato a scuola quello grammatico traduttivo.*
"I don't care about a syntactically correct sentence; what matters to me is that you know how to speak it or that this is the most important thing. I care that you are fluent because then the syntactic correction comes with time, and obviously, we're talking about 2010, so early 2010 is a teaching method far ahead compared to what is always taught in school, the grammatical-translation approach."

A teacher must also be 'acknowledging,' valuing students' skills and competences, as asserted by SAA0556, *'più che spronarmi hanno riconosciuto le mie abilità'* ('More than just encouraging me, they recognised my abilities'). In this statement, the interviewee emphasises that acknowledging students' abilities is an effective way to empower them. She later mentions that she has never been in situations where a teacher told her that she was worse than she

---

[5] Each extract reports the clip name created for analysis, which includes: a) information regarding the interviewee's degree course (SAA GROUP, HUMANITIES; SAB GROUP, SOCIAL SCIENCES; SAC GROUP, STEM; see section on Participants); b) a unique identification number; c) one or more abbreviated words capturing the salient features of the data contents.

actually was, highlighting the importance of interacting with educators who never diminish their value.[6]

Teachers can also hinder students' motivation and progress. These language instructors, perceived as 'pedagogically challenged', evoke negative feelings related to a particular subject, as highlighted by an interviewee who states that her teacher did not motivate her and instilled fear of English in her (SAA0521: '*quella del primo anno non mi motivava infatti avevo quasi paura diciamo del di questa materia*'; 'The one from the first year didn't motivate me; in fact, I was almost afraid of this subject.').

Poor and demotivating relationships with teachers can stem from various reasons. For example, a student suspects that her teacher was jealous or felt 'offended' by the student's pre-existing knowledge of English: (SAA0517: '*Alle superiori in realtà è stato controproducente il mio sapere già un pochino l'inglese perché avevo professoresse posso dire che si sentivano un po' offese dal fatto che io sapessi già qualcosa*'; 'In high school, it was actually counterproductive that I already knew a little English because I had teachers, I can say, who felt somewhat offended by the fact that I already knew something'). For this reason, she perceived that the teacher 'threw a bit of a spanner in the works for me' (SAA0517: '*Però questa è una mia illazione e quindi diciamo che mi ha messo un po' i bastoni tra le ruote*'; 'However, this is just my speculation, so let's say she threw a bit of a spanner in the works for me'). A student underlines that, as a personal initiative, to compensate for the poor instruction he received in school, he attended a course for language certification (SAA0575: '*durante gli anni del liceo invece sapevo di quella di quel lato negativo quindi ho provveduto a studiare con altri docenti all'esterno*'; 'During high school years, however, I was aware of that negative aspect, so I took it upon myself to study with other teachers outside of school'). Thus, non-formal language instruction was a compensatory strategy for instruction offered by institutional state education. Still, informal learning in autonomy is not mentioned as a solution to the problem. Besides, SAA0575's experience emphasises the intricate nature of the system and the intertwined contextual factors that contribute to the maintenance or emergence of motivation. We see here that motivation, arising organically from a person's unique experiences and interactions, is protected like a precious asset. A similar scenario unfolds with SAA0517. In her narrative, she emphasises how personal interests and self-directed learning act as safeguards against the erosion of motivation for learning English within the school environment. Reflecting on her encounter with a teacher lacking motivation and who

---

[6] SAA0556_FAM_TEACHER_DYNAMICITY

SAA0556: *una cosa importante da dire, così come anche a livello di insegnamento, non ho mai avuto situazioni in cui un professore mi ha detto che ero peggio di quello che ero*.

"An important thing to say, as well as in terms of teaching, I have never had situations where a teacher told me that I was worse than I was."

devalued her (as discussed earlier), she underscores the crucial role of her own initiative. Had she not sought resources beyond formal education, her enthusiasm for English might have dwindled. However, by prioritising cultural immersion over mere language acquisition, she successfully sustains her interest autonomously. In fact, she explains: '*se non mi fossi messa io da sola, magari a fruire di materiali esterni al contesto didattico, probabilmente il mio interesse per l'inglese sarebbe morto sinceramente, però forse per questa attenzione un po' più all'aspetto culturale e non linguistico, sono riuscita a portarlo avanti personalmente*' ('If I hadn't taken the initiative myself, perhaps by accessing materials outside the educational context, my interest in English would probably have honestly died however, perhaps due to this slightly more cultural rather than linguistic focus, I managed to sustain it personally').

In conclusion, the findings align with established research, indicating that learners view teachers as influential figures affecting their motivation and behavior (Atkinson 2000; Radel *et al.* 2010; Roth *et al.* 2007). Interviewees describe effective teachers as catalysts for motivated learning, while those with pedagogical challenges are seen as demotivators. This interaction between student and teacher motivation can lead to either positive or negative outcomes. The data also highlight the significance of real-life experiences in understanding motivation dynamics. For instance, one interviewee actively seeks lessons outside the institutional setting to counteract potential demotivation from ineffective teachers, supporting Ushioda's (2009) assertion that cause-effect relationships in motivation are complex and not always predictable. Importantly, not all students respond to pedagogically challenged teachers with demotivation; individual attitudes can drive learners to pursue alternative enriching experiences that help preserve their language competence.

Overall, the investigation reveals both similarities and differences in interviewees' perceptions of teachers' roles. While the influence of teachers is consistently recognised across different students, its impact on shaping behavior varies. Additionally, the complexity of motivation is underscored by the interplay of environmental and individual factors that either promote or inhibit motivation for English language exposure.

3.1.2. *Parental and peer Influence on motivation: Interviewees' perspectives*

In the interviews, parents and peers do not appear to play an active role in motivating students, as they are not frequently mentioned. However, when parents are discussed, they are often portrayed as supportive. For instance, they may suggest experiences that enhance their children's learning (SAA0159_FAM: '*questi viaggi studio in realtà non c'entravano con la*

scuola, ma sono è stata proprio mia mamma a scoprirli e a consigliarmi';
'These study trips weren't actually school-related but it was my mom who
discovered them and recommended them to me'), or they may be depicted as
individuals who encourage students to pursue their interest in the language. In
extract 2 SAA0169 recounts how her parents paid for a certification course for
her and explains their rationale and supportive attitude.

> (2) SAA0169_PARENTS
> I: *Quindi questo [...] corso che hai fatto che si può definire extra
> scolastico non l'hai fatto perché sei stata motivata, stimolata o perché
> semplicemente volevi farlo tu, un po' tutti e due?*
> SAA0169: *Perché i genitori mi hanno detto visto che sei così brava,
> perché non devi avere delle certificazioni che dimostrano che tu hai un
> livello di inglese che ti potrebbe servire per il mondo del lavoro. E io mi
> sono convinta per questo.*
> "I: So, for the course you took, which can be considered extracurricular,
> did you do it because you were motivated, stimulated, or simply because
> you wanted to do it, or a bit of both?
> SAA0169: My parents told me, 'Since you're so good, why don't you get
> certifications that prove you have a level of English that could be useful
> for the world of work?' And I was convinced because of this."

The correlation between various family characteristics and practices and
academic performance has been extensively studied by educational
psychologists. To the best of our knowledge, there are currently no studies that
explore the correlation between family characteristics and exposure to the
English language outside of formal educational settings, independent of
institutional solicitation (although some research has addressed this topic in the
Italian context; see Sordella 2016). However, when considering research on
school achievements, it emerges that a supportive family atmosphere plays a
significant role compared to the other three parenting factors traditionally
recognised as influential in shaping student motivation: appropriate timing of
achievement demands, confidence in children's abilities, and having highly
motivated role models (Eccles *et al.* 1998; Dörnyei, Ushioda 2011).

When considering social factors, peers do not appear to be particularly
influential. Students do not perceive a temporal evolution in friends and
schoolmates motivating them, but rather, the interviewees highlight the varying
degrees of motivation exerted by friends as a situational influence. There are
contexts in which a student feels positively influenced by her peers and contexts
in which she is not, as underlined by SAB0327 in the following extract.

> (3) SAB0327_FRIENDS
> I: *ci sono state occasioni in cui magari confrontandoti con non lo so un
> amico un collega un po' più motivato questo è servito da stimolo a te per
> approcciarti ancora meglio alla lingua oppure no*

> SAB0327: *Sì e no. Nel senso, dipende anche dal contesto, magari in contesto scolastico, con qualcuno con cui diciamo eravamo sulla stessa linea d'onda e allora la motivazione veniva fuori. Nel senso che magari non lo so, anche per scherzare ti mettevi a masticare qualche parola così, anche se non era corretto. Invece magari con chi? Appunto. Come ho detto prima alcuni colleghi anche adesso con cui magari l'inglese non lo calcolano proprio diciamo rimane lì rimane piatta la questione.*
> "I: There have been occasions when, perhaps, comparing yourself with, I don't know, a friend, a colleague who was a bit more motivated, served as a stimulus for you to approach the language even better, or not.
> SAB0327: Yes and no. I mean, it also depends on the context. Maybe in a school context, with someone with whom, let's say, we were on the same wavelength, then motivation would come out. In the sense that maybe, I don't know, even just for fun, you started chewing on some words, even if they weren't correct. Instead, maybe with whom? Exactly. As I said before, some colleagues even now, with whom maybe they don't really consider English, it just stays there, the matter remains flat..."

In conclusion, parents and friends do not appear to exert significant influence. It should be observed, however, that this research likely lacks sufficient data in this regard for two main reasons. First, the scope of the interview was limited, focusing on whether students acknowledged the direct influence of parents and peers on their decisions. As a consequence, it did not examine how parental and peer lifestyles and behaviors may have affected the interviewees' choices. Second, the adult age of the interviewees may contribute to explaining the lack of data. Previous research indicates that parents and peer groups have a greater impact during earlier educational stages (e.g. Pomerantz *et al.* 2007 for parental influence; Berndt, Keefe 1995; Lewis, Sullivan 2007 for peer influence). Particularly among adolescents, peer groups can exert substantial influence, often surpassing that of parents during this period (Dörnyei, Ushioda 2011). By contrast, as adults, the interviewees' identities and self-concepts are more established. Additionally, as university students transitioning into new environments and social circles, social influences tend to play a less prominent role in their decision-making.

Dörnyei and Ushioda's theory suggests that the process of identification, underlying integrativeness in Gardner and Lambert's (1959, 1972) model, might be better explained as an internal process of identification within the individual's self-concept, rather than identification with an external reference group, whether peers or the community speaking the target language. This internal dimension of L2 learners' motivational thinking is further discussed in the following subsection.

3.2. Ought -to and ideal L2 selves: Exploring evolving identities

As previously mentioned, the ought-to self-guides have a prevention focus, regulating the absence or presence of negative outcomes, and are concerned with safety, responsibilities, and obligations (Dörnyei 2005, p. 101), whereas the 'ideal L2 self' orientation has a promotion focus, concerned with hopes, aspirations, advancements, growth, and accomplishments. Our examination reveals numerous factors influencing changes in the dimension of students' ought-to self, alongside discernible patterns. In fact, certain factors appeared to be consistent across some students, for instance the temporal factor along with personal evolution of their self-perception. As students progress through developmental stages, their growth is marked by transitions towards heightened autonomy in decision-making processes, reflecting a decreased reliance on external pressures, such as parental or pedagogical expectations, and a concomitant increase in self-directed motivations, as made explicit in extracts 4 and 5:

(4) SAC0523_OUGHTTOSELF_DYNAMICITY_PARENTS
I: *è importante per te non deludere le aspettative della tua famiglia in merito all'argomento lingua inglese?*
SAC0523: *Forse quando ero più giovane sì, però poi chiaramente maturando e passando gli anni si comprende che certe scelte, fare un corso di lingua o anche studiarlo da autodidatta − dipende sempre però dalle basi che hai − è essenziale se vuoi raggiungere determinati obiettivi o se vedi che quelle competenze potrebbero servirti in ambito non solo istituzionale ma anche di vita e quindi poi diventa una spinta intrinseca quella di migliorarsi.*
"I: Is it important for you not to disappoint your family's expectations regarding the English language topic?
SAC0523: Maybe when I was younger, yes, but then as I matured and the years passed, you understand that certain choices, taking a language course or even studying it as a self-taught learner – it always depends on the foundations you have – are essential if you want to achieve certain goals or if you see that those skills could be useful not only in institutional but also in life settings. So, it gradually becomes an intrinsic drive to improve."

(5) SAA0170_OUGHTTOSELF_DYNAMICITY_PARENTS
SAA0170: *Prima quando ero più piccola e quindi ero meno sicura delle mie abilità sicuramente di più adesso che ritengo che siano appunto consolidate magari di meno però c'è stato sicuramente questo questo fattore.*
"Before, when I was younger and therefore less confident in my abilities, definitely more so now that I believe they are indeed consolidated, perhaps to a lesser extent, but there has certainly been this factor."

Both SAC0523 and SAA0170 exhibit a progression in their attitudes towards English language learning, initially shaped by familial expectations. SAC0523

acknowledges that family expectations played a significant role in his younger years, but as he matured, he recognised the intrinsic value of language proficiency for personal and professional development. This shift reflects a movement from external pressures to a focus on self-improvement and the practical benefits of language skills, aligning with the ideal-self dimension, which emphasises self-promotion over avoidance of negative outcomes.

Similarly, SAA0170 highlights how her family's expectations influenced her early on, particularly when she lacked confidence in her language abilities. Over time, as her skills solidified, the pressure to meet these expectations lessened, indicating a transition from external validation to a more self-assured sense of competence. Both cases reflect a common theme of evolving motivations, shifting from external to internal drivers for learning, resonating with the ideal-self framework.

Concurrently, environmental catalysts play a crucial role in shaping the dynamicity of individuals' 'ought-to L2self.' Life-altering events, such as the loss of a parent, can provoke profound shifts in self-conceptualization, particularly in instances where the deceased individual represented a source of external pressure, as in the case of the student SAA0556:

(6) SAA0556_OUGHTTOSELF_DYNAMICITY_PARENTS
SAA0556: *Mia madre purtroppo è venuta a mancare nel 2021 e in questa situazione devo dire che la sua morte mi ha portato sicuramente dei lati molto negativi, però mi ha comportato, mi ha portato anche dei lati estremamente positivi. Per esempio, il fatto che io sia riuscito ad appropriarmi delle mie capacità senza applicarle sugli altri, perché ripeto gran parte della mia vita in generale non solo a livello di dell'utilizzo dell'inglese ma in generale su tutto su tutti gli ambiti derivava dal fatto di in qualche modo doverle far piacere ed è una cosa che non poteva restare per sempre [...]. Per questo motivo io oggi con mio padre mi trovo molto meglio perché lui vuole risultati e io i risultati glieli porto perché non ha non ha bisogno di sapere tutto [...] mi sento molto meno diciamo attanagliato da questa cosa, semmai la sensazione di dover rendere al meglio possibile adesso lo applico a me stesso come senso di sfida.*
"Unfortunately, my mother passed away in 2021, and in this situation, I have to say that her death certainly brought some very negative aspects, but it also brought some extremely positive ones. For example, the fact that I managed to develop my skills without applying them to others, because I repeat, much of my life in general, not only in terms of the use of English but in general in all areas, stemmed from the need to please her, and that couldn't last forever [...]. For this reason, I now get along much better with my father because he wants results, and I bring him the results because he doesn't need to know everything [...]. I feel much less overwhelmed by this, and if anything, the feeling of having to perform at my best now applies to myself as a sense of challenge."

In this excerpt, the student reflects on the multifaceted impact of his mother's passing: while acknowledging the negative aspects, he also highlights the positive outcomes, particularly in terms of personal growth and autonomy. He describes a shift from seeking external validation, primarily from his mother, to a more self-directed approach in utilising his skills and meeting expectations, particularly with his father.

Regarding the dynamicity within the ideal-self dimension, our data unveil a notable trend: certain students, who previously held a clear vision of their future selves embracing English, have demonstrated a shift in perspective, as elucidated in the following excerpt:

> (7) SAA0169_IDEAL L2 SELF_DYNAMICITY
> I: *Per ragioni lavorative per ragioni di studio, perché ti piace?*
> SAA0169: *Cioè sì, mi riesce bene utilizzare l'inglese principalmente nella mia vita, nel futuro però sinceramente cercherei di non di non accantonare la lingua e cercare sempre di mantenere un minimo di l'italiano lo spagnolo giapponese delle altre lingue preferirei magari 90% usare l'inglese e quel 10% lo divido tra le altre lingue. Comunque sarebbe un peccato lasciare l'italiano come lingua, lasciarla così nel cassetto, dimenticatoio, certo, anche perché quando ero piccola volevo tantissimo andare a vivere in Inghilterra. Mi vedevo già quindi a vivere principalmente in con usando la lingua inglese. Però riflettendoci vorrei non lasciare le altre lingue. Vorrei comunque dare un 50 e 50 a tutte le lingue.*
> "I: For work reasons, for study reasons, why do you like it?
> SAA0169: Yeah, I'm good at using English mainly in my life. In the future, however, I would honestly try not to set aside the language and always try to maintain a minimum of Italian, Spanish, Japanese, and other languages. I might prefer to use English 90% of the time and divide that remaining 10% among the other languages. However, it would be a shame to leave Italian as a language, to leave it in the drawer, forgotten, of course, because when I was little, I really wanted to go and live in England. I could already see myself living mainly by using the English language. But on reflection, I would not want to leave the other languages. I would still like to give an equal share to all languages."

The student reflects on her language use and future aspirations. She perceives herself as proficient in English and acknowledges its predominant role in her life. However, she expresses a desire to maintain a balance between English and other languages, such as Italian, Spanish, and Japanese. Despite her childhood aspirations to live in England and predominantly use English, she now values the importance of preserving her proficiency in other languages. She wishes to allocate equal attention to all languages, emphasising the significance of linguistic diversity and the retention of her plurilingual skills. This sensibility for languages beyond English may stem from her enrollment in a Linguistic Mediation degree program, which exposes her to various

languages and fosters a profound understanding of plurilingualism and its significance in her life. The need to preserve personal identity is exemplified also in the following excerpt:

> (8) SAA0556_IDEAL L2 SELF_DYNAMICITY
> SAA0556: *Io vorrei essere cioè nella mia testa io voglio essere un cittadino del mondo cioè non voglio essere una persona che viene da Lecce o da Taranto o da Bari voglio essere una persona che non si capisce. Però la verità dei fatti è che anche se quella è un po' la parte, un po' l'anello debole di tutto il discorso, se cambiassi troppo il mio accento non mi sentirei più me.*
> I: *Ho capito.*
> SAA0556: *Cioè, credo che sia un tratto caratteriale, un tratto personale.*
> I: *Ok, è una cosa bella questa c'è comunque, lo vuoi tra virgolette mantenere, vuoi preservare, però magari lo vuoi anche un po'*
> SAA0556: *Voglio affinarlo, quello senz'altro. Forse se dovessi affinare qualcosa in questo momento mi concentrerei molto sul migliorare il tipo di pronuncia, magari avvicinarlo, cioè, fare un po' di ibrido cercare di trovare una mia versione della pronuncia ovviamente senza fare strani mix cose, cioè, mantenendo sempre il mio inglese che a questo punto io direi che è globalizzato perché non ha influenze relative al British o all'americana o all'Australia.*
> "SAA0556: I mean, in my head, I want to be a citizen of the world, you know? I don't want to be someone who comes from Lecce or Taranto or Bari; I want to be someone who isn't easily identified. But the truth is, even though that's a bit of the weak link in the whole story, if I were to change my accent too much, I wouldn't feel like myself anymore.
> I: I see.
> SAA0556: I think it's a personality trait, a personal characteristic.
> I: Okay, that's a beautiful thing to have, though. Do you want to maintain it, to preserve it, but maybe also.
> SAA0556: I want to refine it, definitely. Perhaps if I were to refine something right now, I would focus a lot on improving my pronunciation, maybe getting it closer, you know, making it a bit of a hybrid, trying to find my own version of pronunciation, obviously without making strange mixes, you know, always maintaining my English, which at this point I would say is globalised because it doesn't have influences related to British or American or Australian."

The student reflects on his evolving identity in relation to the English language and his accent. He expresses a desire to be recognised as a global citizen rather than being tied to a specific city or region. While he acknowledges that his accent is a core element of his identity, he recognises that altering it too much could undermine his sense of self. Therefore, he seeks to strike a balance – refining his accent to align with a global identity while preserving its original essence. This evolution in his ideal L2 self is driven by his understanding that English is now a global language, no longer confined to British or American norms. His

Researching motivation for informal English exposure and learning. A case study

123

perspective highlights the dynamic and evolving nature of context, demonstrating how it shapes his motivations and identity.

3.3. Dynamic media access and practices: Exploring evolving trends

Analysis reveals that students' engagement with English-language media is shaped by both environmental factors and personal interests. Shifts in living conditions and access to technology significantly influence media consumption habits, while individual preferences and the positive emotions from these activities also play a key role. The COVID-19 pandemic, for example, has driven increased reliance on digital platforms for entertainment and learning due to changes in lifestyle and routines, as exemplified in the following excerpt:

(9) SAA0517_COVID_READING
I: *Ok, di solito sui social network incontri più testi scritti o orali?*
SAA0517: *Orali ma perché credo che siano anche più facilmente fruibili, soprattutto dopo la pandemia, o almeno io mi rendo conto che la mia soglia dell'attenzione dopo la prima quarantena del 2020 si è estremamente abbassata. Quindi un contenuto ma anche ad esempio un articolo di giornale in cui vedo sette sette colonne in fila di articolo. Sono un po' demotivata dal leggerlo o comunque me lo spezzetto in più fasi perché mi rendo conto che anche se continuo a leggerlo tutto, non riesco a interiorizzare magari il concetto che l'autore o l'autrice vuole esprimere. Invece un video, essendo anche magari più coinvolgente visivamente, nel senso magari colori, gesti, ma anche semplicemente il fatto che sia una persona a parlare interrompendo il discorso, magari con una battuta o comunque con una parentesi, penso che sia per me più facilmente fruibile.*
"I: Okay, usually on social networks, do you encounter more written or spoken texts?
SAA0517: Spoken, but because I believe they are also more easily accessible, especially after the pandemic, or at least I realise that my attention span has greatly decreased after the first quarantine in 2020. So, for example, a newspaper article where I see seven columns in a row of text. I'm a bit demotivated to read it or, in any case, I break it down into multiple phases because I realise that even if I continue to read it all, I may not be able to internalise the concept that the author wants to express. Instead, a video, being perhaps more visually engaging, in the sense maybe of colours, gestures, but also simply the fact that it's a person speaking, interrupting the discourse, maybe with a joke or anyway with a parenthesis, I think it's more easily accessible for me."

The finding that videos are more engaging than written texts resonates with previous research on informal exposure to English through media. Prior studies have highlighted that an inherent attraction to language is closely intertwined

with the hedonic dimension, emphasising the importance of enjoyment when accessing audiovisual content, even in a foreign language (Pavesi, Ghia 2020, p. 99). This observation is especially pertinent in the context of the COVID-19 pandemic, which led to increased isolation among students and a shift in their media consumption habits towards social media, TV series, etc., consequently diminishing their attention levels.

Personal factors, including individual interests and preferences, exert a considerable influence on the frequency and intensity of their interactions with English-language media, serving also as pivotal catalysts for students' motivation. For instance, music is one of the factors that stimulate learners' interest toward the language, as stated by SAA0538:

(10) SAA0538_MEDIA_ACCESS_DYNAMICTY
SAA0538: *La musica mi ha aiutato molto in questo senso, mi ha fatto interessare tantissimo all'inglese [...] mi piaceva il ritmo, però non capivo cosa stessero dicendo' [...] 'per poi arrivare a una fase in cui vedevo che anche la traduzione magari proposta [...] era sicuramente una traduzione automatica e cercavo di diciamo capire quello che poteva essere realmente il senso di là del della musica [...] e quindi mi ci sono avvicinata in questo modo poi man mano le serie tv [...] Facebook e poi da lì Instagram tutto il resto YouTube tantissimo anche quindi diciamo siamo stati esposti molto di più rispetto alle generazioni precedenti all'inglese quindi credo che nasca da questo anche l'interesse.*
"Music helped me a lot in this sense, it got me very interested in English because I was there, maybe I liked the rhythm, but I didn't understand what they were saying. So, gradually, first there was the search for what they were saying, let's see what it means, then the translation, and then reaching a phase where I saw that even the translation proposed was definitely an automatic translation and I was trying to understand what the real meaning of the music was, and so I approached it in this way, then gradually the TV series."

Initially drawn to the rhythm and melody, the student progressed from receptive listening to actively seeking out song lyrics, realising the limitations of automated translations and striving to grasp the true meaning behind the music. Overall, the student credits the growing accessibility of English-language content, compared to past generations, for sparking her interest in the language.

This latter excerpt also highlights the interconnectedness of media consumption habits, where exposure to one form of media can serve as a gateway to exploring other types of content (audiovisuals and social networks). This trend is evident in the experiences of other students as well, such as SAA0556, who recounts how English-language Disney videotapes served as a significant catalyst during his childhood, stating: '*devo dire che in quel periodo forse un altro fattore molto importante sono state le videocassette che davano*

in lingua inglese della Disney. Mia madre me ne comprava diverse e ricordo distintamente che quelle furono forse il fire starter, se si può dire' ('I have to say that at that time perhaps another very important factor was the videotapes they were giving in English from Disney. My mother bought me several of them and I distinctly remember that those were perhaps the fire starter, as it were'). The term 'fire starter' suggests that these tapes were instrumental in sparking or initiating his interest and enthusiasm for learning English. He further emphasises the significance of another form of media in his life: video games. He reflects:

> (11) SAA0556_MEDIA_ACCESS_DYNAMICITY1
> SAA0556: *Io praticamente ho in qualche modo imparato a parlare in lingua inglese già dalle dalle scuole medie sostanzialmente non grazie allo studio diciamo scolastico ma grazie ai videogiochi perché mi sono avvicinato a due giochi nello specifico che erano Kingdom Hearts due e GTA San Andreas che mi ha dato due varianti completamente diverse dell'inglese e da cui ho preso particolare spunto senza neanche rendermi conto [...] Quindi l'interesse per l'inglese io credo che sia scaturito soprattutto da quel momento specifico da dal fatto che il gioco mi ha divertito mi ha emozionato [...] come già ho detto in altri momenti cioè mi sono sentito trasportato e questo trasporto continuo è un po' con me come se fosse stato un big bang e quindi continua quel moto di espansione fino alla fine [...] insomma fino all'infinito.*
> "I practically learned to speak English from middle school, essentially not thanks to school studies but thanks to video games because I got into two specific games, Kingdom Hearts II and GTA San Andreas, which gave me two completely different variations of English and from which I took particular inspiration without even realising it. So, my interest in English, I believe, mostly stemmed from that specific moment, from the fact that the game amused me, excited me... as I've already said on other occasions, I felt transported, and this continuous transport is somewhat with me as if it were a big bang, and so it continues that expansion motion until the end... in short, until infinity."

This excerpt highlights the transformative impact of video games on the student's language learning journey. Immersion in games exposed him to various forms of English, boosting both his proficiency and motivation. He describes a strong emotional connection to playing games in English, saying he feels 'transported', and likens his language learning motivation to a 'big bang' – a continuously expanding force. This reflects the powerful role of positive emotional experiences in enhancing language acquisition, reinforcing findings from previous studies on the influence of affective (see Sylvén, Sundqvist 2012).

Continuing with the same student's narrative, we may also notice another relevant factor, i.e. personal circumstances influence students'

extracurricular engagement with English media, as exemplified in the following excerpt:

(12) SAA0056_MEDIA_ACCESS_DYNAMICITY2

> I: *Secondo te le attività che svolgi fuori dal contesto istituzionale le svolgi in modo maggiore o minore rispetto ai tuoi colleghi coetanei a livello proprio di attività extra attività extra istituzionale extra accademica?*
>
> SAA0556: *Fino all'anno scorso sì, erano molto erano molto più presenti molto più influenti ma perché il mio tempo di studio era molto minore. Io ho avuto una vita accademica un po' particolare perché a causa di situazioni sono stato costretto a dover rinunciare agli studi. [...] Mi sono ritrovato con tantissimo tempo libero tra le mani e quindi mi sono ritrovato spesso e volentieri a dovermi riversare in nei videogiochi che sono sempre stati diciamo la costante nella mia vita. Sostanzialmente quindi se dovessi fare un paragone in quel periodo direi di sì era sicuramente molto più presente il videogioco oggi dato che comunque è una diciamo le lezioni sono diverse c'è molto più tempo in università da dover gestire allora è già più difficile però credo che resti comunque di più perché tendenzialmente io ritaglio quelle 2 ore alla giornata quell'oretta e mezza alla giornata per dedicarmi a giocare e quindi o comunque guardare i video i video che trattano di argomenti che mi interessano, che non sono solamente sui videogiochi ma anche su cartoni animati, su anime, cose che possono interessarmi in larga parte e quindi tendenzialmente sono è sempre uno spazio maggiore credo rispetto al resto.*
>
> "I: Do you think that the activities you engage in outside the institutional context are more or less than those of your peers in terms of extracurricular activities?
>
> SAA0556: Until last year, yes, they were much more present, much more influential, but because my study time was much shorter. I had a somewhat particular academic life because due to certain situations, I was forced to give up my studies. [...] I found myself with a lot of free time on my hands, so I often found myself having to immerse myself into video games, which have always been basically a constant in my life. So, if I had to make a comparison during that period, I would say yes, video games were definitely much more present. Today, since the classes are different and there is much more time to manage at university, it's already more difficult, but I still think it remains more because I generally carve out those 2 hours a day or that hour and a half a day to dedicate myself to playing, or watching videos on topics that interest me, which are not only about video games but also about cartoons, anime, things that can interest me to a large extent. So, generally, it's always a larger space, I believe, compared to the rest."

The student reflects on a shift in his activity levels, noting that his engagement in extracurricular activities was higher in the past due to having more free time during a break from his studies. During that period, he dedicated significant time to video games, a constant in his life. Despite changes in his academic

schedule, he still prioritises gaming and leisure activities, consistently setting time aside for them. This suggests that his engagement with English-language media is driven not only by external circumstances but also by personal interests and the positive emotions these activities provide.

Our data analysis reveals a dynamic shift in students' media access experiences, which indirectly influences their motivation. In response to the question about why she accesses English media, SAA0538 acknowledged that initially she felt compelled to engage with English-language media to improve her skills ('*Inizialmente me lo imponevo*'; 'Initially, I felt compelled to do so'). However, over time, this engagement became more spontaneous and natural ('*poi man mano iniziò a diventare sempre più spontanea come cosa*'; 'but over time it became increasingly spontaneous'). This shift highlights a transition from obligation-driven media use to a more intrinsic and voluntary involvement, reflecting an evolving relationship with English media.

(13) SAA0538_07_07_2023_MEDIA_ACCESS_DYNAMICITY2
I: *Lo fai te lo imponi perché devi avvicinarti all'inglese o lo fai proprio nel tempo libero in modo spontaneo?*
SAA0538: *Inizialmente me lo imponevo, poi man mano iniziò a diventare sempre più spontanea come cosa.*
"I: Do you do it because you need to get closer to English or do you do it spontaneously in your free time?
SAA0538: Initially, I felt compelled to do so, but over time it became increasingly spontaneous."

In conclusion, our examination of students' media access and practices reveals a complex interplay of personal, environmental, and experiential factors influencing language acquisition and motivation. Students' varied experiences with music, TV series, and video games underscore the significant role of media exposure. Notably, qualitative shifts in behavior, which cannot be captured by quantitative analysis, indicate that learning experiences can transform initial obligations into spontaneous habits of engaging with the English language outside the classroom. This suggests that exposure and learning experiences are critical variables in shaping students' motivational thinking. Understanding these dynamics is essential for developing tailored strategies that effectively support language learning and foster student engagement beyond the classroom, ultimately enhancing their interests.

4. General discussion and conclusions

Our qualitative case study explored the dynamic nature of motivational thinking. Using content analysis and two cycles of interview coding (structural

and pattern coding), we aimed to capture the nuanced and contextually embedded shifts in students' commitment to extramural English exposure.

The findings confirm the intricate, multifaceted influence of social and family relationships, internal factors, and learning experiences on extramural English exposure, while also highlighting emerging patterns. Interestingly, a pattern of *similarity* in students' experiences appeared during the transition from adolescence to adulthood, suggesting a release from external pressures typically perceived as imposed by parents and teachers.

Conversely, a noticeable pattern of *difference* was observed in how external factors, such as the influence of teachers, affected students' motivation in various ways. The data underscore the pivotal role of teachers in either fostering or diminishing motivation to learn a second language (*frequency pattern*). Effective teachers inspire students, but notably, not all encounters with pedagogically challenged teachers lead to demotivation; some learners actively seek alternative ways to enhance their language proficiency, such as non-formal instruction, which sometimes results in certification. This response was also identified as a *sequence pattern*. Essentially, a release from the perceived pressures imposed by parents and teachers marks a key transition in students' motivational thinking, as many began taking ownership of their language learning, aligning with the development of their Ideal L2 Self. This finding reinforces the dynamic and evolving nature of motivation, as described in Dörnyei's model.

A *causation pattern* also emerged, particularly when students discussed the role of media in motivating language learning. Video games, English-language videos, and other forms of media were frequently cited as catalysts for interest in the English language, aligning with previous research (Sylvén, Sundqvist 2012). These media sources evoked positive emotions, which is a component of motivation, and contributed to sustained motivation, illustrating the significant role of external motivators in shaping learning behaviors.

A *correspondence pattern* was observed during the COVID-19 lockdowns, when students increased their use of video-based content, as opposed to reading. This shift was interpreted by interviewees as a reaction to the emergence of the pandemic and subsequent restrictions, which decreased their level of attention.

The study highlights the necessity of seeing each learner as a unique and complex individual, shaped by a distinct combination of experiences, personal traits, and external influences. Rather than viewing learners as impersonal figures, the findings emphasise the importance of recognising the dynamic interplay of factors that influence motivation in individuals. This approach supports a more holistic and nuanced perspective on language learning motivation, consistent with Dörnyei's and Ushioda's frameworks.

Future research could focus on two promising goals to advance knowledge in this field. First, exploring whether students in language degree programs exhibit different motivational patterns compared to those in other disciplines could reveal variations in English learning motivation trajectories across academic settings. Second, examining the influence of parents and peers as role models – considering factors such as their English proficiency, lifestyle, and informal exposure to foreign languages – could provide deeper insights into the impact of family and social relationships on students' motivational thinking in language learning.

Bionotes:
Paola Leone is an Associate Professor of Language Teaching Methodology at the University of Salento in Italy. Her recent research interests include: 1) computer-mediated interaction for language learning; 2) interaction in multilingual contexts (intercomprehension); and 3) plurilingual education for young learners. She has produced numerous papers, both as a sole author and in collaboration with others, across various academic publishing venues. She has participated in studies funded by both national and international institutions and is currently involved in a nationally funded project on motivation and informal English language exposure.
Emanuela Paone is a research fellow at the University of Bologna. Previously, she held a non tenure-track position at the University of Salento, leading a project funded by the Puglia Region aimed at enhancing foreign language skills of professionals in the tourism sector. Her research interests include the acquisition and teaching of pragmatic and prosodic skills in Italian as a second language. Additionally, she is a co-author of an Italian textbook for foreigners and a volume on the development of emotional competence in a second language.

Authors' addresses: paola.leone@unisalento.it; emanuela.paone@unibo.it

*L*ingue e
*L*inguaggi

References

Al-Hoorie A.H and MacIntyre P.D. (Eds.) 2020, *Contemporary language motivation theory: 60 years since Gardner and Lambert* (1959), Multilingual Matters, Bristol.

Atkinson E.S. 2000, *An investigation into the relationship between teacher motivation and pupil motivation*, in "Educational Psychology" 20 [1], pp. 45-57.

Berndt T.J. and Keefe K. 1995, *Friends' influence on adolescents' adjustment to school*, in "Child Development" 66, pp.1312-1319.

Caruana S. 2021, *An Overview of Audiovisual Input as a Means for Foreign Language Acquisition in Different Contexts*, in "Language & Speech" 64 [4], pp. 1018-1036.

Charmaz K. 2001, *Grounded theory*, in Emerson R.M. (ed.), *Contemporary field research: Perspectives and formulations*, 2nd ed., Waveland Press, Prospect Heights, IL, pp. 335-52.

Dörnyei Z. 2005, *The psychology of the language learner*, Lawrence Erlbaum, Mahwah, New Jersey.

Dörnyei Z. 2009, *The Psychology of Second Language Acquisition*, Oxford University Press, Oxford.

Dörnyei Z. and Ushioda E. 2011, *Teaching and researching motivation*, 2nd ed., Longman/Pearson, London.

Dörnyei Z., MacIntyre P.D. and Henry A. (Eds.) 2015, *Motivational Dynamics in Language Learning*, Multilingual Matters, Bristol.

Eccles J.S., Wigfield A. and Schiefele A. 1998, *Motivation to succeed*, in Damon W. and Eisenberg N. (Eds.), *Handbook of Child Psychology*, 5th ed., Vol. 3: *Social, Emotional, and Personality Development*, John Wiley & Sons, New York, pp. 1017-95.

Gardner R.C. 1985, *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*, Edward Arnold, London.

Gardner R.C. and Lambert W.E. 1959, *Motivational variables in second language acquisition*, in "Canadian Journal of Psychology" 13, pp. 266-72.

Gardner R.C., Lambert W.E. 1972, *Attitudes and Motivation in Second Language Learning*, Newbury House, Rowley, MA.

Lewis M. and Sullivan M.W. 2007, *The development of self-conscious emotions*, in Elliot A.J. and Dweck C.S. (Eds.), *Handbook of Competence and Motivation*, The Guilford Press, New York/London, pp. 185-201.

McKay S.L., Wong S.L.C. 1996, *Multiple Discourses, Multiple Identities: Investment and Agency in Second-Language Learning among Chinese Adolescent Immigrant Students*, in "Harvard Educational Review" 66, pp. 577-606.

Neuendorf K.A. 2017, *The Content Analysis Guidebook*, SAGE Publications Inc, London.

Norton B. and McKinney C. 2011, *An identity approach to second language acquisition*, in Atkinson D. (Ed.), *Alternative approaches to second language acquisition*, Routledge, New York, pp. 73-94.

Pavesi M. ans Ghia E. 2020, *Informal Contact with English. A case study of Italian postgraduate students*, ETS, Pisa.

Pavesi M. this volume, *The informalisation of English language learning. A first national research project*.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English. An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Pomerantz E.M., Grolnick W.S. and Price C.E. 2007, *The role of parents in how children approach achievement*, in Elliot A.J. and Dweck C.S. (Eds.), *Handbook of Competence and Motivation*, The Guilford Press, New York/London, pp. 259-78.

Radel R., Sarrazin P., Legrain P. and Wild T.C. 2010, *Social contagion of motivation between teacher and student: Analyzing underlying processes*, in "Journal of Educational Psychology" 102 [3], pp. 577-587.

Roth G., Assor A., Kanat-Maymon Y. and Kaplan H. 2007, *Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning*, in "Journal of Educational Psychology" 99 [4], pp. 761-774.

Saldaña J. 2013, *The coding manual for qualitative researchers*, 2nd ed., SAGE, London.

Sordella S. 2016, *"Con parole mie": la lingua per lo studio in una classe multilingue*, in Andorno C. and Grassi R. (eds.), *Le dinamiche dell'interazione. Prospettive di analisi e contesti applicativi*, Officina 21, Milano, p. 109-120.

Spolsky B. 2000, *Language motivation revisited*, in "Applied Linguistics" 21 [2], pp. 157-169.

Sundqvist, P., and Sylvén, L. K. 2016, *Extramural English in Teaching and Learning: From Theory and Research to Practice*, Palgrave Macmillan UK, London.

Sylvén L.K., Sundqvist P. 2012, *Gaming as extramural English L2 learning and L2 proficiency among young learners*, in "ReCALL" 24 [3], pp. 302-321.

Ushioda E. 1998, *Effective motivational thinking: A cognitive theoretical approach to the study of language learning motivation*, in Soler E.A. and Espurz V.C. (Eds.), *Current Issues in English Language Methodology*, Universitat Jaume, Castelló de la Plana, pp. 77-89.

Ushioda E. 2001, *Language learning at university: Exploring the role of motivational thinking*, in Dörnyei Z. and Schmidt R. (Eds.), *Motivation and Second Language Acquisition*, University of Hawaii Press, Honolulu, pp. 93-125.

Ushioda E. 2009, *A person-in-context relational view of emergent motivation, self and identity,* in Dörnyei Z. and Ushioda E. (eds.), *Motivation, language identity and the L2 self,* Multilingual Matters, Bristol, pp. 215-228.

Ushioda E. 2011, *Language learning motivation, self and identity: current theoretical perspectives*, in "Computer Assisted Language Learning" 24 [3], pp. 199-210.

Ushioda E. 2020, *Researching L2 motivation: Re-evaluating the role of qualitative inquiry, or the 'wine and conversation' approach*, in Al-Hoorie A.H. and MacIntyre P.D. (eds.), *Contemporary Language Motivation Theory: 60 Years Since Gardner and Lambert (1959)*, Multilingual Matters, Bristol, pp. 194-211.

Ushioda E. and Dörnyei Z. 2017, Beyond *Global English: Motivation to Learn Languages in a Multicultural World: Introduction to the Special Issue*, in "The Modern Language Journal" 101 [3], pp. 451-454.

Yashima T., Nishida R. and Mizumoto A. 2017, *Influence of Learner Beliefs and Gender on the Motivating Power of L2 Selves*, in "The Modern Language Journal" 101 [4], pp. 691-711.

EXPLORING ITALIAN UNIVERSITY STUDENTS' SELF-DIRECTED EXPOSURE TO ENGLISH INPUT IN INFORMAL SETTINGS
Insights into motivation and agency

CRISTINA MARIOTTI
UNIVERSITÀ DI PAVIA

Abstract – Exposure to comprehensible input is essential for second language (L2) acquisition. In our globally interconnected era, technological advancements offer ample opportunities for accessing English. Understanding how Italian university students navigate this landscape is important, particularly considering the challenges they encounter in achieving English proficiency. In this study, we explored the reasons for university students' self-directed exposure to English input during their leisure time. Using the self-determination theory (SDT) framework, we conducted and analysed 21 semi-structured interviews with students at a medium-sized university in northern Italy, and conducted a thematic analysis to classify their motivations as intrinsic or extrinsic. A custom scale was also developed to assess the elaborateness of their expressions of agency. Our analysis revealed the presence of both intrinsic (e.g. immersion, entertainment, and personal interests) and extrinsic (e.g. improving English proficiency) motivations, together with elaborate expressions of agency being linked to specific behaviours, particularly when accessing English to enhance language skills. The results suggest that university students can capitalise on the abundant input that is available via technological resources by leveraging their motivation and agency.

Keywords: Self-directed informal contact with English; L2 English input; intrinsic motivation; extrinsic motivation; agency.

1. Italian university students and the learning of English: A complex relationship

The English language is evolving continuously and has a unique position on the global stage, resonating differently with diverse communities and exerting varying degrees of influence on language policies and education (Aronin *et al.* 2013). It functions as a lingua franca across multiple domains, including academia, business, technology, and diplomacy (Seidlhofer 2018). Proficiency in English has significant implications for individual mobility, economic opportunities, and cultural exchanges, representing both symbolic and tangible investments of time and energy (Bourdieu 1991; De Swaan 2001; Doiz *et al.* 2014; Norton 2016).

However, Italian students have reported difficulties in learning English, as shown by surveys conducted amongst university students. Landolfi (2012) created a corpus of written texts from over 500 first-year university students majoring in English as a second language (L2) in Naples. These students expressed their beliefs and attitudes regarding their experiences with English. A qualitative analysis of the corpus revealed an overly optimistic perception of their proficiency and an inaccurate awareness of their weaknesses in areas such as syntax, sentence construction, and paragraph development. Costa and Mariotti's (2020) survey explored undergraduate students' perspectives regarding English-taught courses at the University of Pavia, a medium-sized northern Italian university, and compared the views of local and international students across all disciplinary areas. This survey involved 357 participants and revealed notable differences between the two groups regarding their attitudes towards learning English in formal settings. While 78% of the international students perceived learning English positively, only 52% of the local students shared this sentiment. Furthermore, the international students exhibited higher levels confidence across all their English skills compared to their local counterparts. For example, in informal settings, 93% of the international students felt confident speaking and understanding English compared to 64% of the local students. Disparities were also evident in academic contexts. These findings revealed significant differences in the perceived English proficiency and confidence of local and international students in both formal and informal settings.

More recent data, based on the assessment of actual performances, have shown that Italy still lags behind in the European context, although it is slowly improving. According to the EF 2023 English Proficiency Index (EF EPI[1] 2023), an online, adaptive test of reading and listening skills comprising the results of 2.2 million adults, Italy is reported as being in 25th position out of 34 European countries, which suggests that competence in English is still somewhat modest amongst Italian speakers. In addition to the findings of the EF EPI, results obtained by Italian students in their final year of upper secondary school in the 2024 Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI - National Institute for the Evaluation of the Education and Training System) tests[2] further highlighted the challenges in English proficiency in Italy. These tests, which are conducted annually for all Italian students, assess fundamental skills in Italian, mathematics, and English. The results indicated that 60.3% of the students achieved level B2 (CEFR) in the reading test, an increase of

---

[1] EF EPI 2023, EF English Proficiency Index, https://www.ef.com/epi (4.4.2024)., https://www.ef.com.tr/epi/regions/europe/ (4.4.2024).
[2] Rapporto INVALSI 2023 https://www.invalsi.it/invalsi/index.php (4.4.2024).

4.2 percentage points compared to 2023, and 45.4% in the listening test, an increase of by 3.7 percentage points from the previous year. Despite these improvements, the majority of the students were still around the 50% mark, indicating substantial room for improvement.

The increased availability of digital technologies has expanded opportunities for accessing authentic, multimodal materials in English and, potentially, for learning the language. Contemporary English learners are increasingly engaging with informal English across diverse digital platforms, including subscription video-on-demand (SVOD) services, social media, online communities, and Massive Multiplayer Online Role-Playing Games (MMORPGs) (Dressman, Sadler 2020; Pavesi, Ghia 2020; Sockett 2014; Sundqvist, Sylvén 2014). Considering that exposure to comprehensible input is essential for second language acquisition (SLA) (Krashen 1985), the fact that there is an abundance of media available in English, both in the form of telecinematic and Internet-based affordances (Dressman, Sadler 2020; Pavesi, Ghia 2020; Reinhardt 2022), presents promising opportunities. SLA research has indicated that engagement in these settings correlates with positive effects on various aspects of L2 acquisition, including enhancements in L2 vocabulary (Sundqvist 2019), L2 speaking skills (Lee, Dressman 2018), scores on standardised English tests (Lee, Dressman 2018), and school grades (Lai *et al.* 2015). In addition, these activities are linked to psychological factors associated with SLA, such as motivation (Lee, Drajati 2019), confidence (Lai *et al.* 2015; Lee 2019; Lee, Drajati 2019), and L2 enjoyment (Talebzadeh *et al.* 2020).

However, the availability of abundant English input does not inherently guarantee that young people will engage with these opportunities. In an information society, this access has significant potential power in determining the quantity and quality of the knowledge one seeks to acquire. Research on motivation and agency can provide valuable insights into university students' behaviour regarding informal exposure to language input, offering potentially useful information to stakeholders who are involved in the process of learning English. Within the wider framework of informal contact with English, in this article, we focus on how Italian university students enrolled at a medium-sized university in northern Italy described informal, self-directed exposure to L2 English by adopting the theoretical framework of self-determination theory (SDT), a psychological perspective that investigates motivation in human behaviour. Before delving into the study and its outcomes, we will provide a concise review of motivational theories in SLA research (Section 2) and in psychology (Section 3). Subsequently, we will explore the relationship between motivation and agency (Section 4) and the relevance of these constructs for research on informal contact with English (Section 5). After outlining the study's rationale (Section 6) and describing our research design (Section 7), we

will then present and discuss our findings (Sections 8 and 9) and outline their implications for future research (Section 10).

2. Motivation in SLA

Affective factors in SLA have been explored since the 1970s (Chastain 1975; Dewaele, MacIntyre 2014; MacIntyre *et al.* 2016). Drawing on the concept of the affective filter proposed by Dulay and Burt (1977), Krashen (1982, 1985) suggested that emotions could either hinder or facilitate the processing of L2 input. For example, learners who experienced enjoyment while studying English would have a lower affective filter, thus promoting language acquisition. Expanding on Krashen's (1985) affective filter hypothesis, research has examined the impact of both negative emotions, such as anxiety (Gkonou *et al.* 2017), and positive emotions, like foreign language enjoyment (FLE), on L2 learning in classroom settings (Dewaele, MacIntyre 2014; Dewaele *et al.* 2018).

Amongst the affective factors identified as influencing the L2 learning process, motivation is acknowledged as a crucial factor (Gardner 1985; Dörnyei 2005). For example, Dörnyei (2020) maintained that, without a sufficient level of motivation, even L2 learners with the highest levels of language aptitude may not complete language-learning tasks or achieve long-term learning goals successfully.

Ushioda and Dörnyei (2012) delineated four distinct phases in the evolution of L2 motivation research in SLA, each characterised by different foci and methodologies. The first phase, called the social-psychological period (from 1959 to 1990), was marked by the foundational work of Gardner and Lambert (1972) and Gardner (1985) in Canada, who explored attitudes to language learning and motivation, and highlighted the importance of integrative and instrumental orientations in L2 acquisition. In particular, Gardner (1985) proposed that integrative and instrumental motivations played key roles in language learning. Integrative motivation refers to language learners' desire to integrate into the target language community, while instrumental motivation focuses on the achievement of practical goals. It is important to highlight that integrativeness has been subjected to significant re-evaluations in recent decades (Coetzee-Van Rooy 2006; Ushioda 2009) as a result of the evolution of English and its current hyper-central status (De Swaan 2001). Accordingly, English is no longer solely linked to specific geographical or cultural Anglophone communities, but is increasingly becoming associated with globalised cultures (Aiello 2018).

The second phase, the cognitive-situated period, began in the 1990s and focused on the role played by the integration of cognitive theories into research on L2 motivation. Scholars aimed to understand motivation within

specific learning contexts, such as classrooms, and explored concepts like intrinsic motivation and self-efficacy. The studies that were conducted during this period showed that motivation could influence learners' oral and written communication skills, cognitive processes, and their ability to self-regulate their learning (Dörnyei 2001).

The process-oriented period (the third phase) in motivation studies, which began at the turn of the century, emphasised motivational change over time. The subsequent socio-dynamic phase (the fourth phase) represented a shift towards understanding L2 motivation as a complex, dynamic system influenced by internal, social, and contextual factors by drawing on complexity theory, which posits that the learner is a complex adaptive system (CAS); in other words, a dynamic, multicomponent system that is composed of different subsystems that interact with each other in a non-linear way (Larsen-Freeman 1997, 2012). The concepts of self-organisation and emergent behaviour, which refer to unpredictable learner behaviour that cannot be deduced from the simple sum of the elements that constitute the system (Dörnyei *et al.* 2015), are typical of complex systems. Using this theoretical framework, scholars have explored the interaction between individuals and their learning environments by considering factors such as motivation and identity (Dörnyei 2005, 2009). Dörnyei's L2 Motivational Self System (L2MSS) integrates motivational concepts with theories pertaining to the self and identity, highlighting the importance of future self-images in motivating language learning. In a recent strand of research, the L2MSS has been applied to informal learning settings, such as self-directed language-learning activities performed in out-of-class (Lin 2023) and informal digital learning of English (IDLE) settings (Lee, Lee 2021).

This discussion highlights a common trend in SLA research: Scholars often investigate motivation by focusing on the participants' ability to reflect on their own L2 learning. However, we explored motivation in self-directed, informal exposure to English from a broader perspective without assuming an inherent intention to learn. The next section presents insights from the psychology of motivation to enable a more comprehensive analysis of the reasons that guide university students to seek informal contact with English.

3. Motivation in psychology

The study of motivation investigates the reasons underlying human behaviour by exploring an organism's needs and their connection to actions. Historically, these theories ranged from mechanistic to organismic views. Mechanistic views saw humans as reacting to physiological drives and external stimuli, while organismic views saw humans as active agents with

intrinsic needs who were capable of initiating behaviour (Deci, Ryan 1985). Early in the history of formal psychology, James (1890) emphasised the importance of concepts such as will, while Woodworth (1918) directly addressed the issue of motivated behaviour. Most psychologists then downplayed or ignored the importance of motivation for decades. Theories of motivation gained prominence starting with psychoanalytic psychology (Freud 1915); the initial focus was on drives, investigated either in the sphere of psychopathology or in behavioural psychology, in which behaviour organisation was attributed exclusively to associative bonds between stimuli and responses (Hull 1943).

In the 1950s, researchers realised that these views were ineffective in terms of accounting for pathology-free developmental patterns and non-stimulus-dependent behaviours, such as the need for exploration (Harlow 1950) and intrinsic motivation (White 1959). In line with this paradigm shift, psychology began to address volitional or non-drive-related aspects of motivation, leading to the development of several theories, each considering volition from a different perspective; examples include the hierarchy of needs (Maslow 1943), expectancy value theory (Vroom 1964), cognitive dissonance (Festinger 1957), and self-regulation (Carver, Scheier 1998).

In recent years, the dual-process theory (Kahneman 2011) has also made significant contributions to research; this theory posits that human behaviour is governed by two distinct systems, namely a reflective, deliberative system and an automatic, impulsive system, which interact and influence decision making, while the neurobiological models (Salamone, Correa 2012) have integrated insights from neuroscience and psychology to contribute to understanding the neural mechanisms underlying motivational processes. Of these, SDT is one of the most influential models in contemporary behavioural science. Developed in the 1970s by Edward L. Deci and Richard M. Ryan, SDT emphasises intrinsic motivation in human behaviour and explores how individuals' perceptions of social contexts influence motivation and well-being.

3.1 Self-Determination Theory (SDT)

SDT considers the study of motivation to be the exploration of two key principles, namely energisation and the direction of human behaviour. Energisation refers to the activation or arousal of motivation within an individual (Ryan, Deci 2000, p. 74); this concept encompasses the initiation of behaviour and the amount of effort an individual is willing to exert in relation to a particular goal or task. Accordingly,

> energy in motivation theory is fundamentally a matter of needs. An adequate
> theory of motivation must therefore take into account both the needs that are

*L*ᵍingue e
*L*ᵍinguaggi

innate to the organism […] and those that are acquired through interactions
with the environment.

Direction concerns "the processes and structures of the organism that give
meaning to internal and external stimuli, thereby directing action towards the
satisfaction of needs" (Deci, Ryan 1985, p. 3). SDT posits that individuals are
active organisms that are naturally inclined to grow, develop, and act on the
drives and emotions in their internal environment and on their external
environment, which has positive and negative forces. Individuals seek to
integrate these influences into their internal structure, or self, to function
effectively and to satisfy their full range of needs (Deci, Ryan 1985, p. 8). In this
regard, Ryan and Deci (2017) identified innate psychological needs for
autonomy, competence, and relatedness, which play crucial roles in fostering an
individual's motivation and well-being. Firstly, the need for autonomy refers to
the need to experience self-direction and volition in one's actions; individuals
with a strong sense of autonomy feel that their actions are self-endorsed and are
in harmony with their personal values and interests. Secondly, the need for
competence refers to the need to feel effective and capable in one's interactions
with the environment. Thirdly, the need for relatedness is the need to feel
connected to others and to experience a sense of belonging and intimacy in
interpersonal relationships. According to Ryan and Deci (2017), intrinsic
motivation arises from the satisfaction of these three basic psychological needs,
as opposed to extrinsic motivation that is driven by external rewards or pressures
rather than by personal satisfaction and fulfilment.

   To account for these phenomena, Deci and Ryan (1985, p. 43)
developed the cognitive evaluation theory (CET), which asserts that intrinsic
motivation increases when the environment supports an individual's needs
for autonomy and competence. Conversely, intrinsic motivation declines
when autonomy is disregarded or hindered by controlling measures (such as
bribes, demands, or pressuring language) or when perceived competence is
undermined by negative or unhelpful feedback. For example, early research
demonstrated that external motivators such as money could decrease intrinsic
motivation by undermining perceived autonomy (Deci 1971). Subsequent
studies showed that other external factors that were perceived as being
controlling, such as deadlines (Amabile *et al.* 1976) and surveillance (Plant,
Ryan 1985), also reduced intrinsic motivation. Similarly, interpersonal
environments can impact on intrinsic motivation depending on whether they
are viewed as informative or controlling. For example, while positive
feedback is generally seen as informative (supporting competence), it can be
perceived as controlling (undermining autonomy) if given in a pressuring
manner (Ryan 1982). Furthermore, internal factors, such as one's own
perceptions, emotions, and thoughts, can also influence motivation. This may
occur when feelings of self-value or identity become intertwined with

performance to the extent that engaging in a specific behaviour becomes essential in order to feel worthy or esteemed (Mageau *et al.* 2009; Plant, Ryan 1985).

In educational settings, the application of the SDT framework has emphasised the importance of autonomy, competence, and relatedness to increase students' motivation and academic success (Niemiec, Ryan 2009). It has been observed that, even if extrinsic rewards have a place in these settings, they are not always effective for motivating individuals in tasks that require creativity and problem solving (Pink 2009). SDT has also been used to explore SLA in both formal and informal contexts (Lamb *et al.* 2019; Noels *et al.* 2000), emphasising the importance of autonomy, competence, and relatedness for L2 learning. A psychometric instrument, called the Language Learning Orientation Scale (LLOS) (Noels *et al.* 2000), has been created to assess motivational orientations in L2 learning. Its applications consistently show that greater perceived autonomy, competence, and relatedness are positively associated with intrinsic motivation and internalised forms of extrinsic motivation (Agawa, Takeuchi 2016; Noels 2001).

4. The relationship between motivation and agency

In social cognitive psychology, the human mind is seen as being "generative, creative, proactive, and reflective, not just reactive", and "to be an agent is to intentionally make things happen by one's actions" (Bandura 2001, p. 4). Agency is particularly important when studying motivation because it highlights an individual's ability to act independently and to make decisions that influence their own life and the world around them. Agency involves a sense of control, self-efficacy, and the ability to make intentional decisions (Bandura 2001).

Agency and motivation are closely related concepts. From a complexity theory perspective, motivation can be viewed as a precursor to the expression of agency as emergent behaviour (cf. "motivation can be regarded as a precursor to engagement" Arndt 2023, p. 4). The concept of agency has attracted significant attention in SLA, often combined with discussions about L2 speakers' identity (Mercer 2012; Zuengler 1989) and learner engagement in formal settings (Mercer 2012; Reeve 2012). This perspective suggests that learners are not passive recipients in the language-learning process; instead, they have the capacity to make conscious choices, to exert influence, to resist (such as by remaining silent or discontinuing courses) or to comply, despite potential limitations imposed by their social circumstances.

5. Motivation and agency in informal contact with English

There has recently been a significant increase in empirical research on informal second language practices (ISLPs). However, limited systematic attention has been paid to the affective factors influencing individuals' access to English outside of the classroom. Studies addressing university students' motivation to access English input informally include those by Trinder (2017), Kusyk (2020) and, concerning Italy, Pavesi and Ghia (2020), and De Riso (2023).

In a study focusing on drivers for ISLPs, Trinder (2017) examined a group of Austrian economics students' access to English-language media and technologies. Using open-ended questions, the researcher investigated users' reasons for accessing various types of media and observed a phenomenon termed "dual-purpose engagement", whereby participants accessed English media not only to enhance their L2 skills but also for leisure and entertainment, making language learning an enjoyable and effortless activity. Similarly, other studies exploring motivations for ISLPs have found a combination of factors, particularly a balance between L2 learning objectives and recreational pursuits.

Kusyk (2020) conducted research on the online informal learning of English (OILE) amongst French and German university students, and revealed that their exposure habits were guided by a combination of explicit language-learning goals and entertainment purposes. While the participants mainly engaged with English media for recreation and language learning, their focus shifted depending on the type of media, such as watching AVs and playing video games for leisure, and reading online texts for learning purposes.

In a comprehensive survey of postgraduate Italian students, Pavesi and Ghia (2020) found that language-learning and hedonic factors were the main reasons for accessing English input informally. The participants often sought authenticity and naturalness by watching films, television series, and other AV programmes in the original language, and expressed admiration for the language itself.

De Riso's (2023) subsequent research on Italian university students corroborated these findings, with the participants engaging in leisure activities in English to experience authenticity, to pursue their interest in the language, and to seize opportunities to enhance their L2 skills.

Current research is providing more information about motivation in informal contact with English (Ghia this volume; Leone, Paone this volume) within the wider framework of the extensive PRIN-funded study of "The informalisation of English language learning through the media" (Pavesi this volume). To date, we are not aware of any research that has employed an SDT perspective to examine motivation in self-directed access to English input in informal settings. Building on the aforementioned considerations, we

investigated the potential implications of this theoretical framework for investigating the motivated behaviour of Italian university students in contexts of informal access to English as an L2.

6. Rationale for investigating motivation and agency in self-directed informal settings

It is important to clarify that our research concerns the self-directed *access* to English input rather than the self-directed *learning* of the English language; the latter, rooted in learner autonomy, emphasises students' control over their learning journey (Holec 1981; Reinders 2010).

In our approach, we focused on exploring motivation and agency in Italian university students' decisions to access English input without formulating assumptions about the reasons that guided their behaviour. This approach enabled the participants' experiences and motivations to emerge naturally within the data. Based on the view that "the active organism views stimuli not as causes of behaviour, but as affordances or opportunities that the organism can utilise in satisfying their needs" (Deci, Ryan 1985, p. 4), we adopted the SDT framework due to its comprehensive nature and adaptability to various contexts, including informal contact with English settings. With this context in mind, our research questions (RQs) focused on self-reported descriptions of motivated and agentic behaviour, as follows:

- RQ1: What drives our informants to access English input during their free time?

- RQ2: How do they elaborate on the criteria that guide their behaviour?

We aimed to answer these questions through a qualitative analysis of our informants' narratives.

7. Methodology

7.1 Research design

Our study is qualitative in that it investigates data collected through semi-structured interviews with Italian university students from an emic perspective. Qualitative research is well suited for exploring multifaceted phenomena such as motivational thinking and expressions of agency because it provides in-depth insights into the participants' perspectives (Duff 2012). Semi-structured interviews serve as a flexible tool that enables researchers to elicit rich, context-dependent information while maintaining the possibility of

cross-case comparisons (Smith 2018). Emic data were corroborated by information deriving from the Informal English Contact and Learning (IECoL) questionnaire, which was devised as part of the on-going, large-scale national project (PRIN) "The informalisation of English language learning through media: Language input, learning outcomes and sociolinguistic attitudes from an Italian perspective" (Pavesi this volume; Pavesi *et al.* 2023). The questionnaire data concerned the type of university course, the frequency and intensity of exposure, the attitude towards English on a 0-10 scale, and the participants' levels of English proficiency, which was evaluated using CEFR language certificates or, if these were unavailable, via a CEFR-based test (cf. Pavesi *et al.* 2023 for a detailed description of the questionnaire).

7.2 Participants

The participants were 21 undergraduate and postgraduate students, 11 males and 14 females, whose native language was Italian. They were enrolled at the University of Pavia, a medium-sized university in northern Italy, and were taking courses that were distributed equally across three macro areas: Macro area A, Arts and Humanities (including students in the Arts, History, Foreign Languages, and Psychology), Macro area B, Economics/Law/Political and Social Science, and Macro area C, Science and Technology (Pavesi *et al.* 2023).

The informants were selected using a combination of criterion and convenience sampling. Criterion sampling was employed to identify potential participants amongst students who accessed English input during their leisure time based on their responses to the IECoL questionnaire. Convenience sampling involved recruiting participants through invitation. Students who specialised in English were excluded from the sample to mitigate potential bias. Students who chose to participate voluntarily in the semi-structured interviews regarding their exposure behaviour were included. While convenience sampling offered practicality and accessibility in the data collection, we acknowledge the potential biases that can be introduced by the self-selection process. This was indicated by the underrepresentation of low-exposure informants in our sample. The majority of the volunteers (18 out of 21) had very high exposure to English input (several hours every day) and a very positive attitude towards English (rated from 8 to 10 out of 10), with varied proficiency levels (B1 = 5; B2 = 3; C1 = 9; C2 = 4). Only three of the 21 informants had low exposure. These participants also had a B1 proficiency level and negative attitudes towards English (rated from 3 to 5 out of 10).

7.3 Instrument

The data analysed in this article were collected through responses to the following two open-ended questions:

1) *Quali sono i motivi per cui ti esponi all'inglese nel tempo libero?* (What are the reasons for exposing yourself to English in your free time?)

2) *Quali sono i criteri che guidano il tuo comportamento?* (What are the criteria that guide your behaviour?)

The researcher explicitly considered the formulation of questions and the social context during the recruitment and data collection phases due to recognising their influence on the specific research circumstances. The questions were designed to be open-ended to gather comprehensive data, and were formulated using terms from the semantic field of decision and choice to capture the concepts of motivation and agency, consistent with SDT principles (Ryan, Deci 2017). The wording was concise to clarify the desired information while minimising priming. To create a comfortable environment and to facilitate effective communication, the researcher adopted a co-adaptive approach, actively engaging with each participant to build a feedback loop during the interview process. The interviews were conducted in Italian, the participants' native language, to enhance the clarity and ease of communication.

7.4 Data collection procedure

The interviews were conducted in a quiet, comfortable environment to encourage open communication. The participants received an informed consent form one week prior to the interviews, which provided detailed information about the project and allowed the students ample time to review the information and to decide whether to participate. Participation was incentivised via a small monetary compensation. The interviewer collected signed informed consent forms before each interview began. The interviews lasted for 15 minutes on average and included questions that explored additional aspects of the informants' experiences, although these are not the focus of the present research. Audio recordings of the interviews were made using a portable device.

7.5 Data analysis

The interviews were transcribed using Speechmatics. A thematic analysis approach was employed to analyse the data, adopting a lexical-semantic principle (Braun, Clarke 2022). This involved searching for expressions that revealed the reasons that motivated our informants to access English input

and expressions of agency, operationalised as the expression of one's ability "to intentionally make things happen by one's actions" (Bandura 2001, p. 4). Specifically, to detect expressions of agency, we examined the types and moods of verbs and modal auxiliaries, adverbs, and adjectives following Duff's (2012) suggestions for research on self-reported agency in qualitative studies. In addition, we examined personal pronouns, which can serve as intensity modulators in Italian (Bazzanella, Gili Fivela 2009).

The thematic analysis was conducted as an inductive, iterative process that involved reading and re-reading the data for initial and axial coding (Saldaña 2021). Codes were assigned to expressions describing motivation by drawing on categories adapted from Deci and Ryan (2008)* and Noels *et al.* (2019), as reported in Table 1.

Motivation type and code	Description
Intrinsic motivation (INT)	Individuals are internally driven to access input in English in informal settings due to the inherent enjoyment and satisfaction they derive from the process.
Extrinsic motivation (EXT) Subtypes: 1) Integrated 2) Identified 3) Introjected 4) External	Individuals are driven to access input in English in informal settings to satisfy needs that are not connected to the inherent enjoyment and satisfaction they derive from the process. Extrinsic motivation encompasses the following subtypes, ranging from most to least internalised: 1. *Integrated*: Informants recognise accessing input in English as an integral part of their identity and self-expression; they find personal meaning in doing it and see it as consistent with other aspects of their identity. 2. *Identified*: Informants recognise the personal significance of accessing input in English in achieving meaningful goals, such as study or career advancement. 3. *Introjected*: Informants feel internally obligated to access input in English due to self-imposed or normative pressures. They may focus on avoiding negative emotions or boosting their ego rather than freely choosing the activity for its personal meaning. 4. *External*: Informants feel compelled to access input in English due to external demands or circumstances.

Table 1
Description of codes for motivation.

Codes were assigned to expressions describing agency using the Agency Elaborateness Scale (AES), a scale that was created specifically for this study following Duff's (2012) suggestion to develop ad hoc instruments for detecting and measuring agency in qualitative research based on the adopted theoretical framework. Although other agency scales exist, they are designed to measure different phenomena (such as perception) and do not allow for qualitative analyses (cf. the Sense of Agency Scale in Tapal *et al.* 2017).

The AES indicators included endorsement of one's behaviour, proactivity, and expressions of confidence. As shown in Table 2, these indicators were inspired by SDT principles (Ryan, Deci 2017) and were adapted to suit the context of the present study.

Indicator	Code	Description
Endorsement	END	Endorsement of one's behaviour refers to an individual's identification with, support for, and acceptance of their own actions while accessing input in English. This concept is linked to the internalisation process, in which individuals view their behaviours as aligned with their own values, beliefs, and/or goals.
Proactivity	PRO	Proactivity refers to the ability to identify and pursue opportunities to access input in English, rather than waiting for instructions or external cues; it indicates self-direction.
Confidence	CONF	Expressions of confidence relate to individuals' beliefs in their ability to effectively pursue and achieve their goals as they access input in English.

Table 2
Description of indicators in the AES.

The elaborateness score ranged from 0 (indicating the absence of any indicators) to 3 (indicating the presence of all three indicators). The codes were inserted manually and retrieved automatically using data analysis software to ensure accuracy. The data were then reviewed by two peers who were not part of this study, but who were knowledgeable about the topic. The inter-rater reliability was 80%; the raters then discussed the inconsistencies and reached agreement about the remaining 20%. We present our results in Section 8 by highlighting recurring patterns and providing descriptions of some of the informants' behaviours.

7.6 Ethical aspects

To ensure data anonymity during the interviews, voice recordings were encrypted, and the files were stored on a device that was not connected to the Internet.

8. Results

To address RQ1 (What drives our informants to access English input during their free time?), we classified their responses by applying the SDT categories of motivation.

Our informants accessed English input for the following reasons, listed in decreasing order of frequency: seeking entertainment, enjoying immersive experiences, improving their English language skills, accessing a wide range of content and information that aligned with their personal interests, and because they perceived English-language information - particularly news - as being more reliable than its Italian counterpart.

By applying the SDT paradigm, our data revealed that the intrinsically motivated behaviour encompassed seeking entertainment, enjoying immersive experiences, and pursuing personal interests. The informants who reported intrinsic motivation had a very high exposure index (at least three

hours every day) and extremely positive attitudes towards English (from 9 to 10), but not necessarily high proficiency. An example of how these informants described this type of motivation is provided in Excerpt 1:

> (1) *Ormai mi sono abituata a vederli direttamente in inglese e se uso i sottotitoli li metto direttamente in inglese, ma tendenzialmente lo faccio perché lo trovo, trovo più interessante approcciarmi al contenuto nella sua forma originale, quindi senza passare attraverso il doppiaggio, perché lo trovo più interessante al fine diciamo, di usufruire del contenuto audiovisivo nella sua forma integrale, più pura diciamo, per un'esperienza più piena.*
>
> "Now I'm used to watching them directly in English, and if I use subtitles, I use them in English as well. But I tend to do it because I find it more interesting to approach the content in its original form, so without going through dubbing, because I find it more interesting to enjoy the audiovisual content in its complete, purer form, for a fuller experience."

As can be seen, this informant's access to English input was mainly motivated by the intrinsic enjoyment and satisfaction experienced during the process. An example of seeking entertainment is provided in Excerpt 2:

> (2) *Utilizzo quasi esclusivamente materiale in inglese per, perché ritengo che i media e lingua inglese siano migliori, le serie TV sono proprio di qualità migliore.*
>
> "I almost exclusively use material in English because I believe that English-language media is superior, especially when it comes to TV series which are of higher quality."

With regard to extrinsic motivation, our informants mainly reported the identified extrinsic subtype, as they recognised the personal significance of accessing input in English to achieve meaningful goals such as improving their proficiency levels, as shown in Excerpt 3:

> (3) *Guardo principalmente magari vlog di persone che riprendono la propria vita e che magari utilizzano i termini che non ho mai sentito perché sono slang o comunque termini che vengono utilizzati magari in una zona o nell'altra, quindi principalmente per utilità ecco, l'inglese comunque come lingua mi è sempre piaciuto, quindi vabbè, quello l'ho dato forse per scontato, però lo ritengo anche molto utile comunque per tutto, e quindi, cioè io ritengo di avere comunque un buon livello, però posso sempre migliorarlo e questo modo per migliorarlo senza studiare attivamente.*
>
> "I mainly watch vlogs of people documenting their lives, and maybe they use terms I've never heard before because they're slang or terms used in one area or another. So mainly for utility, you know? Anyway, I've always liked English as a language, so that's a given, but I also find it very useful for everything. So, I mean, I think I have a good level of

proficiency already, but I can always improve, and this is a way to improve without actively studying."

This informant, an engineering student, described the behaviour he adopted in detail when accessing English input to achieve a purpose he deemed useful, likely for his future professional life, and stressed the fact that it was important for him to expose himself to online content that he found interesting as a means to "improve without actively studying".

Our informants also frequently mentioned the most internalised subtype, integrated extrinsic motivation, which corresponded to accessing input in English as an integral part of an individual's identity and self-expression (see Table 1 for a detailed description), as illustrated in Excerpt 4:

(4) *Quando leggo una notizia lo faccio prima in lingua inglese diciamo da quotidiani inglesi o comunque americani statunitensi invece poi la leggo in quelli italiani mi rendo conto anche di come entrambi hanno diciamo una narrativa diversa, e questa per me è una cosa importante, se noi leggiamo una notizia in inglese abbiamo una pluralità di punti di vista perché abbiamo quello americano quello europeo quello che ne so dell'emisfero australe per dire perché l'inglese comunque è una lingua che si parla in tutti questi in queste aree del mondo quindi effettivamente ci esponiamo contemporaneamente a tanti punti di vista diversi.*
"When I read a news article, I read it first in English from British or American newspapers, and then in Italian ones, I realise how both have different narratives. And that's an important thing to me. If we read news in English, we have a plurality of perspectives because we have the American one, the European one, and even perspectives from the southern hemisphere, because English is spoken in all these areas of the world. So, effectively, we are exposed simultaneously to many different viewpoints."

This informant, who was a political science student, emphasised the importance of accessing unbiased news from multiple sources in an attempt to uphold her commitment to understanding the contemporary world accurately.

Only less proficient informants described their behaviour using words that could be analysed as expressions of introjected regulation (see Table 1 for a detailed description), a more controlled subtype of extrinsic motivation. These informants apparently felt obligated to access input in English due to self-imposed pressures. An example is provided in Excerpt 5:

(5) *Guardo video di YouTube in inglese perché voglio cercare di andare un po' più avanti, però successivamente insomma, perché poi quando uno inizia a sentirsi in imbarazzo di fronte alla lingua, non è nemmeno più portato ad apprenderla, quindi prima di tutto ci si abitua a un livello base poi magari si sale insomma sì sì quindi crearsi una zona*

> *comunque di comfort e poi da lì eventualmente decidere come muoversi.*
> "I watch YouTube videos in English because I want to try to improve a bit more, but then, you know because when someone starts feeling embarrassed about the language, they're not even inclined to learn it anymore. So first of all, you get used to a basic level, then maybe you move up, you know. So, yes, yes, creating a comfort zone and then from there, eventually deciding how to proceed."

In this excerpt, the informant mentioned mainly accessing English input to avoid negative emotions related to limited proficiency, rather than choosing the activity due to its personal significance. This inclination may be attributed to the informant's personal characteristics or past internalised negative experiences with English, which may have caused feelings of embarrassment when using the language.

It was interesting that our data confirmed a trend that was observed in prior research in which telecinematic input served two primary functions: It was associated with intrinsic motivation due to the appreciation of the original language and with a quest for naturalness and authenticity (Pavesi, Ghia 2020, p. 95), and was also the most frequently used type of input for enhancing English proficiency, particularly for expanding L2 vocabulary (cf. dual-purpose engagement in Trinder 2017). An example is presented in Excerpt 6:

> (6) *Allora, come motivi principali, se un materiale è disponibile anche in inglese preferisco vederlo magari in lingua originale, appunto, anche perché lo vedo comunque come un modo per unire svago a un qualcosa di utile.*
> "So, as main reasons, if a material is available in English, I prefer to watch it in the original language, precisely because I see it as a way to combine leisure with something useful."

Consistent with the notion of self-directed exposure, none of our participants reported feeling compelled to access English input due to external demands or uncontrollable circumstances. As a result, no instances of external extrinsic motivation stemming from an obligation to engage with the L2 were identified.

Our data did not reveal identifiable patterns concerning gender or affiliation with a specific macro area.

We addressed RQ2 (How do they elaborate on the criteria that guide their behaviour?) by analysing expressions of agentic behaviour. The highest elaborateness score (AES = 3) was consistently associated with descriptions that reflected intrinsically motivated behaviour (for instance, when seeking entertainment). High scores (AES = 2 and 3) were also associated with

internalised, self-regulated extrinsic motivation of an identified subtype, such as when accessing English input to improve proficiency in the language or to seek reliable news. Informants who scored high for agency elaborateness typically demonstrated very high exposure and highly positive attitudes towards English (9 or 10).

Distinctive patterns of behaviour emerged when the informants reported accessing English input to learn the language. An example is presented in Excerpt 7:

> (7) *Dipende da cosa guardo, perché se è una serie, magari un film che in cui parlano molto velocemente o che non capisco bene decido in base al tempo che ho, magari metto i sottotitoli in italiano così intanto capisco la trama e in base quindi al livello di apprendimento poi cerco di mettere i sottotitoli in inglese, magari prima metto una serie, magari un film in inglese, poi se mi rendo conto che non riesco a capire determinate cose, allora passo ad altro così non sto a perdere tempo con qualcosa che per me è troppo difficile, ma magari no, se parlano più piano cerco magari di mettere di mettere subito i sottotitoli in inglese, magari piano piano.*
> "It depends on what I'm watching, because if it's a series, maybe a movie where they speak very quickly or that I don't understand well, I decide based on the time I have. Maybe I'll put on Italian subtitles so that I can understand the plot, and then, based on my learning level, I try to switch to English subtitles. Perhaps I'll start with an English series or movie, and if I realise I can't understand certain things, then I switch to something else so I don't waste time on something too difficult for me. But if they speak more slowly, maybe I'll try to put on English subtitles right away, perhaps gradually."

It could be suggested that this informant, a mathematics students, employed a form of self-assigned, self-directed viewing behaviour that reflected his needs, expressing endorsement of his behaviour ("maybe a movie where they speak very quickly or that I don't understand well, I decide based on the time I have"), proactivity ("if they speak more slowly, maybe I'll try to put on English subtitles right away, perhaps gradually"), and confidence in his self-efficacy ("based on my learning level, I try to switch to English subtitles"). This informant appeared to have developed a detailed procedure supported by reasoning and rationale tailored to his specific needs and preferences, indicating a thoughtful approach to his viewing habits.

Those who provided the least detailed descriptions of their agentic behaviour (AES = 1 or 0) tended to have lower proficiency and negative attitudes towards English (scoring from 3 to 5 out of 10).

9. Discussion

We examined how motivation and agency could impact on informal L2 English exposure and highlighted the nuanced nature of these affective factors. Firstly, our data confirmed that our informants accessed English input to fulfil various functions in their daily lives, leveraging the wide range of opportunities provided by the multimodality of the affordances (Hasebrink *et al.* 2007). This finding aligns with previous studies in the field (De Riso 2023; Kusyk 2020; Pavesi, Ghia 2020; Trinder 2017).

Secondly, we noted that intrinsic motivation, defined as the decision to access input in English based on the inherent enjoyment and satisfaction derived from the process, was associated with very high levels of exposure frequency and intensity, as well as highly positive attitudes towards the language. It is interesting that it was not necessarily linked to advanced proficiency, as it was also reported by informants with B1 and B2 proficiency levels. Given the nature of the data and the design of this study, it was not possible to determine whether increased exposure stemmed from a positive attitude towards the language or, conversely, if immersing oneself in English input provided a deeply ingrained and enjoyable activity that iteratively amplified the need for exposure (Pavesi, Ghia 2020). Internalised forms of extrinsic motivation also played an important role in guiding our informants' exposure behaviour. Specifically, the desire to enhance their English proficiency, driven by identified extrinsic motivation, significantly influenced L2 exposure practices.

Telecinematic input was associated with two types of motivation: It was linked to intrinsic motivation, as the informants accessed it due to appreciation of the original language and a desire for naturalness and authenticity (see Pavesi, Ghia 2020, p. 95). In addition, it was associated with identified extrinsic motivation, as it was the most frequently used type of input for self-assigned, self-directed viewing practices aimed at improving English skills (see dual engagement in Trinder 2017).

Finally, with regard to agency, the informants with low exposure, low proficiency, and negative attitudes towards English provided minimal details when describing their behaviour in informal contact with English. Those with very high exposure and highly positive attitudes towards English used elaborate expressions, particularly when discussing accessing English input to immerse themselves in the language, to improve their English proficiency, and to seek news online due to the perceived greater availability and reliability compared to news that was delivered in Italian. In these instances, our informants provided detailed accounts demonstrating endorsement (END) of their behaviour, proactivity in seeking input (PRO), and confidence (CONF) in their self-efficacy. Thus, it can be suggested that more exposure

and positive attitudes enhanced the informants' agency, leading to more meaningful engagements with English.

9.1 Limitations of the present study

While qualitative research can provide valuable insights, its findings may lack generalisability to broader populations. Moreover, convenience sampling can introduce a self-selection bias (Dörnyei, Dewaele 2023). It is also important to recognise that the data generated from interviews are social constructions, and that the content and depth of the responses may be influenced by various factors such as the setting, the interviewer, or the interviewee (e.g. personality traits, mood on the day of the interview, or the reason for participating in the interview). Of note, the observer interacts with the observed, and participants' presentations of themselves may be shaped by perceptions of the researcher's expectations (Duff 2012; Talmy 2010).

Furthermore, we acknowledge that motivation, like language development, is dynamic and subject to fluctuations (Larsen-Freeman 2012, p. 78), and is unique to each individual within different systems. Therefore, this study only provided a situated snapshot and aimed to lay the methodological groundwork for potential future expansion.

10. Conclusions and implications for future research

In this article, we explored motivation and agency in the self-directed exposure to English input in informal contexts amongst Italian university students. Despite its limitations, the findings of this study might contribute to a more nuanced understanding of how motivation and agency influence behaviour in informal contact with English. Moreover, they may have potential implications for various stakeholders. In particular, students could explore the reasons that guide them to access English input during their leisure time by identifying the reasons that are intrinsic and those that are extrinsic, yet which are internalised and aligned with their sense of self, making them easier to act upon. By leveraging their motivation, they can capitalise on the abundant input that is available, which is not only rich and varied but can also be made comprehensible and tailored to their needs through technological resources. This can lead to greater autonomy for L2 English learners and to a satisfying experience that can extend to other learning contexts. We believe that it would be valuable for future research to continue to investigate motivation in informal settings to understand how students utilise the resources at their disposal.

Bionote: Cristina Mariotti is Associate Professor in English Language in the Department of Political and Social Sciences at the Università di Pavia, Italy. She has done research in the cognitive processes involved in the content-based learning of English and has published CLIL teaching materials. She has then expanded her interests to ICLHE (Integrating Content and Language in Higher Education), EMI (English-medium Instruction), and the use of subtitled audiovisual materials in language teaching and learning.

Author's address: cristina.mariotti@unipv.it

References

Agawa T. and Takeuchi O. 2016, *Validating self-determination theory in the Japanese EFL context: Relationship between innate needs and motivation*, in "Asian EFL Journal" 18, pp. 7-33.

Aiello J. 2018, *Negotiating Englishes and English-Speaking Identities*, Routledge, London.

Amabile T.M., DeJong W. and Lepper M.R. 1976, *Effects of externally imposed deadlines on subsequent intrinsic motivation*, in "Journal of Personality and Social Psychology" 34 [1], p. 92-98.

Arndt H. 2023, *Construction and validation of a questionnaire to study engagement in informal second language learning*, in "Studies in Second Language Acquisition" 45 [5], pp. 1456-1480.

Aronin L., Fishman J.A., Singleton D. and Ó Laoire M. 2013, *Introduction. Current multilingualism: A new linguistic dispensation*, in Singleton D., Fishman J.A., Aronin L. and Ó Laoire M. (Eds.), *Current Multilingualism: A New Linguistic Dispensation*, Mouton De Gruyter, Berlin, pp. 3-24.

Bandura A. 2001, *Social cognitive theory: An agentic perspective*, in "Annual Revue of Psychology" 52, pp. 1-26.

Bazzanella C. and Gili Fivela B. 2009, *Introduzione*, in Gili Fivela B. and Bazzanella C. (Eds.), *Fenomeni di Intensità nell'Italiano Parlato*, Cesati, Firenze, pp. 13-24.

Bourdieu P. 1991, *Language and Symbolic Power: The Economy of Linguistic Exchanges*, Harvard University Press, Cambridge, MA.

Braun V. and Clarke V. 2022, *Thematic Analysis: A Practical Guide*, Sage, London.

Carver C.S. and Scheier M.F. 1998, *On the Self-regulation of Behaviour*, Cambridge University Press, New York.

Chastain K. 1975, *Affective and ability factors in second-language acquisition*, in "Language Learning" 25 [1], pp. 153-161.

Coetzee-Van Rooy S. 2006, *Integrativeness: Untenable for world Englishes learners?*, in "World Englishes" 25 [3/4], pp. 437-450.

Costa F. and Mariotti C. 2020, *EMI students' "international coexistence" at one Italian university*, in Bowles H. and Murphy A.C. (Eds.), *English-Medium Instruction and the Internationalization of Universities, International and Development Education*, Palgrave Macmillan, London, pp. 229-258.

De Riso C. 2023, *Out-of-the-classroom English Input and Informal Second Language Acquisition: A Longitudinal Study of Italian University Students*, unpublished doctoral thesis, University of Pavia.

De Swaan A. 2001, *Words of the World: The Global Language System*, Polity Press, Malden, Mass.

Deci E.L. 1971, *Effects of externally mediated rewards on intrinsic motivation*, in "Journal of Personality and Social Psychology" 18, pp. 105-115.

Deci E.L. and Ryan R.M. 1985, *Intrinsic Motivation and Self-determination in Human Behaviour*, Plenum Press, New York.

Deci E.L. and Ryan R.M. 2008, *Self-determination theory: A macrotheory of human motivation, development, and health*, in "Canadian Psychology/Psychologie Canadienne" 49 [3], pp. 182-185.

Dewaele J.-M. and MacIntyre P.D. 2014, *The two faces of Janus? Anxiety and enjoyment in the foreign language classroom*, in "Studies in Second Language Learning and Teaching" 4 [2], pp. 237-274.

Dewaele J.-M., Witney J., Saito K. and Dewaele L. 2018, *Foreign language enjoyment and anxiety: The effect of teacher and learner variables*, in "Language Teaching Research" 22 [6], pp. 676-697.

Doiz A., Lasagabaster D. and Sierra J.M. 2014, *Language friction and multilingual policies at higher education: The stakeholders' view*, in "Journal of Multilingual and Multicultural Development" 35, pp. 345-360.

Dörnyei Z. 2001, *New themes and approaches in second language motivation research*, in "Annual Review of Applied Linguistics" 21, pp. 43-59.

Dörnyei Z. 2005, *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*, Routledge, London/New York.

Dörnyei Z. 2009, *The L2 motivational self system*, in Dörnyei Z. and Ushioda E. (Eds.), *Motivation, Language Identity and the L2 Self*, Multilingual Matters, Bristol, pp. 9-42.

Dörnyei Z. 2020, *Innovations and Challenges in Language Learning Motivation*, Routledge, London/New York.

Dörnyei Z. and Dewaele J.-M. 2023, *Questionnaires in Second Language Research: Construction, Administration, and Processing*, Routledge, London/New York.

Dörnyei Z., MacIntyre P.D. and Henry A. 2015, *Introduction: Applying complex dynamic systems principles to empirical research on L2 motivation*, in Dörnyei Z., MacIntyre P.D. and Henry A. (Eds.), *Motivational Dynamics in Language Learning*, Multilingual Matters, Bristol, pp. 1-7.

Dressman M. and Sadler R.W. (Eds.) 2020, *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Duff P. 2012, *Identity, agency, and SLA*, in Gass S.M. and Mackey A. (Eds.), *Handbook of Second Language Acquisition*, Routledge, London/New York, pp. 410-426.

Dulay H. and Burt M. 1977, *Remarks on creativity in language acquisition*, in Burt M., Dulay H. and Finocchiaro M. (Eds.), *Viewpoints on English as a Second Language*, Regents, New York, pp. 95-126.

Festinger L. 1957, *A Theory of Cognitive Dissonance*, Stanford University Press, Stanford.

Freud S. 1915, *The Unconscious*, Standard Edition, Hogarth, London, pp. 166-204.

Gardner R.C. 1985, *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*, Edward Arnold, London.

Gardner R.C. and Lambert W.E. 1972, *Attitudes and Motivation in Second Language Learning*, Newbury House Publishers, Rowley, MA.

Ghia E. this volume, *What drives L2 viewers' informal access to English? A case study on audiovisual input*.

Gkonou C., Daubney M. and Dewaele J.-M. (Eds.) 2017, *New Insights into Language Anxiety: Theory, Research and Educational Implications*, Multilingual Matters, Bristol.

Harlow H.F. 1950, *Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys*, in "Journal of Comparative and Physiological Psychology" 43 [4], pp. 289-294.

Hasebrink U., Berns M. and de Bot K. 2007, *In the presence of English: A resume after step one of an international study*, in Berns M., de Bot K. and Hasebrink U. (Eds.), *In the Presence of English: Media and European Youth*, Springer, New York, pp. 111-119.

Holec H. 1981, *Autonomy and Foreign Language Learning*, Pergamon, Oxford.

Hull C.L. 1943, *Principles of Behavior: An Introduction to Behavior Theory*, Appleton-Century-Crofts, New York.

Kahneman D. 2011, *Thinking, Fast and Slow*, Straus and Giroux, Farrar.

Krashen S 1982, *Principles and Practice in Second Language Acquisition*, Pergamon Press, Oxford.

Krashen S. 1985, *The Input Hypothesis: Issues and Implications*, Longman, London.

Kusyk M. 2020, *Informal English learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 333-348.

James W. 1890, *The Principles of Psychology*, Vol. 1, Henry Holt and Company, New York.

Landolfi L. 2012, *P.Æ.C.E. An Italian-English Corpus Based on EFL Students*, Liguori Editore, Napoli.

Lai C., Zhu W. and Gong G. 2015, *Understanding the quality of out-of-class English learning*, in "TESOL Quarterly" 49 [2], pp. 278-308.

Larsen-Freeman D. 1997, *Chaos/complexity science and second language acquisition*, in "Applied Linguistics" 18, pp. 141-65.

Larsen-Freeman D. 2012, *Complexity theory*, in Gass S. and Mackey A. (Eds.), *Handbook of Second Language Acquisition*, Routledge, London, pp. 73-87.

Lee J.S. 2019, *Quantity and diversity of informal digital learning of English*, in "Language Learning & Technology" 23 [1], pp. 114-126.

Lee J.S. and Drajati N.A. 2019, *Affective variables and informal digital learning of English: Keys to willingness to communicate in a second language*, in "Australasian Journal of Educational Technology" 35 [5], pp. 168-182.

Lee J.S. and Dressman M. 2018, *When IDLE hands make an English workshop: Informal digital learning of English and language proficiency*, in "TESOL Quarterly" 52 [2], pp. 435-445.

Lee J.-S. and Lee K. 2021, *The role of informal digital learning of English and L2 motivational self system in foreign language enjoyment,* in "British Journal of Educational Technology" 52 [1], pp. 358-373.

Leone P. and Paone E. this volume, *Researching motivation for informal English exposure and learning. A case study*.

Lin J. 2023, *The structural relationships among L2 motivation, out-of-class informal learning, and oral proficiency: A multiple-group structural equation modeling study*, in "Language Teaching Research" 0 [0], pp. 1-30. https://journals.sagepub.com/doi/10.1177/13621688231189030 (12.12.2024).

MacIntyre P., Gregersen T. and Mercer S. (Eds.) 2016, *Positive Psychology in SLA*, Multilingual Matters, Bristol.

Mageau G.A., Vallerand R.J., Charest J., Salvy S.J., Lacaille N., Bouffard T. and Koestner R. 2009, *On the development of harmonious and obsessive passion: The role of autonomy support, activity specialization, and identification with the activity*, in "Journal of Personality" 77 [3], pp. 601-646.

Maslow A.H. 1943, *A theory of human motivation*, in "Psychological Review" 50 [4], pp. 370-396.

Mercer S. 2012, *The complexity of learner agency*, in "Apples – Journal of Applied Language Studies" 6 [2], pp. 41-59.

Niemiec C.P. and Ryan R.M. 2009, *Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice*, in "Theory and Research in Education" 7 [2], pp. 133-144.

Noels K.A. 2001, *New orientations in language learning motivation: Towards a model of intrinsic, extrinsic and integrative orientations*, in Dörnyei Z. and Schmidt R. (Eds.), *Motivation and Second Language Acquisition*, University of Hawai'i Second Language Teaching and Curriculum Centre, Honolulu, HI, pp. 43-68.

Noels K.A., Pelletier L.G., Clément R. and Vallerand R.J. 2000, *Why are you learning a second language? Motivational orientations and self-determination theory*, in "Language Learning" 50 [1], pp. 57-85.

Noels K.A., Lou N.M., Lascano D.I.V., Chaffee K.E., Dincer A., Zhang Y.S.D. and Zhang X. 2019, *Self-determination and motivated engagement in language learning*, in Lamb M., Csizér K., Henry A. and Ryan S. (Eds.), *The Palgrave Handbook of Motivation for Language Learning*, Palgrave Macmillan, London, pp. 95-115.

Norton B. 2016, *Preface*, in Preece S. (Ed.), *The Routledge Handbook of Language and Identity*, New York, Routledge, pp. xxii-xxiv.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English: An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Pavesi M. this volume, *The informalisation of English language learning. A first national research project*.

Pink D.H. 2009, *Drive: The Surprising Truth about What Motivates Us*, Riverhead Books, New York.

Plant R.W. and Ryan R.M. 1985, *Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally controlling styles*, in "Journal of Personality" 53 [3], pp. 435-449.

Reeve J. 2012, *A self-determination theory perspective on student engagement*, in Christenson S.L., Reschly A. and Wylie C. (Eds.), *Handbook of Research on Student Engagement,* Springer, New York, pp. 149-172.

Reinders H. 2010, *Towards a classroom pedagogy for learner autonomy: A framework of independent language learning skills*, in "Australian Journal of Teacher Education" 35 [5], pp. 40-55.

Reinhardt J. 2022, *The history of language learning and teaching beyond the classroom*, in Reinders H., Lai C. and Sundqvist P. (Eds.), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, Oxon, pp. 9-23.

Ryan RM. 1982, *Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory*, in "Journal of Personality and Social Psychology" 43 [3], pp. 450-461.

Ryan R.M. and Deci E.L. 2000, *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being*, in "American Psychologist" 55 [1], pp. 68-78.

Ryan R.M. and Deci E.L. 2017, *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*, Guilford Press, New York.

Salamone J.D. and Correa M. 2012, *The mysterious motivational functions of mesolimbic dopamine*, in "Neuron" 76 [3], pp. 470-485.

Saldaña J. 2021, *The Coding Manual for Qualitative Researchers* (4th ed.), Sage, London.

Seidlhofer B. 2018, *Understanding English as a Lingua Franca: A Complete Introduction to the Theoretical Nature and Practical Implications of English Used as a Lingua Franca*, Oxford University Press, Oxford.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave Macmillan, London.

Sundqvist P. 2019, *Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary*, in "Language Learning & Technology" 23 [1], 87-113.

Sundqvist P. and Sylvén L.K. 2014, *Language-related computer use: Focus on young L2 English learners in Sweden,* in "ReCALL" 26 [1], pp. 3-20.

Talebzadeh N., Elahi Shirvan M. and Khajavy G.H. 2020, *Dynamics and mechanisms of foreign language enjoyment contagion*, in "Innovation in Language Learning and Teaching" 14, pp. 399-420.

Talmy S. 2010, *Qualitative interviews in applied linguistics: From research instrument to social practice*, in "Annual Review of Applied Linguistics" 30, pp. 128-148.

Tapal A., Oren E., Dar R. and Eitam B. 2017, *The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment*, in "Frontiers in Psychology" 8, pp. 1-11.

Trinder R. 2017, *Informal and deliberate learning with new technologies*, in "ELT Journal" 71 [4], pp. 401-412.

Ushioda E. 2009, *A person-in-context relational view of emergent motivation, self, and identity*, in Dörnyei Z. and Ushioda E. (Eds.), *Motivation, Language Identity and the L2 Self*, Multilingual Matters, Bristol, pp. 215-228.

Ushioda E. and Dörnyei Z. 2012, *Motivation*, in Gass S.M. and Mackey. A. (Eds.), *The Routledge Handbook of Second Language Acquisition*, Routledge, London, pp. 396-409.

Vroom V.H. 1964, *Work and Motivation*, Wiley, New York.

White R.W. 1959, *Motivation reconsidered: The concept of competence*, in "Psychological Review" 66 [5], pp. 297-333.

Woodworth R.S. 1918, *Dynamic Psychology*, Columbia University Press, New York.

Zuengler J. 1989, *Identity and IL [interlanguage] development and use*, in "Applied Linguistics" 10 [1], pp. 80-96.

COMPLEXITY ACCURACY AND FLUENCY
IN INFORMAL L2 LEARNING
AMONG UNIVERSITY STUDENTS
Four longitudinal case studies

CAMILLA DE RISO
UNIVERSITÀ DI PAVIA

Abstract – With English increasingly moving out of the classroom and into informal settings, studies on informal second language learning (ISLL) have supplied evidence on how prolonged exposure to informal L2 input can enhance advanced receptive and productive skills in L2 learners (Sockett 2014; Sundqvist, Sylvén 2016; Arnbjörnsdóttir, Ingvarsdóttir 2018; Dressman, Sadler 2020; Toffoli *et al.* 2023). However, only recently has research addressed longitudinal development within ISLL. In this respect, the three constructs of complexity, accuracy and fluency (CAF) have provided a comprehensive framework for analysing L2 learners' acquisitional process longitudinally (Polat, Kim 2014; Kusyk 2017). CAF measures aim to assess the richness, correctness, and naturalness of learners' L2 production and use. Furthermore, the interplay between these three dimensions across time appears to echo the non-linearity of L2 development (Ferrari 2012; Lambert, Kormos 2014; Vercellotti 2015; Michel 2017; Pallotti 2021). Starting from these premises, a longitudinal study was carried out at the University of Pavia, aiming to monitor the evolution of four participants' L2 spoken production over time. CAF trajectories were observed through monologic storytelling tasks, investigating any possible relationship between L2 development and frequency, intensity and type of exposure to informal English input. Data analysis has revealed distinctively different profiles which appear to be affected by type and intensity of media input. It has further confirmed that L2 development is multifaceted, multi-layered and multidimensional in nature, as several individual factors, such as learners' identity and beliefs, also play a key role in language development.

Keywords: Informal English learning; complexity, accuracy, fluency; longitudinal study; individual differences; Italy.

1. Introduction

Over the last two decades, the landscape for second language acquisition and language learners has changed drastically, moving from near-total dependence on the knowledge, expertise and planning of others to an unprecedented level of autonomy and opportunity for self-teaching and 'picking up' new languages (Dressman 2020). Given the broad and

diversified availability of L2 input outside institutional settings, research on informal language learning is flourishing in Europe and beyond, exploring informal second language learning (ISLL) scenarios in different countries with various population groups (i.e. age, degree of instruction, language competence levels), and other social and sociolinguistic factors such as educational and translational policies (Sundqvist, Sylvén 2016; Arnbjörnsdóttir, Ingvarsdóttir 2018; De Wilde *et al.* 2020; Dressman, Sadler 2020; Muñoz, Cadierno 2021; Toffoli *et al.* 2023, among others). While findings to date substantiate the effectiveness of prolonged exposure to online and offline informal input on L2 proficiency (Peters 2018; Kusyk 2020; Ludke 2020; Pattemore, Muñoz 2022; Lyrigkou 2023), only recently has research within ISLL addressed longitudinal L2 development. A longitudinal case study design provides insights into the dynamic, individually owned process of L2 development, enabling to thoroughly monitor the developmental process of a single individual while exploring individual differences related to motivation and attitudes towards the L2. This abundance of data challenges L2 researchers seeking analytical approaches that guarantee comparability and replicability of results. In this respect, the three dimensions of complexity, accuracy and fluency (CAF) – although controversial (Housen *et al.* 2012; Michel 2017; Tavakoli, Wright 2020; Pallotti 2021) – have provided a comprehensive framework for evaluating L2 learner-users' acquisitional paths longitudinally (Polat, Kim 2014; Vercellotti 2015; Kusyk 2017; Peng *et al.* 2022).

Starting from these premises, this article describes the results of a longitudinal study carried out in a medium-sized university in Italy aiming to monitor the evolution of four participants' L2 spoken production over time. CAF trajectories were observed through monologic storytelling tasks, concurrently investigating any relationship between L2 development and frequency, intensity and type of exposure to informal English input. By positioning itself in the eclectic field of ISLL, the study moves from the assumptions that L2 development is multifaceted, multi-layered and multidimensional in nature (Housen *et al.* 2012; Peng *et al.* 2022; Suzuki, Kormos 2023), as several individual factors such as learners' identity and beliefs also play a key role in language development (Arndt 2019; Li *et al.* 2022).

In what follows, Section 2 reviews recent literature on ISLL in the European context and gives a summary overview on CAF theoretical and methodological issues. Section 3 describes the study and provides data on participants' backgrounds and types of exposure to informal English input. Results are discussed in 4 and presented in three sub-sections: (i) an overview of participant's habits when accessing informal English input, (ii) the relationship with English, attitudes and behaviours towards the L2 and (iii) the analysis of CAF trajectories in L2 oral productions through time. As

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

161

remarked in the conclusive section, language develops dynamically in non-linear fashion and it is, therefore, rather challenging to identify a shared process of CAF development, especially when multiple individual factors such as motivation, self-confidence and engagement also come into play.

2. Background of the study

2.1. Informal second language learning

ISLL is defined as a form of second language acquisition which is thought to result from leisure-oriented activities such as watching films or TV series, listening to music, reading, playing video games, and browsing social networks in the L2. These informal activities take place outside educational contexts, are independent of formal frameworks for learning or testing and are self-directed by learners themselves who typically engage in these activities out of personal interests rather than for language learning. The growing access to English outside the classroom goes hand in hand with today's language and media-saturated world – a world where globalisation and mobility have expanded language affordances and multiplied the settings where language acquisition can occur (Benson 2021; Reinhardt 2022).

Over the last 20 years, systematic research on ISLL of English has been conducted in several European countries, at first mostly in the northern area where informal contact with English starts at a very young age and is strongly rooted in everyday lives (Verspoor *et al.* 2011; Sundqvist, Sylvén 2016; Puimège, Peters 2019; Arnbjörnsdóttir, Ingvarsdóttir 2018). Due to the fact that they are traditionally subbing countries, contact with English in northern Europe is extensive among teenagers and continues throughout adulthood, with differences in preferred input sources according to age and gender (Sundqvist 2009; Peters 2018). The massive L2 exposure through media led to higher English proficiency levels and is most evident in vocabulary development, reading, listening, and speaking skills (Arnbjörnsdóttir 2018). Receptive activities like watching television and browsing the web are most popular, though listening to music and playing video games are also widespread (Sylvén, Sundqvist 2012). Because of media specificity, users tend to acquire colloquial and informal language registers, to the detriment of academic writing skills required in the educational and professional sphere (Arnbjörnsdóttir, Ingvarsdóttir 2018). More recently, studies on ISLL conducted in France and Germany (Sockett 2014; Kusyk 2017, 2020; Arndt 2019), Spain, Italy and Greece (Muñoz 2020; Pavesi, Ghia 2020; Lyrigkou 2023), called for a systematic comparison of the results. Although informal contact with English seems to occur at a later age in

traditionally dubbing countries, i.e. in secondary school or university, some trends appear to be quite generalised and shared amid learner-users coming from central and southern Europe when approaching English in extramural contexts. Among audiovisuals, TV series appear to be the most popular choice and for the most part learner-viewers rely on the support of subtitles, though preferred viewing modalities vary according to content availability and L2 proficiency. More advanced users tend to opt for bimodal subtitles or no subtitles at all whereas interlingual subtitles are preferred at lower L2 proficiency levels (Pavesi *et al.* 2019; Pujadas, Muñoz 2020; De Wilde *et al.* 2020; De Riso 2021; Ghia, Pavesi 2021; Muñoz, Cadierno 2021; see Pattemore *et al.* 2024). L2 proficiency and familiarity with extramural English also recur in the polarisation between receptive and productive skills, as learners from central and southern European countries tend to primarily engage with English informally in receptive tasks such as watching AVs, listening to music and searching the web (Kusyk 2020; De Riso 2023; Toffoli 2023). The divide in learner-users' frequency of access to receptive vs. productive input is wider than their northern European counterparts as it appears that higher degrees of linguistic self-confidence can lead to greater willingness to engage in more activities that involve the English language, including those that require L2 production (Li *et al.* 2022; Arndt 2023).

Previous research has hinted at a lack of English proficiency among Italians as self-assessed by participants themselves, as perceived by other Europeans, and as persistently indicated in European reports (see Aiello 2018). Compared to other European countries, Italy appears to fall behind in terms of citizens' competence in L2 English and extent of exposure to the language in informal settings. To date, these results have been much rebutted as new studies – including this – suggest encouraging trends in the growing incidence of informal contact with English in the country, higher self-perceived L2 competence and increased L2 identity awareness among learner-users (Cravidi 2016; Pavesi, Ghia 2020; De Riso 2023; Cicero 2023).

2.2. CAF and L2 development

The multidimensionality of language proficiency and performance can be thoroughly described by the notions of complexity, accuracy and fluency (CAF). Complexity refers to the size, elaborateness, richness and diversity of the L2 performance. Accuracy is a measure of the target-like and error-free use of the L2. Fluency is described as the ability to produce eloquent speech without undue pausing, hesitation or reformulation. Research in Second Language Acquisition (SLA) has frequently used CAF measures as dependent variables to gauge L2 performance operated by independent variables such as task complexity or repetition (Yuan, Ellis 2003; Tavakoli,

Skehan 2005; Skehan, Foster 2008; Foster, Skehan 2012; Lambert *et al.* 2017; Suzuki 2017). In recent years, studies within the field have used CAF to identify developmental change and observe longitudinal learner trajectories (Ferrari 2012; Polat, Kim 2014; Vercellotti 2015; Kusyk 2017; Michel 2017; Peng *et al.* 2022). Regardless of the considerable amount of research available on CAF (Housen, Kuiken 2009; Norris, Ortega 2009; Housen *et al.* 2012; Lambert, Kormos 2014; Pallotti 2009, 2021; Hasnain, Halder 2022 for recent reviews), the constructs' theoretical and operational definitions are still controversial topics of discussion.

Considered as the most controversial construct of the triad, the term complexity is commonly used to refer both to a structural or linguistic dimension, i.e. the elements and intricacies of a system, and to a cognitive dimension, that is the effort needed to process said system (Pallotti 2021). Since linguistic complexity itself is also multidimensional, further disambiguation is needed between lexical and grammatical complexity as both, in turn, feature different subdimensions (e.g. lexical diversity, lexical sophistication, syntactic complexity). To make way in the vast, intricate framework of complexity, several measures have been developed and critically scrutinised to generate a plausible rationale beyond their usage. To date, however, there is still a lack of consensus among researchers as the quest for omnibus measures of complexity is somewhat misguided as it "fails to adequately capture the multidimensional space of variation among complexity features" (Biber *et al.* 2023, p. 23). Unlike complexity, there is less disagreement on accuracy, defined as any deviation from a norm or native-like use of language including diverse types of errors (Housen et al. 2012). Despite the well-established tradition of adopting the 'accuracy perspective', i.e. counting and weighing errors, it can be problematic to define exactly what constitutes an error, especially when lexical, morphological, syntactic, phonological, and spelling errors are bundled together. In this respect, new studies have embraced the 'interlanguage approach' wherein accurate productions "may depend on a more proceduralised norm-oriented interlanguage system and be thus positively related to increased fluency" (Pallotti 2021, p. 205). From a cognitive perspective, fluency is connected to the proceduralisation and automatisation of a language, with functions such as planning and monitoring language production (Tavakoli, Skehan 2005). Pertaining first and foremost to spoken language, three main dimensions of fluency are usually identified: speed, breakdown and repair. Although this fluency model has been thoroughly examined and validated (Skehan et al. 2016; Tavakoli, Wright 2020; Suzuki, Kormos 2023 among others), normal dysfluency phenomena are to be expected when analysing spoken language. Whether it is a monologic/dialogic task or spontaneous conversation, it is quite natural for a

speaker's flow to be characterised by pauses, hesitators and repetitions "at points where the need to keep talking threatens to run ahead of mental planning, and the planning needs to catch up" (Biber *et al.* 2021, p. 1048). While sometimes dysfluency is so pervasive a feature that, by the standards of written language, spoken language is grammatically inchoate, it is normal for speakers to produce utterances with minor dysfluencies which do not interfere with understanding (Biber *et al.* 2021, p. 1052-53).

When measuring L2 learners' proficiency, performance and development, the three dimensions of CAF need to work together to relay exhaustive outcomes. However, earlier research on the multidimensionality of L2 development suggests that due to working memory's limited capacity for processing input and output, L2 learners have difficulties in catering simultaneously to meaning and form, and therefore need to choose which content to prioritise at the expense of the other (Ellis, Barkhuizen 2005; Vercellotti 2015; Michel 2017 for recent reviews). In terms of input, working memory extracts and temporarily stores information from both input and long-term memory. For L2 learners the load on working memory increases, because the extract-and-store processing rely considerably more on controlled processing rather than automatic. For output, L2 learners need to retrieve information from long-term memory and hold it in their short-term memory to produce messages that are meaningful, coherent, and grammatically correct. In real-time conversation, with limited planning time, the overload on working memory could prevent the simultaneous processing of comprehension (meaning) and acquisition (form), leaving L2 learners to prioritise one or the other depending on context and orientation (Ellis, Barkhuizen 2005, p. 142). To simplify, focusing on one CAF construct may result in lower performance in one or both other constructs, that is, trade-off effects. To investigate CAF and L2 development, research has analysed learners' performance by testing the predictions of two theoretical models, i.e. Skehan's (1998) Limited Attention Capacity model and Robinson's (2011) Cognition Hypothesis. Skehan's information-processing model predicts a conflicting relationship among the three CAF components where learners emphasise meaning (fluency) over form, potentially hindering further L2 development. When learners do focus on form, there is a secondary contrast between control of form (accuracy) and use of more advanced language (complexity) (Skehan 2009). Because of limited mental resources, i.e. limited attentional capacity and working memory, all language learners have these tensions during performance and thus entailing a single-source view of attention (Skehan 2015). On the other hand, researchers who reject a single-source capacity limitation accept CAF competitiveness as explained by attentional control and interference (Robinson 2003). Robison's Cognition Hypothesis (2011) expects tasks to promote either fluency or combined complexity and accuracy, a theory which supports Skehan's primary

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

165

trade-off but contrasts with the second. For example, simple monologic tasks are likely to foster fluency to the detriment of complexity or accuracy, whereas accuracy and complexity (but not fluency) are promoted during complex monologic tasks (Robinson 2011). Over the years, many studies have set out to contribute to the debate and despite contrasting results, researchers working in different theoretical frameworks have concluded that trade-off effects impact language performance, supporting either Skehan's or Robinson's theoretical model. Grammatical complexity has been reported to increase at the expense of fluency (Bygate 2001), and a trade-off between fluency and accuracy seems to be a particularly robust finding in the literature (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011). Conversely, findings also suggest a trade-off between complexity and accuracy (Ferrari 2012).

The large variety of research designs and the plethora of CAF measures generating conflicting results call for more standardisation and a unified methodological approach in what is, in all respects, an ongoing debate. After a thorough review of CAF measures and researchers' perspectives regarding the choice of particular metrics, the present study sets out to gain preliminary insights on L2 development among a sample of Italian learner-users of English during a one-year longitudinal study.

3. The study: aims, sample, methods

The present study analyses if, how and to what extent access to several types of informal English input may affect L2 development as part of a wider study on informal contact with English (De Riso 2023). Four longitudinal case studies gauge how different access patterns to different types of informal English input may play a role in learners' L2 development and address the following research questions: i) To what extent does prolonged exposure to informal English input affect L2 development? ii) Is there any relationship between access patterns and learners' L2 proficiency? iii) Which paths do complexity, accuracy and fluency trajectories follow in elicited production tasks over a 12-month period? The study additionally provides qualitative observations about learner-users' attitudes towards the English language and how their motivations may vary in time and in relation to other factors such as increased exposure and better knowledge of the language.

From an initial large-scale questionnaire survey (*n* 605) that explored the degrees and modalities of spontaneous and naturalistic access to English by Pavia University students, 264 respondents consented to take part in the longitudinal study. With the aim of carrying out longitudinal studies, the majority of the selected cohort was made up of non-language specialists, in order to sample a wider variety of first-year undergraduates and postgraduates with a view to carrying out longitudinal studies. Early selection

criteria included self-assessed levels of competence in the language (CEFR B2 and C1) and self-reported exposure indexes calculated by multiplying frequency by length of exposure to the different types of informal input surveyed in the questionnaire (from average to high-input learner-users). 73 respondents were selected and eventually 20[1] participants confirmed their attendance. By the end of the data collection, only 4 subjects had completed the steps of the study successfully (Table 1).

Subject	Gender	Age	Course	CEFR	Exposure index
Delia	F	23	MA Biology	B2	57/140
Grace	F	20	BA Chemistry	B2	58/140
Marta	F	25	MA Biology	C1	64/140
Lea	F	20	BA Biotechnology	B2	105/140

Table 1
Personal information of the four subjects analysed as case studies.

The longitudinal study lasted throughout 2021 with one-on-one online meetings[2] scheduled on Zoom every other month for a total of six meetings per subject. Each data collection involved two main monologic oral tasks elicited through storytelling techniques, i.e. the retelling of a muted film clip,[3] and the recollection of an event from the past.[4] A total of 48 monologues were recorded, transcribed and further analysed according to 15 measures of complexity, accuracy and fluency (Table 2) to observe participants' L2 development trajectories.[5]

Complexity included: grammatical complexity, calculated as mean length of AS-units[6] (Foster *et al.* 2000); lexical diversity, calculated using *vocd* (McKee *et al.* 2000); and lexical sophistication, measured with Lexical

---

[1] 16 females and four males distributed among different courses as follows: Biology and Biotechnology (6), Medicine (4), Communication Studies (2), Pharmacy (2), Chemistry (1), Computer Engineering (1), Environmental Engineering (1), Philosophy (1), Political Science (1), Psychology (1).

[2] Due to COVID-19 restrictions, online meetings were strongly encouraged as opposed to face-to-face.

[3] 5-7 minutes long. January: *The Wedding Date* (Kilner 2004); March: *Sex and the City: The Movie* (King 2008); May: *Four Weddings and A Funeral* (Newell 1994); July: *Runaway Bride* (Marshall 1999); September: *Made of Honor* (Weiland 2008); November: *Mamma Mia!* (Lloyd 2008).

[4] Participants were encouraged to reflect and provide a narrative on their past experiences as follows (from the first to the last meeting): your first time on holiday with friends; your first day at university; your first time surrounded by a big crowd; the first item you purchased something with your own money; your first time going to the cinema; last days of high school.

[5] Data were standardised to bring 15 CAF measures to the same scale (z-score).

[6] Analysis of Speech unit, defined as a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either (Foster *et al.* 2000, p. 365).

Frequency Profile[7] (LFP) (Laufer, Nation 1995). The analysis of *vocd* is based on the probability of new vocabulary being introduced into longer and longer samples of text. The index is automatically calculated[8] through a series of type-token ratio samplings and curve fittings (McNamara *et al.* 2014), and it has been shown to be a useful measure for L2 data (McCarthy, Jarvis 2007, 2010). LFP describes the lexical context of a text in terms of frequency scores (Laufer, Nation 1995).

Complexity
Lexical
1 - Diversity: *vocd*
2 - Sophistication: Lexical Frequency Profile (LFP)
Grammatical:
3 - Length: mean length of AS-unit
Accuracy
4 - Global: Normalised error rate (per 1000 words)
Lexis:
5 - Words in L1
6 - Non-normative word formation
7 - Non-normative word choice
Syntax:
8 - Non-normative use of singular-plural forms
9 - Non-normative use of articles, pronouns, and prepositions
10 - Word order or sentence structure based on L1
Verb:
11 - Non-normative form (conjugation)
12 - Non-normative choice of tense, aspect, and mood
13 - Other (missing word)
Fluency
14 - Reformulations (normalised per 1000 words)
15 - Repeats (normalised per 1000 words)

Table 2
Selected set of measures for complexity, accuracy and fluency.

The BNC-COCA corpus was used as reference for this analysis as it offers a comprehensive sample of modern English usage. This ensures that word frequency categorisation is based on a broad linguistic dataset, enhancing the reliability of the LFP and providing insights into the text's lexical richness and sophistication. Accuracy was measured as percentage of error-free clauses. Looking at the non-normative use of the L2 is recommended (Ellis, Barkhuizen 2005), since specific measures may misinterpret learners' knowledge if learners avoid certain forms or constructions. While the study was initially set to measure speed, breakdown and repair fluency, data were collected during the pandemic when lessons in all institutions, from primary school to university, were held online. It often happened, then, that Zoom video calls would end unexpectedly or that the video or audio would lag whenever the servers were

---

[7] Automatically calculated with Compleat Web VP (frequency frameworks by Laufer *et al.* coded by Tom Cobb, available at https://www.lextutor.ca/vp/comp/).
[8] Using Coh-Metrix (Graesser *et al*.), available at http://cohmetrix-new.memphis.edu/.

overloaded. These unforeseen events compromised the reliability of speed and breakdown fluency; hence, it was decided to focus on repair fluency, i.e. repeats and reformulations.

By embracing a mixed-method approach, the study relied on semi-structured interviews to gain insights on learner-users' personal and L2 identity; every other week participants were asked to keep a journal[9] of their informal activities in the L2 and record whatever thought might have arisen from the experience.

4. Results

Each case study is analysed separately and is divided into three sections: (i) an overview of participant's habits when accessing informal English input, (ii) the relationship with English, attitudes and behaviours towards the L2 (iii) the analysis of CAF trajectories in L2 oral productions through time. Section (iii) analyses two monologic tasks (i.e. storytelling) elicited over a 12-month period, whereas section (i) and (ii) draw data from participants' questionnaire responses, semi-structured interviews, and journals.

4.1. Exposure to informal English input

Participants L2 usage profiles were representative of the informal habits surveyed in the questionnaire (De Riso 2023) insofar as they engaged in multiple informal activities while showing strong preferences for a specific type of input. More specifically, Delia preferred to watch films and TV series, Grace loved to read fantasy novels, Marta listened to podcasts and livestreams and Lea used to play video games online, read manga and watch anime. Their participation in informal activities did not vary particularly over the 12 months, though at times content consumption increased considerably whenever binge-watching (or binge-reading, binge-listening, binge-playing, etc.) occurred. Among the dominant factors leading to engaging in informal activities in English, participants mentioned the hedonic component, a search for naturalness and authenticity, the immersive experience, sociability factors and earlier availability of content in the English language.

Participants' first contacts with informal English input date back to secondary school when they first approached music, YouTube videos, books and video games recreationally. In lack of an Italian counterpart and unwilling to wait longer for an eventual translation, participants turned to

---

[9] An online Word document per participant, shared exclusively with the researcher in order to gain access to data while maintaining privacy among participants.

content in the L2 without having any language-related goal in mind. During high school they began to notice a certain ease with the English language and attributed this easiness to what they were doing out-of-the-classroom rather than their formal instruction. Nevertheless, the main reasons for accessing content in English did not vary. To date, participants' input choices are affected by habit and enjoyment rather than language learning purposes, whilst being aware of the beneficial effect that prolonged exposure to English has on their L2 competence.

Despite the significant exposure to informal English input, it should be mentioned that participants occasionally turn to Italian whenever the activity in the L2 is perceived as too demanding. For example, with AVs, dubbing (when done properly) requires less attention and results in a more relaxing experience, whereas L1 subtitling offers an aid to comprehension when dialogues are too technical. As regards reading, participants opt for Italian fantasy novels and manga whenever they get the chance: as niche genres, there is still little to no original publications (nor translations) available in Italy, thus English is 'necessarily' preferred.

4.2. Attitudes towards English

Data on individual differences, i.e. learner-user traits and characteristics that may have an impact on learning processes, behaviours and outcomes, were gathered through journal entries, semi-structured interviews and spontaneous conversations with participants. On different occasions, participants were asked to recall their first approach to informal English input and give a one-word definition of their relationship with the language. Although attitudes towards the L2 were overall positive, the participants' experiences were quite different as reflected in their motivations and beliefs.

In the case of Delia, her appreciation of the language clashed with a low self-perceived L2 competence, highlighting a complicated relationship with English. In other words, her use of the L2 is partially clouded by the pressure of other people's judgment:

> (1) "When I speak in English, I know what I want to say […] but I get nervous since I don't want to make mistakes, because I imagine that people will assume I'm not good at it […] sometimes it gets harder to transform what I'm thinking into words and sentences, and I begin to stutter."

This conflict, however, appears to be strictly related to conversation as Delia eventually admits that her insecurities might be related to a mild social anxiety rather than a lack of L2 competence:

(2) "I had one paper published, one was recently accepted plus I'm working on my thesis in English, and I don't find it hard at all […] I basically do everything in English it's just that I'm not really extroverted so that could be part of the problem too, speaking in public in general."

Similarly, Grace too underestimates her English proficiency level; however, she believes she has a natural predisposition for the language that led to a strong L2 confidence and language awareness:

(3) "When I was 13 or 14 I began to read books in English because I wanted to read them in their original language. At first it was hard I remember that I used to write all the words that I didn't know in a notebook. […] Then something just clicked […] I was able to read and speak effortlessly."

She defines her relationship with English as "a wonderful love story" and one that is meant to last:

(4) "Every time I get the chance to speak, read, listen to it I'm always very happy […] When I'm doing something and I can choose between Italian or English, I chose English. I will always choose English unless of course the original language is Italian."

That of Marta is one interesting case, as she believes that her bilingual upbringing (in Italian and Spanish) gave her an "extra boost in picking up foreign languages". According to her, "having a bilingual brain is like having multiple language microchips in your head". She describes the process of switching from one language to another as a "change input language command":

(5) "I know I'm privileged from a linguistic point of view; I know that people who grew up in bilingual environments are more inclined to learn new languages […] I consider myself extremely lucky because for me it was a rather spontaneous thing [to learn English]."

Whilst acknowledging her 'head start' in L2 learning, Marta is quite proud of her relationship with English because: "knowing other languages and cultures is both necessary and fundamental nowadays and it's something to be proud of". Quite often, Marta finds herself thinking or writing things down in English and she says that most times she hardly notices which language she is using:

(6) "The more I listen to it the more I think in English. I occasionally write my thoughts in a personal journal and sometimes I find myself writing in English, like it's natural to me"

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

171

Rather similar feelings were shared by Lea. She considers English almost as a "routinized second language" and is extremely proud of her ability to communicate effectively with both native and non-native speakers – an activity in which she engages at every chance she gets:

(7) "I simply couldn't be without English, it allowed me to meet new people and cultures from all around the world […] I use it so often almost as if it was my second language"

(8) "I was playing [a video game] online with my Norwegian friend and a friend of mine joined us, he's Italian. I was chatting in English as usual and noticed that my Italian friend was very quiet, so I began to translate things to him, I was like an interpreter. And later he thanked me, because he was really struggling to follow what we were saying and was like: "How can you talk so easily?" and I didn't know what to say it was a normal thing for me."

4.3. CAF trajectories over time

A total of 48 monologues elicited through storytelling were recorded, transcribed and further analysed according to 15 selected CAF measures (Section 3 Table 2). Each case study was analysed individually; however, few similarities can be observed across the results.

Throughout the longitudinal study, CAF trajectories presented strong variability across participants and within individual L2 development, showing a constant push and pull among measures, i.e. progression and regression. The trajectories of accuracy and fluency show that the two dimensions appear to be particularly sensitive to each other. While this result is consistent with previous literature (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011), the trade-off[10] between accuracy and fluency found in the present study led to different outcomes among participants. Findings from Delia's case study (Figure 1) seem to support Skehan's primary competition between meaning and form, as grammatical complexity (mean length of AS-units) appears to increase at the expense of fluency (measured by the number of repairs).

---

[10] The present data analysis relied on descriptive statistics of a small sample of students, consequently, assumptions of trade-off effects should be made conservatively.
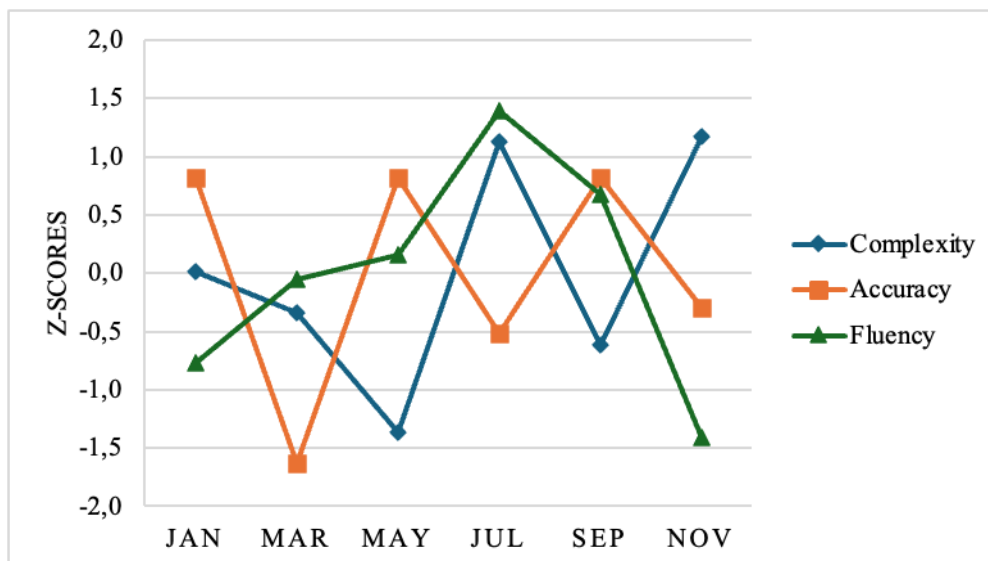
Figure 1
CAF trajectories: Delia's case study.

Conversely, the results from Marta's case study (Figure 2) show that learner-users can have higher accuracy and complexity at the expense of fluency, thus supporting Robinson's Cognition Hypothesis but refuting Skehan's secondary contrast with form.
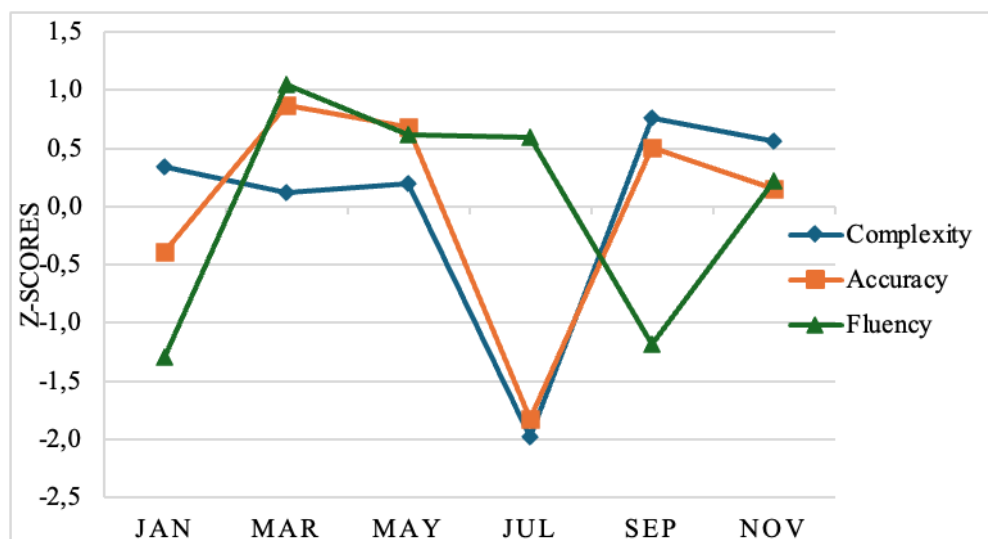
Figure 2
CAF trajectories: Marta's case study.

Even though these findings partially support competitiveness among CAF components, the fluctuations observed among participants show just how complex and dynamic L2 development is and how difficult it is to identify a shared process of CAF development.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

173

The study further compared CAF trajectories with participants' habits of exposure to the different sources of informal English input. Among the several informal activities investigated in the study, extensive reading appears to be by far the most effective for L2 development in the sample. Findings from Grace's case study suggest that reading novels extensively can lead to improved complexity, i.e. more structured utterances accompanied by a diverse, sophisticated vocabulary. The following examples compare Grace's L2 oral production in March (9) and November (10):

> (9) "We were celebrating [New Year's Eve] in the square that is right in front of the castle and there was a spectacular fireworks' show it was exceptional. The fireworks started from the walls of the castle, and they were creating these beautiful, beautiful patterns in the sky and there was this huge crowd, and they were all celebrating. There were people screaming, dancing, people kissing, just having a good time in general."

> (10) "So, I basically found my footing and other than that I don't remember much of my actual exam. I do remember coming out of there and feeling very like relieved like a great burden had been taken off my shoulders, but at the same time very sad because that was the moment that I had waited for, for five years. And if you hear people talking about the, the exam they all- it always seems like this big huge thing something impossible to overcome and, and I know it gives you like so much anxiety and stress and you're so nervous and then you get there and you just blink and it's gone and you did it and you don't really know how you did it, but you got through it and you're like, 'That was it?'"

In the context of SLA, researchers agree on the value of extensive reading for L2 learning but they also warn about the possible risks of text selection. Ultimately, learners could select a text that is too simple compared to their proficiency level to the detriment of L2 learning (Arnold 2009; Grabe, Yamashita 2022). This seems to be the case of Lea who, despite reading plenty of comics and manga, did not show signs of L2 improvement throughout the longitudinal study. The following examples compare Lea's L2 oral production in March (11) and November (12):

> (11) "And during the breaks, I tried to speak with some- somebody else. To get to, to try and know them. And in the end, I met the girl that I went- I went home together on the train. To get to Pavia, I went with my, my guy friend, and when I, when I had to return, I was with him and the girl I met. So, we were all three coming back home together, and it was kind of nice."

> (12) "And then we went- we just went to high school that was like five minutes by foot for my home so not, not that far. I did my exam. It was kind of

> funny because it happened a lot of the stuff that my friends call it lucky? So,
> I was nervous at, at the beginning but it ended up being something I, I laugh
> while I- when I tell it. And when I was done I was, I was still with this
> friend of mine and we went, we went to eat outside- out."

By comparing Grace and Lea's L2 oral productions, it appears that exposure to different literary genres affects L2 development in a different way: if, on the one hand, reading prose led to more structured utterances with rich, complex vocabulary, on the other, reading comics, typically written with short sentences and simple vocabulary, resulted in a similar way of speaking.

Throughout the study, all four participants were largely exposed to informal L2 audiovisual input, and while it would be an overstatement to imply that exposure to AVs did not affect L2 development in the sample, findings differ and, in part, contradict each other. As in Kusyk (2017), it appears that watching AVs informally positively affects L2 development in the measure of lexical sophistication. However, if, in the case of Delia, extensive exposure to informal L2 AV input led to a moderately more sophisticated vocabulary, the same did not occur with Marta, whose lexical sophistication slightly worsened despite the prolonged exposure to AVs (see Appendix). It is worth mentioning that these different L2 development outcomes within exposure to the same source of informal input may have been affected by individual factors, such as motivation and media immersion. Although each participant accessed the different types of informal input extensively, some were more motivated than others and appreciated the informal activity to the point of experiencing media immersion constantly. These subjects (i.e. Grace and Delia) were the ones who ultimately showed signs of improvement compared to the others. Lastly, during the longitudinal study participants used the L2 productively by habitually interacting with both native and non-native speakers of English through face-to-face interactions, face-to-screen interactions (voice calls and video calls), text messages and chats. Although they all perceived improvements in their L2 competence following these activities, this development was not always reflected in their CAF trajectories, especially with regard to fluency.

The evolution of complexity, accuracy and fluency trajectories over time for each case study is available in the Appendix.

5. Conclusions

By positioning itself in the eclectic field of ISLL, the present study observed the evolution of CAF trajectories in L2 oral productions of four case studies over a 12-month period, aiming to investigate any potential relation among

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

175

frequency, intensity, and type of exposure to informal English input and L2 development. Moreover, the study explored the thoughts and feelings of L2 learner-users who accessed different sources of informal English input extensively, by investigating participants' L2 awareness in relation to themselves, others, and the informal activities they access habitually.

While prolonged exposure to informal English input arguably affects L2 development, findings show that individual factors such as motivation and patterns of access to the informal input seem to have influenced L2 competence in different ways. For example, both Grace and Lea engaged in reading extensively. Nevertheless, their L2 development had different outcomes perhaps due to the specificity of the two preferred literary genres (i.e. novels vs. comics) and participants' higher vs lower motivation, which could lead (or not lead) to media immersion. On the one hand, Grace's lexical and grammatical complexity improved greatly throughout the longitudinal study, reflecting the sophisticated structures and rich vocabulary typical of prose. She showed strong tendencies of media immersion, reading up to 5 hours a day to finish the novel she was reading. On the other, Lea used to read comics and manga every day while doing other activities, e.g. commuting on the train, while having breakfast, thus lacking media immersion and motivation to read the whole story at once. Despite extensive exposure, her L2 development lacked grammatical and lexical complexity, perhaps echoing the simple structures and minimal vocabulary of comics and manga and suggesting that less attention and motivation towards the informal activity may interfere with L2 development. Similarly, Delia and Marta were both largely exposed to informal L2 audiovisual input; however, the different modalities of access and frequency of exposure led to different outcomes in terms of lexical sophistication. By the end of the longitudinal study, Delia, who was strongly motivated and used to watch at least one film or episode per day, used moderately more sophisticated vocabulary. Marta, on the contrary, used to leave the video on while doing other activities such as cleaning or cooking and her lexical sophistication slightly worsened throughout the study. Overall, CAF trajectories presented strong variability across participants and within individual L2 development. Among CAF measures, the two dimensions of accuracy and fluency appeared to be particularly sensitive to each other and while this result is broadly consistent with a trade-off interpretation (Yuan, Ellis 2003; Michel *et al.* 2007; Ahmadian, Tavakoli 2011), findings differ and sometimes contradict each other (cf. Figure 1 and Figure 2 in Section 4.3). The fluctuations observed among the four case studies show just how complex and dynamic L2 development is and how difficult it is to identify a shared process of CAF development.

The present contribution wanted to provide better insights on informal L2 learner-users' individual profiles as well as analyse *if*, *how*, and *to what extent* access to informal English input may affect L2 development analysed

by CAF measures. However, some limitations of the data collection emerged and will need to be fully addressed in future research. The modality of data collection (recorded Zoom video calls) and the unexpected issues that came within compromised the reliability of speed and breakdown fluency, thus limiting the analysis to repair fluency, i.e. repeats and reformulations. Although the longitudinal monitoring of CAF trajectories allowed to thoroughly observe the evolution of participants' L2 oral production, the lack of pre/post L2 proficiency tests made it difficult to properly assess L2 development and identify subtle improvements, especially if participants were already proficient in the L2. Finally, results might have been somehow influenced by task design. At each data collection participants were asked to tell the same story, i.e. Test 1: four participants tell story A; Test 2: four participants tell story B, etc. Perhaps, results would have varied if each storytelling task was distributed differently among the four informants every time, ensuring that the task construct did not influence overall task performance over time i.e. Test 1: participant 1 tells story A, participant 2 tells story B, participant 3 tells story C, …; Test 2: participant 1 tells story C, participant 2 tells story A, participant 3 tells story B, etc.[11]

Bionote: Camilla De Riso is a postdoctoral researcher in Linguistics at the University of Pavia, where she earned both her MA in Theoretical, Applied and Modern Languages Linguistics (2019) and PhD in Linguistics (2023). Her doctoral thesis examined the impact of informal English input on second language acquisition through a longitudinal perspective, analysing the development of complexity, accuracy and fluency in students' L2 oral production. She has been a visiting fellow at Cardiff University (2014), Karlsruhe University of Education (2021), and the University of Malta (2022), enriching her expertise in content and language integrated learning (CLIL) and quantitative methods in applied linguistics. Her research interests include informal second language learning, audiovisual translation, language learner identities, motivations and attitudes, as well as the constructs of complexity, accuracy and fluency (CAF) in L2 development.

Author's address: camilla.deriso@unipv.it

---

[11] The author would like to thank the reviewer for this valuable insight which will contribute to improve future research.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies
177

References

Ahmadian M.J. and Tavakoli M. 2011, *The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production*, in "Language Teaching Research" 15 [1], pp. 35-59.

Aiello J. 2018, *Negotiating Englishes and English-speaking identities*, Routledge, London.

Arnbjörnsdóttir B. 2018, *Using English at University*, in Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.), *Language Development Across the Life Span: The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham, pp. 143-162.

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development across the Life Span: The Impact of English on Education and Work in Iceland*, Springer International Publishing, Cham.

Arndt H. 2019, *Informal second language learning: The role of engagement, proficiency, attitudes, and motivation*, unpublished doctoral thesis, University of Oxford.

Arndt H. 2023, *Construction and validation of a questionnaire to study engagement in informal second language learning*, in "Studies in Second Language Acquisition", 45 [5], pp. 1456-1480.

Arnold N. 2009, *Online extensive reading and advanced foreign language learners: An evaluation study*, in "Foreign Language Annals" 42 [2], pp. 340-366.

Benson P. 2021, *Language Learning Environments*, Multilingual Matters, Bristol.

Biber D., Johansson S., Leech G., Conrad S. and Finegan E. 2021, *Longman Grammar of Spoken and Written English*, Longman, London.

Biber D., Larsson T. and Hancock G.R. 2023, *The linguistic organization of grammatical text complexity: comparing the empirical adequacy of theory-based models*, in "Corpus Linguistics and Linguistic Theory", 20 [2], pp. 347-373.

Bygate M. 2001, *Effects of task repetition on the structure and control of oral language*, in Bygate M., Skehan P. and Swain M. (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, Pearson Education Limited, pp. 23-48.

Cicero A. 2023, *Motivation and attitudes in EFL learning: A case study of Italian high school students*, unpublished doctoral thesis, University of Pavia.

Cravidi G. 2016, *Accessing English Input Outside University: A Qualitative and Quantitative Investigation*, unpublished postgraduate thesis, Università degli Studi di Pavia.

De Riso C. 2021, *English audiovisual input: a study of teenagers' habits, attitudes, and responses to subtitling versus dubbing*, in "RILA - Rassegna Italiana di Linguistica Applicata" 1 [2], pp. 245-260.

De Riso C. 2023, *Out-of-the-classroom English input and informal Second Language Acquisition: A longitudinal study of Italian university students*, unpublished doctoral thesis, University of Pavia.

De Wilde V., Brysbaert M. and Eyckmans J. 2020, *Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important?*, in "Bilingualism: Language and Cognition" 23, pp. 171-185.

Dressman M. 2020, *Introduction*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 1-12.

Dressman M. and Sadler R.W. (Eds.) 2020, *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken.

Ellis R. and Barkhuizen G. 2005, *Analysing Learner Language*, Oxford University Press, Oxford.

Ferrari S. 2012, *A longitudinal study of complexity, accuracy, and fluency variation in second language development*, in Housen A., Kuiken F. and Vedder I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam, pp. 277-298.

Foster P., Tonkyn A. and Wigglesworth G. 2000, *Measuring spoken language: a unit for all reasons*, in "Applied Linguistics" 21 [3], pp. 354-75.

Foster P. and Skehan P. 2012, Complexity, accuracy, fluency and lexis in task-based performance, in Housen A., Kuiken F. and Vedder I. (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, John Benjamins, Amsterdam, pp. 199-220.

Ghia E. and Pavesi M. 2021, *Choosing between dubbing and subtitling in a changing landscape*, in "Lingue e Linguaggi" 46, pp. 161-177.

Grabe W. and Yamashita J. 2022, *Reading in a Second Language: Moving from Theory to Practice*, Cambridge University Press, Cambridge.

Hasnain S. and Halder S. 2022, *Intricacies of the Multifaceted Triad-Complexity, Accuracy, and Fluency: A Review of Studies on Measures of Oral Production*, in "Journal of Education" 204 [1], pp. 145-158.

Housen A. and Kuiken F. 2009, *Complexity, accuracy and fluency in second language acquisition*, in "Applied Linguistics" 30 [4], pp.461-473.

Housen A., Kuiken F. and Vedder I. (Eds.) 2012, *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, John Benjamins, Amsterdam.

Kusyk M. 2017, *Les dynamiques du développement de l'anglais au travers d'activités informelles en ligne: une étude exploratoire auprès d'étudiants français et allemands*, unpublished doctoral thesis, University of Strasbourg/Karlsruhe University of Education.

Kusyk M. 2020, *Informal English learning in France*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 333-348.

Lambert C. and Kormos J. 2014, *Complexity, accuracy, and fluency in task based L2 research: Toward more developmentally based measures of second language acquisition*, in "Applied Linguistics" 35 [5], pp. 606-614.

Lambert C., Kormos J. and Minn D. 2017, *Task repetition and second language speech processing*, in "Studies in Second Language Acquisition" 39 [1], pp. 167-196.

Laufer B. and Nation P. 1995, *Vocabulary size and use: Lexical richness in L2 written production*, in "Applied Linguistics" 16 [3], pp. 307-322.

Li S., Hiver P. and Papi M. 2022, *The Routledge Handbook of Second Language Acquisition and Individual Differences*, Routledge, New York.

Ludke K.M. 2020, *Songs and music*, in Dressman M. and Sadler R.W. (eds), *The Handbook of Informal Language Learning*, Wiley-Blackwell, Hoboken, pp. 203-213.

Lyrigkou C. 2023, *Informal second language learning and EFL learners' spoken use of discourse markers*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure*, De Gruyter, Berlin, pp. 91-120.

McCarthy P.M. and Jarvis S. 2007, *vocd: a theoretical and empirical evaluation*, in "Language Testing" 24 [4], pp. 459-488.

McCarthy P.M. and Jarvis S. 2010, *MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment*, in "Behavior Research Methods" 42, pp. 381-392.

McKee G.D., Malvern D. and Richards B. 2000, *Measuring vocabulary diversity using dedicated software*, in "Literary and Linguistic Computing" 15, pp. 323-37.

McNamara D.S., Graesser A.D., McCarthy P.M. and Cai Z. 2014, *Automated Evaluation of Text and Discourse with Coh-Metrix*, Cambridge University Press, Cambridge.

Michel M. 2017, *Complexity, Accuracy and Fluency (CAF)*, in Loewen S. and Sato M. (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition*, Routledge, London, pp. 50-68.

Michel M.C., Kuiken F. and Vedder I. 2007, *The influence of complexity in monologic versus dialogic tasks in Dutch L2*, in "International Review of Applied Linguistics" 45 [13], pp. 241-59.

Muñoz C. 2020, *Boys like games and girls like movies. Age and gender differences in out-of-school contact with English*, in "RESLA" 33, pp. 172-202.

Muñoz C. and Cadierno T. 2021, *How do differences in exposure affect English language learning? A comparison of teenagers in two learning environments*, in "Studies in Second Language Learning and Teaching" 11 [2], pp. 185-212.

Norris J.M. and Ortega L. 2009, *Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity*, in "Applied Linguistics" 30 [4], pp. 555-578.

Pallotti G. 2009, *CAF: Defining, refining and differentiating constructs*, in "Applied Linguistics" 30 [4], pp. 590-601.

Pallotti G. 2021, *Measuring complexity, accuracy, and fluency (CAF)*, in Winke P. and Brunfaut T. (Eds.), *The Routledge Handbook of Second Language Acquisition and Language Testing*, Routledge, New York, pp. 201-210.

Pattemore A. and Muñoz C. 2022, *Captions and learnability factors in learning grammar from audio-visual input*, in "The JALT CALL Journal" 18 [1], pp. 83-109.

Pattemore A., Suárez M.D.M. and Muñoz C. 2024, *Perceptions of learning from audiovisual input and changes in L2 viewing preferences: The roles of on-screen text and proficiency*, in "ReCALL", 36 [2], pp. 135-151.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Pavesi M., Perego E., Ghia E. and Colletto V. 2019, *Assessing subtitling in a dubbing country: an empirical investigation on learner-viewers*, in "Studi Italiani di Linguistica Teorica e Applicata" 48 [1], pp. 5-26.

Peng H., Lowie W. and Jager S. 2022, *Unravelling the idiosyncrasy and commonality in l2 developmental processes: A time-series clustering methodology*, in "Applied Linguistics" 43 [5], pp. 891-911.

Peters E. 2018, *The effect of out-of-school exposure to English language media on learners' vocabulary knowledge*, in "International Journal of Applied Linguistics" 169 [1], pp. 142-167.

Polat B. and Kim Y. 2014, *Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development*, in "Applied Linguistics" 35 [2], pp. 184-207.

Puimège E. and Peters E. 2019, *Learners' English Vocabulary Knowledge Prior to Formal Instruction: The Role of Learner-Related and Word-Related Variables*, in "Language Learning" 69 [4], pp. 943-977.

Pujadas G. and Muñoz C. 2020, *Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles*, in "Studies in Second Language Acquisition" 42 [3], pp. 551-575.

Reinhardt J. 2022, *The History of Language Learning and Teaching Beyond the Classroom*, in Reinders H., Lai C. and Sundqvist P. (Eds.), *The Routledge Handbook of Language Learning and Teaching Beyond the Classroom*, Routledge, Oxon, pp. 9-23.

Robinson P. 2003, *Attention and memory during SLA*, in Doughty C.J. and Long M.H.

(Eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, Oxfrod, pp. 631-678.

Robinson P. 2011, *Second language task complexity, the cognition hypothesis, language learning and performance* in Robinson P. (Ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*, John Benjamins, Amsterdam, pp. 3-38.

Skehan P. 1998, *A Cognitive Approach to Language Learning*, Oxford University Press, Oxford.

Skehan P. 2009, *Modelling second language performance: integrating complexity, accuracy, fluency, lexis*, in "Applied Linguistics" 30 [4], pp. 1-23.

Skehan P. 2015, *Limited attentional capacity and cognition: Two hypotheses regarding second language performance on tasks*, in Bygate M. (Ed.), *Domains and directions in the development of TBLT,* John Benjamins, Amsterdam, pp. 123-155.

Skehan P. and Foster P. 2008, *Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the ealing research*, in Van Daele S, Housen A., Kuiken F., Pierrard M. and Vedder I. (Eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching*, University of Brussels Press, Brussels, pp. 207-226.

Skehan P., Foster P. and Shum S. 2016, *Ladders and snakes in second language fluency*, in "International Review of Applied Linguistics in Language Teaching" 54 [2], pp. 97-111.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave MacMillan, London.

Sundqvist P. 2009, *Extramural English matters: Out-of-school English and Its Impact on Swedish Ninth Graders' Oral Proficiency and Vocabulary*, unpublished doctoral thesis, Karlstad University Studies.

Sundqvist P. and Sylvén L.K. 2016, *Extramural English in Teaching and Learning: From Theory and Research to Practice*, Palgrave Macmillan, London.

Suzuki M. 2017, *Complexity, accuracy, and fluency measures in oral pre-task planning: A synthesis*, in "Second Language Studies" 36 [1], pp. 1-52.

Suzuki S. and Kormos J. 2023, *The multidimensionality of second language oral fluency: Interfacing cognitive fluency and utterance fluency*, in "Studies in Second Language Acquisition" 45 [1], pp. 38-64.

Sylvén L.K. and Sundqvist P. 2012, *Gaming as extramural English L2 learning and L2 proficiency among your learners*, in "ReCALL" 24 [3], pp. 302-321.

Tavakoli P. and Wright C. 2020, *Second Language Speech Fluency: From Research to Practice,* Cambridge University Press, Cambridge.

Tavakoli P. and Skehan P. 2005, *Strategic planning, task structure, and performance testing*, in Ellis R. (Ed.), *Planning and task performance in a second language,* John Benjamins, Amsterdam, pp. 239-273.

Toffoli D. 2023, *Learner profiles and ISLL trajectories*, in Toffoli D., Sockett G. and Kusyk M. (Eds.), *Language Learning and Leisure*, De Gruyter, Berlin, pp. 291-325.

Toffoli D., Sockett G. and Kusyk M. (Eds.) 2023, *Language Learning and Leisure*, De Gruyter, Berlin.

Vercellotti ML. 2015, *The development of Complexity, Accuracy, and Fluency in second language performance: A longitudinal study*, in "*Applied Linguistics*" 38 [1], pp. 90-111.

Verspoor M., de Bot K. and van Rein E. 2011, *English as a foreign language. The role of out-of-school language input*, in De Houwer A. and Wilton A. (Eds.), *English in Europe Today. Sociocultural and Educational Perspectives*, John Benjamins, Amsterdam, pp. 148-166.

Complexity accuracy and fluency in informal L2 learning among university students.
Four longitudinal case studies

181

Yuan F. and Ellis R. 2003, *The effects of pre- task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production*, in "Applied Linguistics" 24 [1], pp. 1-27.

Appendix

L2 development: observing CAF trajectories through time (preferred informal input source in brackets).

a) Findings from the case study: Delia (films and TV series)
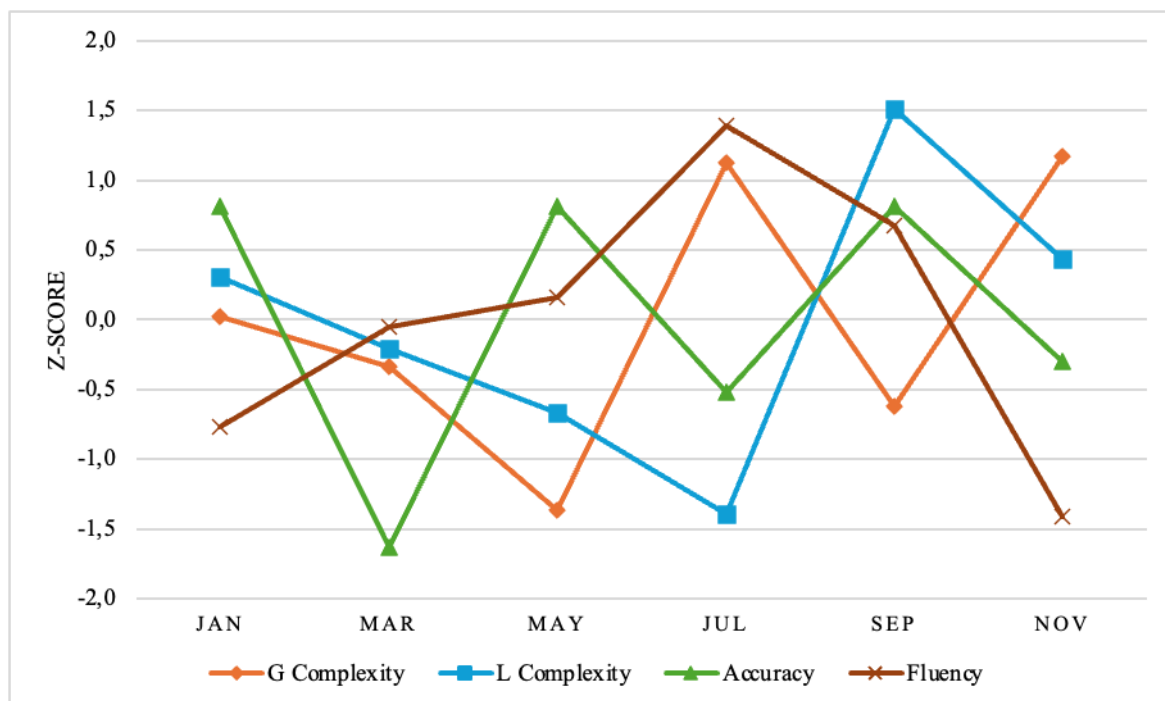
Figure 3
CAF trajectories: Delia's case study.

b) Findings from the case study: Grace (fantasy novels)
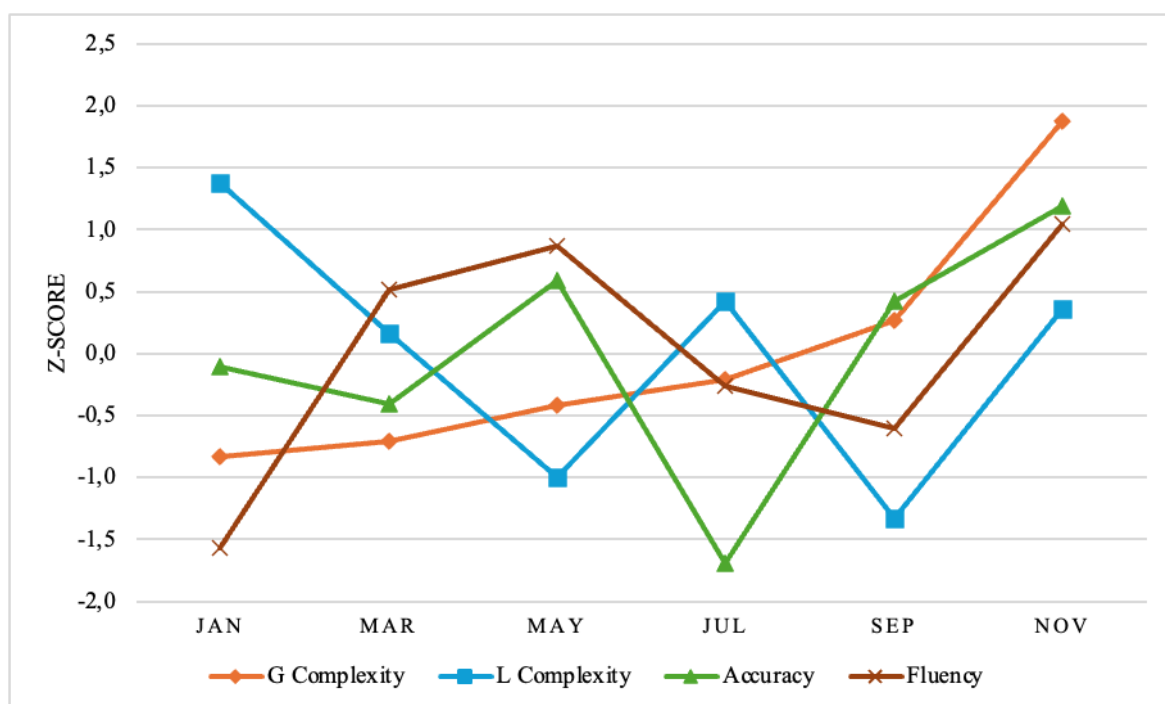
Figure 4
CAF trajectories: Grace's case study.

c) Findings from the case study: Marta (podcasts and livestreams)
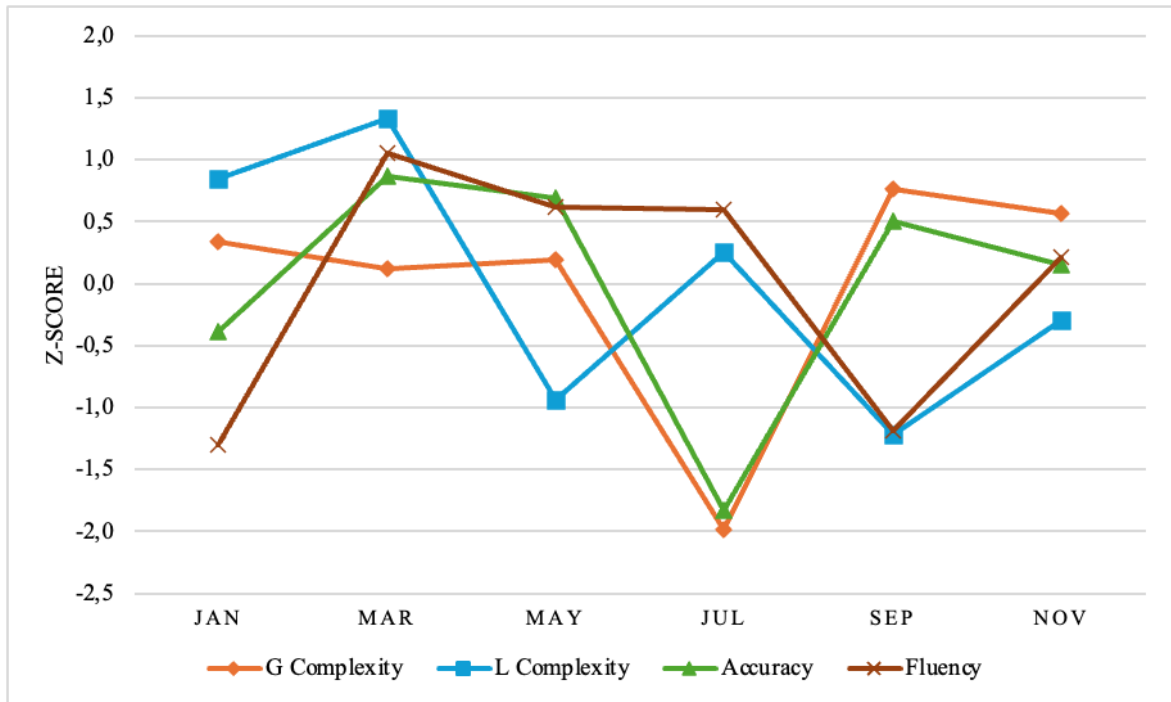
Figure 5
CAF trajectories: Marta's case study.

d) Findings from the case study: Lea (video games, anime and manga)
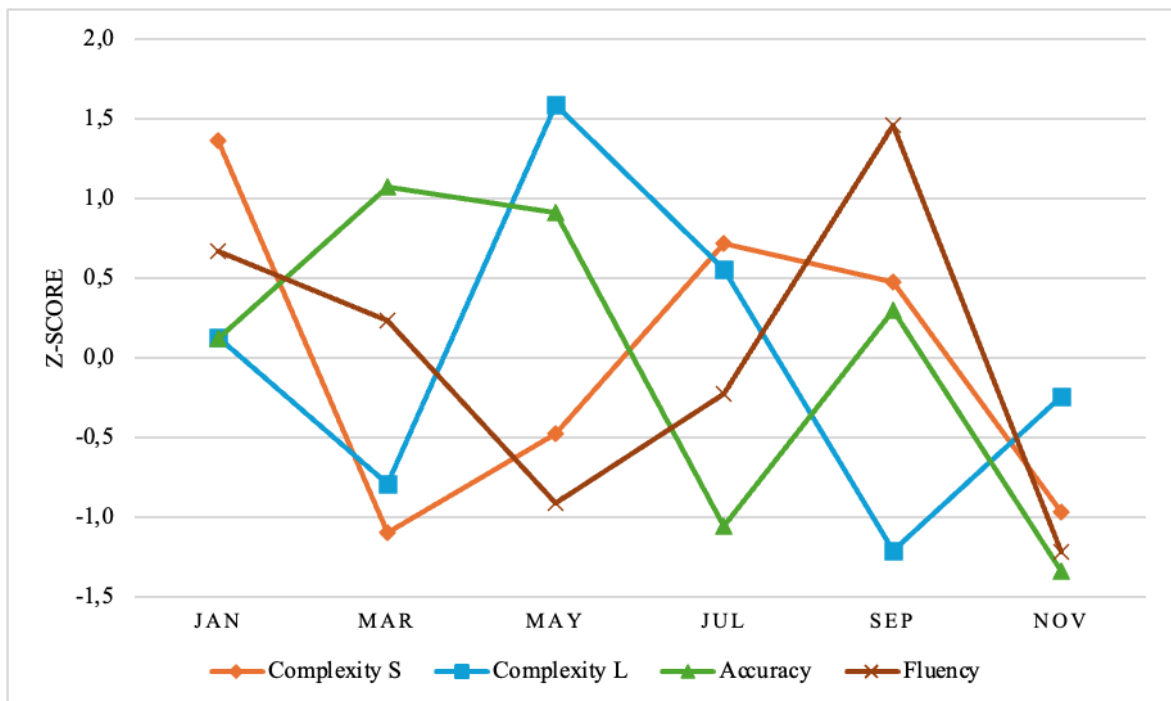
Figure 6
CAF trajectories: Lea's case study.

INSTAGRAM ENGLISH MICROLESSONS
A new genre for non-formal EFL learning[1]

GLORIA CAPPELLI, NICOLETTA SIMI
UNIVERSITÀ DI PISA

Abstract – Academic interest in *Instagram's* role as an informal language learning platform has grown over the past few years. This study systematically analyses *Instagram* accounts specifically dedicated to teaching English as a Foreign Language (EFL), aiming to identify common features that establish them as a genre and to provide a descriptive framework for further empirical research into the learning outcomes of this type of language exposure. This qualitative investigation examines the characteristics of 6 institutional and 21 non-institutional English language teaching accounts. Using categories from traditional classroom language instruction as the analytical framework (non-formal, contents and topics, teachers and learners and methodology and modes of presentation), this study compares informal *Instagram* teaching to traditional formal EFL education. Additionally, it contrasts the strategies employed by institutional accounts with those used by private, micro-celebrity teachers.

Keywords: informal English language learning; *Instagram* ELT; social media; multimodality.

1. Introduction

In recent years, the emergence of social media, the progress in mobile and web-based technology and the socio-cultural impact of global events such as the COVID-19 pandemic have reshaped the language learning and teaching landscape. They have modified learners' interest in face-to-face vs. online education and their awareness of the affordances of social media platforms for foreign language learning that offer dynamic and interactive spaces beyond traditional educational settings (Dos Santos 2022; Teng *et al.* 2022).

For instance, *Instagram* effectively integrates multimodal content, textual communication and social interaction. It is the third most widely used medium by Generation Z users, non-formal, people born between 1997 and 2012, alongside TikTok and Snapchat (Statista 2024)[2] and, according to GWI,

---

[1] The article is the result of joint research. Section 1 was written jointly by the authors. Gloria Cappelli wrote sections 2, 3, 4, and 5. Nicoletta Simi wrote sections 4.1, 4.2, and 4.3.

[2] https://www.statista.com/statistics/1276021/instagram-snapchat-tiktok-gen-z-users/ (2.4.2024).

already in 2022,[3] it was the most used by Generation Alpha users (non-formal, people born after 2013). For this reason, over the past few years, a growing number of studies have focused on its potential for English language teaching and learning, thus contributing to the lively debate about the role of social media in formal, non-formal and informal education in general and foreign language learning in particular (Erarslan 2019; Gonulal 2019; Lailiyah, Setiyaningsih 2020; Teng *et al.* 2022).

Some studies have explicitly focused on *Tiktok* or *Instagram* teachers' and education influencers' accounts from a variety of perspectives. Vizcaíno-Verdú and Abidin (2023) have discussed the way in which influential teachers' identity is built on social media through micro-celebrification practices. Aslan's (2024) analysis has tackled their multimodal meaning-making practices, and Tommaso (2024) has explored the recurrent rhetorical patterns and linguistic features for achieving pedagogical and promotional communicative purposes. Canani and Zulli (2022) have investigated the impact of the content they produce when integrated into classroom practice, while others have studied the effects on learners' motivation (Simsek 2023, Meirbekov *et al.* 2024). However, to the best of our knowledge, no study has systematically analysed this type of English as a foreign language (EFL) resource on *Instagram* as a genre per se by looking at the pedagogic approach adopted, the contents offered and the skills focused on. This contribution aims to fill this gap in the literature through a qualitative investigation of these aspects in 27 *Instagram* accounts specifically aimed at teaching English to establish whether a set of recognizable features can be identified which justify their description as a distinctive genre in the panorama of the digital resources for non-formal English language learning.

Informal language learning has been typically defined as implicit and incidental learning resulting from exposure to materials not specifically designed to foster language acquisition (Toffoli *et al.* 2023). However, it has become increasingly evident that many variables are involved in language learning outside formal education in the 21st century, including but not limited to learners' intentionality and the type of input to which they are exposed. The many labels used to refer to language learning beyond the traditional contexts (e.g. *Extramural English*, *Informal Digital Learning of English, Online Informal Learning of English, Informal Second Language Learning*) bear witness to the complexity of the discussion.

The type of *Instagram* content discussed in this study is specifically designed for English language learning and, therefore, lacks the predominantly unintentional, unaware, or semiconscious nature of the "prototypical" informal language learning outcome which might result from watching a movie in English or listening to English music just for fun. As Sundqvist (2024, p. 2)

---

[3] https://ecommercedb.com/insights/chart/12590 (2.4.2024).

writes, *Instagram* ELT posts and reels are "pedagogical in nature, although technically, they are examples of informal learning". For this reason, for the sake of the present discussion, we will adopt Dressman's (2020) inclusive definition of informal language learning as "all activities undertaken by learners outside a formally organised program of language instruction" (Dressman 2020, p. 4). We consider being exposed to *Instagram* ELT content as one of the many contemporary forms that informal language learning can assume. However, the specific nature of the accounts investigated points to a relevant involvement of at least some form of intentionality, generic interest or linguistic curiosity on the part of the users, having English teaching as their focus. For this reason, we will use the more neutral label "non-formal" to discuss such multimodal input that learners intentionally consume outside of formal education and that engages attentive processes of metalinguistic awareness and noticing (Schmidt 1995), thus leading to prevalently explicit learning. Given the increasing popularity of this form of learning, it seems fair to assume that intentionality cannot be considered a distinctive feature of formal language learning exclusively, but rather, "intention to learn the language is not necessary, but may be present in IDLE" (Informal Digital Learning of English; Kusyk 2023, p. 57).

Several research questions guided our data collection and analysis. Firstly, we aimed to investigate the differences and similarities between English teaching in traditional formal settings and the instructional content for non-formal language learning on *Instagram*. To achieve this, we systematically deconstructed the teaching and learning process into methodology, participants and topics covered, analysing each component individually to determine whether they can contribute to defining *Instagram* ELT as a distinct genre. Secondly, we tried to determine if these dimensions vary between the accounts of "micro-celebrities" and well-established ELT institutions and companies (e.g. comparing accounts such as @britishcouncilenglishonline with those of popular education influencers).

The paper is divided into five parts. Section 2 provides an overview of the most recent literature on *Instagram* as an ELT and EFL learning platform. Section 3 presents the study's methodology. Section 4 is dedicated to analysing the collected materials, while Section 5 offers some concluding remarks.

2. *Instagram* and non-formal English language learning

Social media have revolutionised information sharing by introducing interactivity, moving beyond the unidirectionality of the "broadcast age" (Manning 2014), when knowledge was passed from news outlets to their audience and feedback was (if any) delayed and impersonal. In a way, social media have broken the fourth wall and have allowed fast interaction between

all actors involved in communication. Their portability, affordability and accessibility have favoured their global spread. Their popularity among the younger generations has inspired many educators and researchers to study their potential for teaching and learning purposes, including their affordances for foreign language instruction. Studies on *Instagram* for English language learning have mainly been produced within the broader research domain of Mobile-Assisted Language Learning (MALL), which has attracted increasing attention over the past decade (Crompton, Burke 2018; Kamasak *et al.* 2021; Karakaya, Bozkurt 2022; Lei *et al.* 2022; Teng *et al.* 2022) and has been investigated in its potential for learning across locations and contexts, including informal language learning (Barrot 2022; Guo, Lee 2023; Lee 2022; Reinders *et al.* 2022; Stockwell 2022).

The pedagogic affordances offered by image-based social media platforms such as *Instagram* and *TikTok* are now widely recognised (Lee 2023). *Instagram* posts and reels (non-formal, short videos) are "multimodal ensembles", that is, digital products which combine "a plurality of signs in different modes into a particular configuration to form a coherent arrangement" (Kress 2010 p. 162). They blend photographic and graphic elements, videos, music and sound effects, written and spoken language and elements typical of social media and digital communication, such as emojis, hashtags and geotags (Wagner 2021). Wagner (2021, p. 155) defines *Instagram* as a "multimodal microblogging genre" which requires learners to have multimodal literacy skills because, to benefit from this type of informal ELT materials, they must be able "to analyse each constituent semiotic resource (*working intramodally*) as well as the aggregate whole (*working intermodally*)".

Aslan (2024) has analysed the multimodal composition of micro-celebrity English teachers' reels in one of the few available genre-oriented analyses of *Instagram* for non-formal ELT adopting the conceptual framework of relational pedagogy, which focuses on the relationships between material, social and individual dimensions of learning and meaning-making. He has shown how ELT influencers create opportunities for input noticing and intake and for interaction and participation through "a transductive process from the verbal to visual modes" (Aslan 2024, p. 8) and "intertextuality and resemiotisation" (Aslan 2024, p. 11), that is, by combining carefully selected and culturally meaningful written text, spoken text, body language and images or sounds to make linguistic or cultural concepts understandable and memorable (Bonsignori, Cappelli 2020).

Another affordance for non-formal language learning is the mobility of social media platforms (Lee 2023). *Instagram*'s accounts' creative potential favours learners' engagement and captures their attention beyond the confined spaces of a classroom or the pages of a textbook. They can be accessed from any location where users have Internet access, thus multiplying the

opportunities for exposure to authentic language input (Teng *et al.* 2022). This favours the personalisation of the learning experience, with positive effects on the outcome since the latter depends on the learner's intentionality to engage with the resources the environment makes available (Van Lier 2007).

Autonomy development and motivation increase are, indeed, frequently mentioned in studies investigating teachers' and learners' perceptions of the use of *Instagram* in EFL (Aslan 2024; Gonulal 2019; Kern 2021; Min, Hashim 2022; Zheng *et al.* 2018). Teng *et al.* (2022) point out that *Instagram* provides many opportunities for multimodal social interaction centred on a particular topic or activity offered by the platform. *Instagram* learners-followers can interact with teachers-Instagrammers and other users via private messages, comments, as well as other means such as answering polls or liking content. Such interactions contribute to a collective knowledge-sharing environment that enhances the overall learning experience and leads to the emergence of real virtual communities formed by people sharing similar interests and desires (Erarslan 2019). Interestingly, some participants report being self-conscious and somewhat anxious when interacting on *Instagram*, fearing being judged by more proficient users, but this often leads to more accurate and well-polished language use (Lailiyah, Setiyaningsih 2020; Misnawati *et al.* 2022).

By providing a "window" to observe authentic language use and cultural practices of native speakers and more expert users of the language (Reinhardt 2019, 2020), *Instagram* creates a positive feeling of "upward social comparison", which is a well-known drive for behaviour and choices (Meier, Johnson 2022). Being exposed to what is perceived by learners as "authentic" English is often associated with a perception of "language learning gains and positive affective outcomes" (Lai *et al.* 2018). It provides opportunities for noticing differences between the forms taught in class and those encountered in context, which can enhance incidental learning (Richards 2015). In addition, interacting with other users gives learners the impression that they are having meaningful exchanges in the foreign language, which are more stimulating than those they might have had in artificial formal instruction settings (Lailiyah, Setiyaningsih 2020) and more interesting than the activities and the language they might have found in coursebooks (Nasution 2023; Meirbekov *et al.* 2024).

Most studies focusing on EFL and *Instagram* present teachers' and learners' perceived efficacy of incorporating its use in formal and informal language learning or are discussions of qualitative attitude surveys or structured and semi-structured interviews about exposure to English via social media (Gonulal 2019; Junior 2020; Lailiyah, Setiyaningsih 2020; Mahmud *et al.* 2022; Nasution 2023; Suryanti Tambunan *et al.* 2022). Overall, the surveys show that *Instagram* EFL is perceived as effective in improving all language skills, especially vocabulary and grammar. Many report that *Instagram* helps

them discover new expressions and phrases, learn idioms or lexical items typical of registers other than those represented in classroom material and use humour appropriately (Alotaibi 2023; Rezaie, Chalak 2021). It also promotes learners' creativity and motivation (Nasution 2023). Unfortunately, very few studies present empirical, experimental or quasi-experimental investigations of the actual learning outcome.

Some studies have also highlighted the challenges of using *Instagram* for learning English (Rezaie, Chalak 2021). Some learners see a limitation in the absence of structure and of a syllabus. Moreover, some complain that *Instagram* content is usually not specified for proficiency levels or registers. Therefore, it is difficult to understand whether certain expressions are appropriate for specific contexts of use or the level one sets out to attain. Abbreviations and, in general, the synthetic style used in captions and posts might be detrimental to learners' academic writing skills. If one excludes the most reputable institutional and commercial accounts, there is generally no control over the content published, which might occasionally include spelling, grammar and even pronunciation mistakes (Rezaie, Chalak 2021). Another aspect that might represent a weakness of *Instagram* as a platform for ELT is that communicative skills (e.g. speaking) are not easy to practice. Most accounts help develop receptive abilities (e.g. reading and listening) or present vocabulary and grammar rules, thus potentially resulting in an unbalanced development of the four language skills. Finally, some people have pointed out the possible risks of using mobile devices, including privacy issues, easily getting distracted by too many stimuli and the danger of technology addiction (Metruk 2022).

3. Research questions and methodology

Verifying the actual rather than the perceived impact of non-formal learning is somewhat difficult since, as Sundqvist (2024, p. 6) writes, it "occurs outside educational institutions", so "it obviously cannot be 'controlled' and, thus, poses challenges in terms of methods". Nevertheless, we believe that understanding the features of the sources of such learning might help better interpret possible correlations between self-reported exposure and linguistic development.

No established frameworks for analysing *Instagram*-mediated ELT content could be identified in the existing literature. For this reason, we decided to assume the main dimensions of traditional formal ELT as a heuristic viewpoint, providing a solid descriptive frame of reference. To give structure to our analysis of a still unexplored domain, we decided to organise our investigation focusing on three major categories of foreign language instruction: participants (non-formal, teachers and learners), contents (non-

formal, topics and skills) and methodology and mode of presentation. We then tried to identify and describe the main features of these components in a selection of ELT *Instagram* accounts. In addition, since some studies reported that non-formal resources are felt as less trustworthy than traditional materials (Lai *et al.* 2018; Rezaie, Chalak 2021), we tried to verify whether such features varied across different types of accounts. More specifically, we compared reputable and official institutions' accounts (e.g. @britishcouncilenglishonline) and non-institutional accounts (e.g. @ash_britishenglish) to assess whether users' concerns were justified. For the purpose of the present discussion, we considered "institutional" the accounts of both public institutions such as the British Council or the BBC and of entities such as Cambridge English or Oxford English, which, although having a commercial aim as well, play a vital role in defining the standard of British English usage and can be considered authoritative and trustworthy.

We followed 8 "institutional" accounts and 51 "non-institutional" accounts for 12 weeks. The disparity in number is due to the availability of accounts in the two categories, with private accounts being much more numerous than reputable entities' accounts. We then selected six institutional and 21 non-institutional accounts and qualitatively analysed the posts published between February and April 2024 using the abovementioned three categories. Tables 2 and 3 in the Annexes section provide the list of all the *Instagram* accounts followed and the accounts whose content was included in the study, respectively.

The reasons for exclusion were a) failure to focus on ELT (e.g. @britishcouncil, @stage_door_johnny), b) featuring lessons specifically meant for specific L1 learners and having as the primary language used in the posts a language other than English (e.g. @ english_with_manuela, @antonios.english, both addressing Italian speaking followers), c) not being personal accounts but rather collections of *Instagram* ELT materials from other sources (e.g. @english_squad, @english.ingeneral), d) being prevalently and explicitly for profit and commercial accounts (e.g. @toefl_official, @english_al_fresco), e) the teacher not being a native speaker of English (e.g. @english.with.arezou) and f) the account having less than 80.000 followers (e.g. @lola_speak). These criteria were chosen with the sole purpose of selecting accounts that were comparable in an otherwise extremely vast and diverse offer. Inclusion and exclusion involve neither judgment as to the quality of the accounts excluded nor any specific stance relative to L1 vs. L2 English language teachers or native-speakerism in general. Section 4 describes and discusses the general trends observed and compares them across different types of accounts.

4. Data and discussion

Instagram accounts have a recognisable structure. They typically include the account's username, a profile picture (non-formal, either a photo or a logo that appears next to the username in all posts and interactions), a "bio" in which the account owners introduce themselves or describe the purpose of the account, together with the number of published posts, followers and accounts followed. The profile picture acts as an active link if "Stories" (non-formal, short-lived photos or videos that disappear after 24 hours) are available. Otherwise, it features links to share the profile. The bio can include links to external resources. A series of buttons allows users to follow, message and email the account or see similar profile recommendations. Below the bio, some accounts include "Highlights", collections of stories that are featured permanently. They showcase the best or most important Stories. The main content of an *Instagram* page is the grid of posts. Posts can be photos, videos, carousels (non-formal, multiple images or videos in one post), or IGTV (non-formal, longer videos). Each post can be accompanied by a caption, hashtags and tags of other users. Followers can either scroll the main grid of posts or look only at the available reels. They can usually leave comments. All the accounts analysed allow followers to do so.

Figure 1
Example of an *Instagram* ELT account.

Accounts vary with respect to Highlights. All institutional accounts include several sections, which they use to organise their content into easily accessible categories. Most non-institutional accounts (80%) also include Highlights. Although some categories (e.g. vocabulary, idioms, grammar) are common across all *Instagram* ELT accounts that use this feature, they differ in interesting ways. Non-institutional accounts often include an "about me" category, with stories that explain how they came to teach English or tell followers more about them. The stories most commonly focus on their professional qualifications. Still, they can be very personal too and, occasionally, they introduce private events or non-ELT related aspects of the micro-celebrity teacher's life. Personal information is not found on the institutional accounts, but @cambridgeenglish has a Highlight category dedicated to "the team". In this collection of Stories, the people behind the account are introduced as they answer a question about their favourite resources for learning English. These posts create a personal connection between followers and the "accounts". On the other hand, reputable entities such as Cambridge University Press do not need to prove their undisputed qualities as sources of ELT materials, while private individuals do, which explains the need to provide qualifications. Nevertheless, quite interestingly, Cambridge English and Oxford English accounts also include Highlights about the cities and their famous Universities. At the same time, this contributes to the narrative of these historical institutions as the home of "true English" and provides support for the quality of the materials offered via social media platforms.

By including stories of their academic and professional career, *Instagram* micro-celebrity teachers also become positive models and trigger upward envy, possibly motivating learners to follow a similar path towards mastering English. This is especially true of non-native speakers of English, whose *Instagram* teaching histories thus become professional success stories.

Except for @bbclearningenglish, all institutional accounts include self-promotional Highlights featuring courses or language programmes offered or professional opportunities (e.g. internship positions at Pearson). Some non-institutional accounts also promote books, one-to-one tutoring or other pay services. The commercial and self-promotional nature of these private accounts is quite marked.

Overall, institutional accounts publish more static posts than reels, whereas the opposite is true for non-institutional accounts. Some are exceptions: @pearsonlanguages predominantly posts videos and @english_vocabulary and @idiom.land, which specialise in vocabulary, only post static images. Figure 2 presents the distribution of reels and static posts in the accounts included in the study.
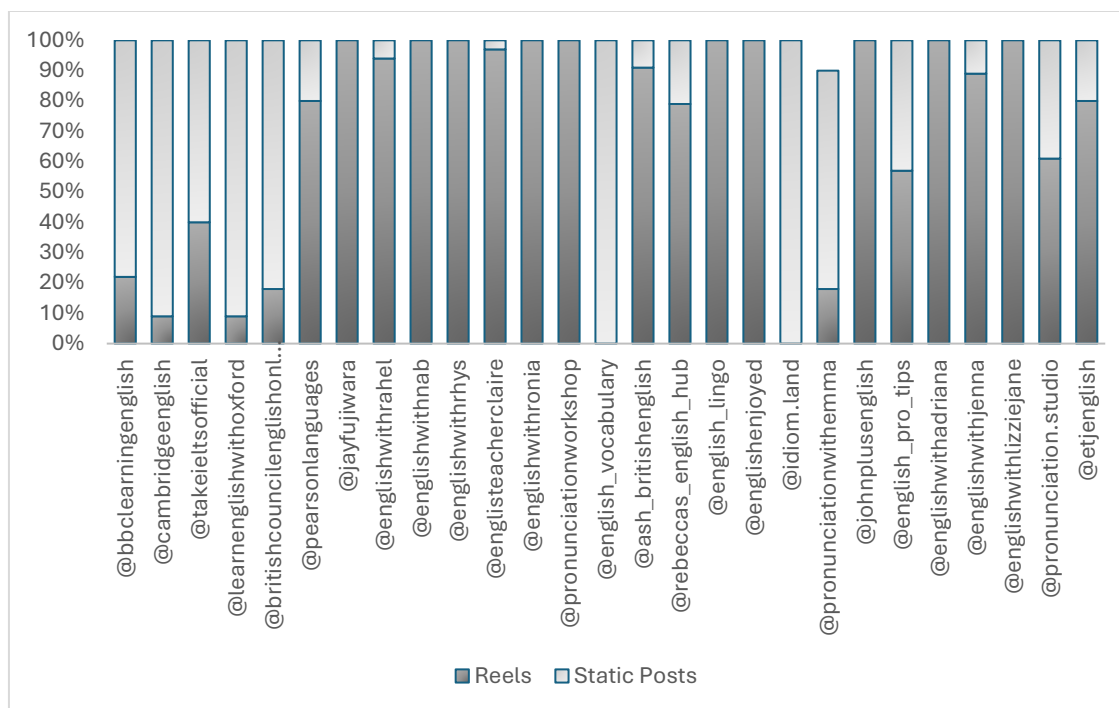
Figure 2.
Frequency of reels vs. static posts across accounts.

4.1. Contents: Topics and skills

The quantity and the contents of the posts vary greatly across accounts. Some add content daily; others post every other day, weekly, or less frequently. However, the qualitative analysis of the accounts included in the study has revealed some trends in terms of topics. Most posts focus on one concept, be it one or more lexical items, a grammar rule, the pronunciation of a word or a specific sound, or some cultural aspect of the English-speaking world. Accordingly, the content published between February and April 2024 can be classified into seven major categories: *vocabulary*, *grammar*, *pronunciation*, *use of English*, *culture*, *tips* and *other*. The macro-categories *culture*, *use of English* and *other* are quite heterogeneous. We classified posts as *culture* if they focused on the culture of English-speaking countries, ranging from books and literature to movies, public holidays and famous people's quotes. We tagged posts as *use of English* if they focused on issues of register, contextual appropriacy, differences in the varieties and other aspects related to the pragmatics of communication in English as a foreign language. Finally, the *other* category includes posts of a very diverse nature, ranging from self-promotion to personal posts, to the occasional post meant to stimulate discussion or written production (non-formal, on complex skills), to the importance of English for career purposes. Figures 3 and 4 present the distribution of topics across accounts.
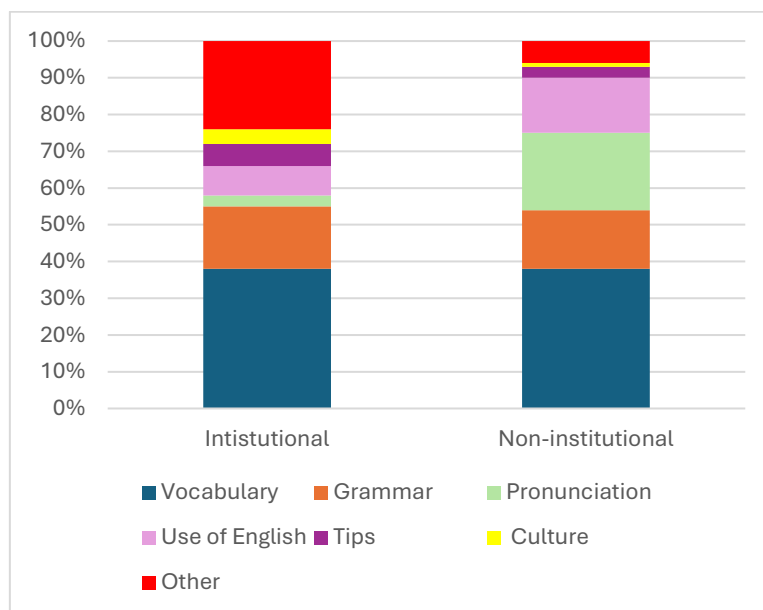
Figure 3.
Average distribution of different types of content within the two types of accounts.
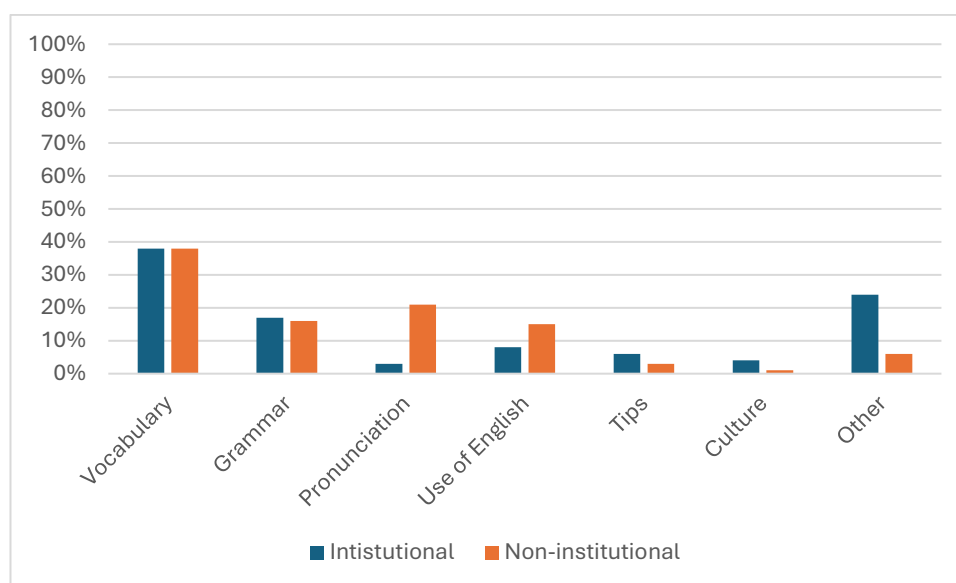
Figure 4.
Comparison of average content distribution between accounts.

Both institutional and non-institutional accounts provide abundant lexical and morpho-syntactic input: vocabulary and grammar are equally represented in the two groups. Vocabulary posts introduce new words or phrases. Static posts usually present a word, an expression or an idiom associated with an image of the referent or with a definition and an example sentence. The presentation of vocabulary in reels has very diverse formats, ranging from people using vocabulary in context or explicitly explaining the meaning of a certain word or

idioms. Many posts present and explain idiomatic expressions, proverbs or the cultural nuances associated with a lexical item, sometimes in the form of more or less interactive quizzes or games and puzzles. Figure 5 offers examples.

Figure 5
Vocabulary posts by (from the left) @englishwithrys (first column), @english_lingo (second column), @learningenglishwithoxford (third column, top), @bbclearningenglish (third column, bottom), @cambridgeenglish (fourth column, top), @idiom.land (fourth column, bottom).

Similarly, grammar posts can include explanations of grammar rules, debunk common mistakes and show how to use specific grammatical structures correctly. As in the case of vocabulary posts, they can be reels in which somebody explains, sometimes simulating a classroom setting, or, more frequently, static posts in the form of multiple-choice quizzes.

This shows that, despite the innovative platform, the type of input provided aligns with that generally offered by traditional textbooks and classroom materials. In the case of well-established entities' accounts, the reason for this might also be the nature of the materials produced. Creating static posts is faster and easier than making videos with professional actors. Still, they have the limitations of any other written document when it comes to conveying information and they are much more suitable for activities that only involve a receptive or structured fruition of the input. *Instagram*, however, offers many affordances to provide a diversified presentation of the input. Non-institutional accounts seem more committed to exploiting its multimodal potential and, therefore, more productive when designing creative multimodal content. Yet, they also opt for more explicit vocabulary and grammar lessons than other input types.

Pronunciation posts are typically in video format. A speaker explains how to pronounce complex sounds or words with examples. In one case, @pronunciation.studio, posts are very technical and even show the articulatory

position of sounds.[4] Most frequently, though, posts in this category adopt a folk linguistics approach and pronunciation is rendered in written format with letter combinations typically pronounced like those to be learnt, as in Figure 6.

Figure 6
Speak fast (@englishwithnab).

Tips usually focus on strategies for effective language learning, including advice on study techniques and maintaining motivation, but also recommendations for apps and external resources that can (allegedly) maximise learning. These posts are usually sponsored (in non-institutional accounts) or collaborations (in institutional accounts).

Use of English and Culture posts are very interesting categories. The former includes information about register-appropriate expressions or situational language (e.g. "how to" posts or simple dialogues "at the airport", "at the restaurant", etc.). They also offer alternatives to express basic concepts in a diversified way. Moreover, they include posts about the differences between English standards. Culture posts range from short reels about popular destinations (e.g. *Welcome to London*, @takeieltsofficial) to micro-documentaries about various topics (e.g. *Shakespeare and the Globe Theatre*, *It's Tea O'Clock*, @britishcouncilenglishonline; *What is Britain*, @ash_britishenglish) and videos where seminal literary works are narrated (e.g. *The Story of Hamlet*, @britishcouncilenglishonline).

Institutional and non-institutional accounts differ in terms of the average distribution of these categories of posts. More specifically, pronunciation and

---

[4] For an example of the pronunciation of the voiced alveolar approximant sound of English /r/ see https://www.instagram.com/p/C6VkSlkICUp/ (2.4.2024).

issues related to the use of English in context are more widespread in non-institutional feeds (and this is in line with the production of more reels). In contrast, posts related to the English-speaking world's culture and tips for improving specific skills or developing specific abilities are more common in institutional accounts. The latter also publish posts on a greater variety of issues of a very diverse nature, including international certificates, testimonials by learners around the globe, interviews with celebrities and inspirational and humorous posts that are not inherently geared towards ELT. These differences can be easily explained by the fact that some of these accounts are connected to international examination boards or publishing businesses. Therefore, they provide followers with useful information about their particular "services". On the other hand, accounts such as @bbclearningenglish or @britishcouncilenglishonline also have the "mission" to represent a lingua-culture and teaching English becomes just one of the components of the cultural world they present and promote. It should be mentioned that some intra-group variability is present, as Table 1 shows. A more thorough and extended quantitative investigation would be necessary to reach more generalisable conclusions.

Instagram account	Vocab.	Gram.	Pron.	Use of Eng.	Tips	Cult.	Other
Institutional accounts							
@bbclearningenglish	56%	26%	4%	5%	0%	0%	9%
@cambridgeenglish	34%	13%	5%	11%	0%	0%	38%
@takeieltsofficial	26%	14%	2%	7%	11%	7%	33%
@learnenglishwithoxford	55%	23%	2%	6%	0%	2%	13%
@britishcouncilenglishonline	42%	26%	0%	9%	5%	10%	8%
@pearsonlanguages	13%	2%	8%	8%	19%	5%	46%
Non-institutional accounts							
@jayfujiwara	14%	43%	14%	29%	0%	0%	0%
@englishwithrahel	49%	3%	3%	34%	11%	0%	0%
@englishwithnab	38%	23%	8%	31%	0%	0%	0%
@englishwithrhys	43%	36%	11%	7%	4%	0%	0%
@englisteacherclaire	58%	24%	3%	12%	3%	0%	0%
@englishwithronia	31%	48%	10%	10%	1%	0%	0%
@pronunciationworkshop	0%	0%	100%	0%	0%	0%	0%
@english_vocabulary	100%	0%	0%	0%	0%	0%	0%
@ash_britishenglish	67%	10%	6%	8%	0%	4%	6%
@rebeccas_english_hub	45%	15%	18%	7%	2%	2%	12%
@english_lingo	71%	14%	2%	12%	2%	0%	0%
@englishenjoyed	26%	5%	32%	26%	0%	5%	5%
@idiom.land	100%	0%	0%	0%	0%	0%	0%
@pronunciationwithemma	0%	0%	64%	0%	0%	0%	36%
@johnplusenglish	87%	0%	9%	4%	0%	0%	0%
@english_pro_tips	45%	8%	8%	4%	0%	0%	34%
@englishwithadriana	17%	17%	8%	33%	17%	0%	8%
@englishwithjenna	22%	33%	33%	11%	0%	0%	0%
@englishwithlizziejane	75%	13%	0%	13%	0%	0%	0%
@pronunciation.studio	0%	0%	94%	0%	0%	0%	6%
@etjenglish	0%	0%	60%	20%	20%	0%	0%

Table 1
Distribution of content type across accounts.

Four out of six institutional accounts promote British English and British culture in general, while @takeieltsofficial and @pearsonlanguages insist more on the instrumental and international nature of English as a global language. However, whereas the former does so by simply posting about differences between British and American English that might be useful for the IELTS exam, the latter's feed features many posts by non-native speakers of English and about the role of English in global mobility and relocation, thus adopting a general view of the language as a lingua franca. The non-institutional accounts generally refer to either British English or American English standards and specify which one. Some have as a selling point the promise to help followers learn to speak "true" British English (e.g. @ash_britishenglish) or "as a real American" (e.g. @englishwithnab).

Occasionally, some posts focus on specific skills, such as reading or listening comprehension or provide prompts for writing and speaking. This happens mostly on institutional accounts' feeds, especially those connected with international examinations. These accounts also offer most tips, usually about performing better in exams (e.g. IELTS, Pearson English Test). However, some non-institutional accounts also provide tips on expanding one's knowledge of English through movies or books or through some third-party app that sponsors them. As is evident, the nature of the content posted is influenced by the nature of the platform. It is undoubtedly easier to provide content that does not require much feedback (e.g. the meaning of a word, a grammar rule) or for which feedback is easy to give (e.g. multiple-choice quizzes about specific aspects). Providing opportunities for speaking or writing requires interaction with an interlocutor, which is certainly more challenging to implement on a platform that primarily relies on asynchronous content production and fruition.

4.2. Participants: Teachers and learners

For the present study, we considered the content source to be *teachers*, be it a real, identifiable person or an Instagram account, and followers to be the *learners*. We analysed the type of content delivery and the way in which learners-followers receive and interact with it, with the teacher account, and with each other. To investigate participants' interaction, we looked at images/videos and their captions and comments.

Knowledge can be passed on in different ways in static posts. Content is presented via single images or carousels. It is often graphically more appealing and creative than content in standard printed materials. It is also understandably more synthetic in terms of the amount of text provided. However, *Instagram* static posts and reference books or printed reference materials often share the same type of information delivery: rules, expressions and words are presented, paraphrased, exemplified and illustrated when appropriate.

Reels exploit the full potential of multimodality. Post-production editing allows teachers to integrate their delivery with additional tools for improving explanations. This potentially results in more effective teaching and better learning. Reels highlighting common mistakes or criticising other *Instagram* teachers' posts which contain inaccuracies represent an interesting product. They are only marginally geared towards ELT. On the one hand, these videos aim to shed light on foreign speakers' frequent errors and, at the same time, they serve as a platform to showcase the account owner's professionalism. By presenting themselves as reliable and capable teachers, creators aim to establish credibility within the *Instagram* ELT community.

Different *Instagram* teacher types can be identified, which we chose to label as *explainers*, *enactors* and *popularisers* based on their preferred strategies for input presentation. *Explainers* sometimes reproduce traditional language teaching forms and settings, talking directly to the followers, explaining rules, exemplifying concepts, practising pronunciation, etc. Explanations can sometimes be delivered at meaningful locations, such as the aquarium to introduce the difference between *fish* and *fishes* (Figure 8). Static posts tend to fall into this category and they are occasionally used to teach complex skills, such as reading or writing. Figures 7 and 8 exemplify this type of teaching action.

Figure 7.
Static post by @britishcouncilenglishonline explaining how to write a text.

Figure 8
Instagram teachers explaining (top, left to right: @jaifujiwara, @englisteacherclaire, @pronunciationworkshop, @ash_britishenglish; bottom: @englishwithrahel).

Enactors present concepts via short sketches in which dialogues are performed as if they were learners of English, examiners, or language users in different situations and from different varieties. Followers can learn by example, inferring relevant structures and words with the help of subtitling or other verbal and non-verbal graphic elements. Figures 9-11 show some examples.

Finally, some videos present different kinds of information (usually cultural knowledge) as short documentaries. We have called these types of *Instagram* teachers *popularisers*. Some static posts have the same function. For example, they might report a quote by a famous person (e.g. Bob Marley's "Life is stronger than death", @bbclearningenglis). Through the caption, followers will learn about the Anglophone World's culture or language (e.g. how to form comparatives). Figure 12 features three examples from @bbclearningenglish.
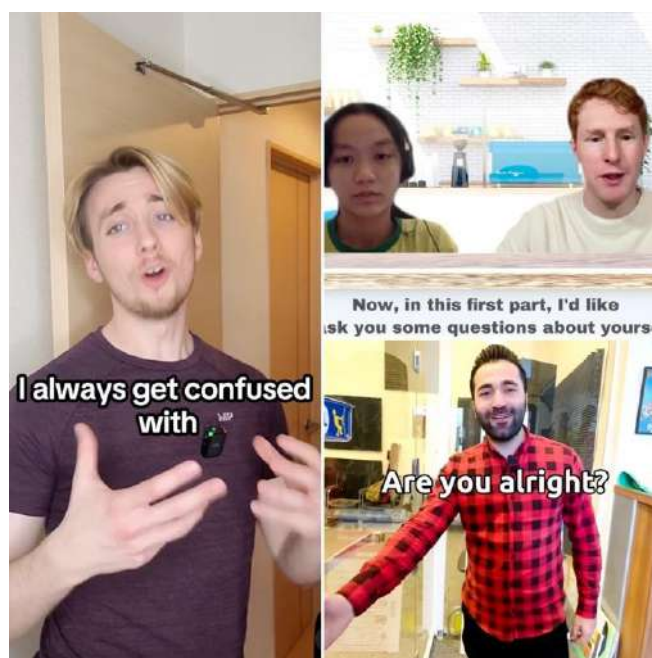
Figure 9
@ash_britishenglish pretending to be a learner confused by specific grammar rules (left),
@englishwithrahel enacting welcoming a colleague (right, bottom), @english_pro_tips
enacting the beginning of the speaking test for a language certificate exam (right, top).

Figure 10.
@englishwithrahel enacting a dialogue at a restaurant with expressions for different
proficiency levels (left), @english_lingo enacting a dialogue with his daughter Sophia
meant to teach the concept of littering (right, top), @englishwithnab enacting a "variety
battle" with his double to present words that differ in BrE and AmE.
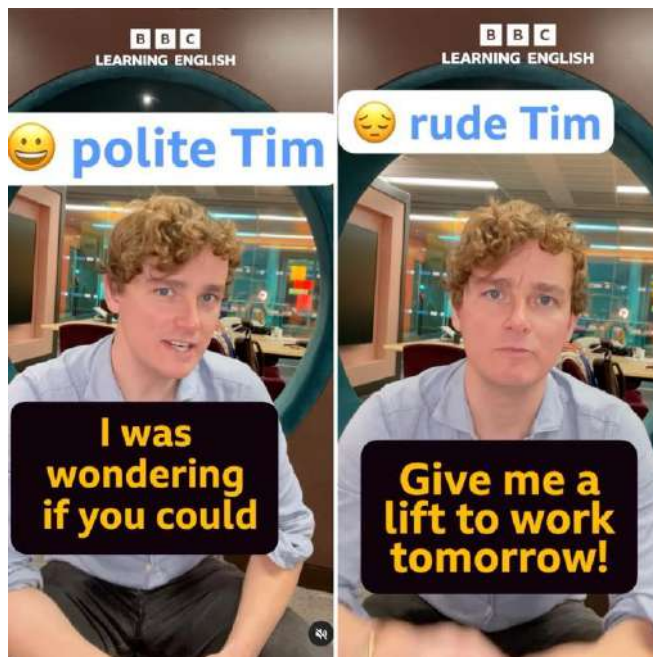
Figure 11
@bbclearningenglish with Tim enacting dialogues to exemplify
appropriate ways to make a request.

Figure 12.
Bob Marley's quote containing the comparative form of the adjective *strong* (top), *Hamlet*
narrated in a short video (left, bottom), a short video about different tea-drinking habits
(right, bottom).

The analysis of captions accompanying static posts reveals that the teacher's role is carried out through several common strategies across accounts. Captions can be *descriptive* and provide a title for the image or video (e.g. "Present Perfect Tense" in association with a video in which the rules and use of the tense are explained) or include the transcription of the spoken component of a video. They can be *explanatory* and provide further insight into the topic of the picture or video to help learners understand better (see Figure 13).
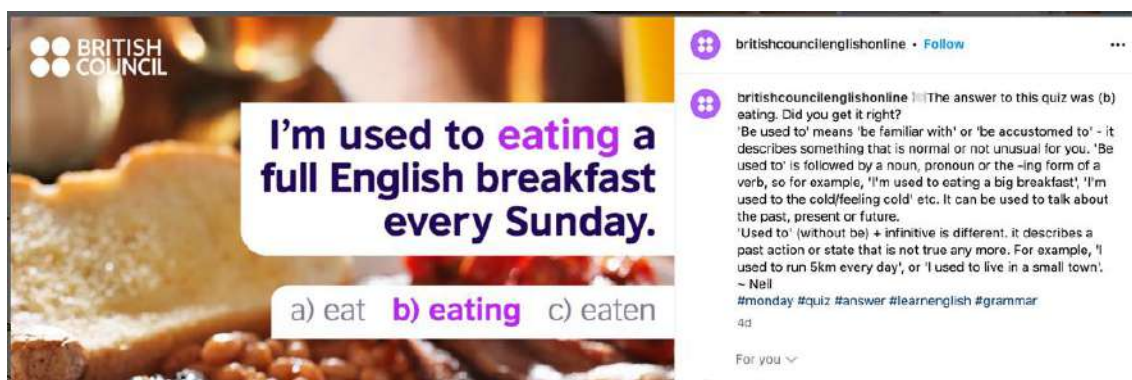
Figure 13
Explanatory caption which explains the correct answer and the meaning of the expressions in the post (by @britishcouncilenglishonline).

Captions can include a question to assess whether the rule presented in the image has been learnt. This can be a multiple-choice or open question. In such a case, they are *complementary* to the visual component because they integrate and complete the teaching act. Most captions, though, are *interactive* or at least include an element of interactivity. Interactive captions include forms of direct address to followers. They can explain something and then ask them to leave a comment that answers a question, offers a personal opinion or anecdote, or expands on the post's topic. Or they can be complementary to the post and directly involve the reader, thus forming a more complex multimodal ensemble and micro-lesson. Figure 14 is an example. "Kate" discusses how you can improve English by watching movies and the caption by "Neil" prompts followers to share their personal experiences with English movies and then directs them towards the clip and other external resources.

Figure 14
Example of complementary + interactive caption (@britishcouncilenglishonline).

Comments can be compared to a virtual classroom space where interaction between the participants in the learning event happens. The analysis of a small selection of posts (non-formal 10 per account) revealed that the majority of the first-level comments (non-formal, those produced as a direct reply to the post and not to other users' comments) are learners-followers' answers to questions asked either in the visual / video component of the post or in the caption (80%), followed by expressions of appreciation for the content (8%), questions for clarification (5%), or criticism (2%). Some comments (5%) are not strictly related to the pedagogic nature of the post and either focus on the people on video (e.g. commenting on their appearance) or are spam comments advertising products or promoting some propaganda.

Second-level comments generally include feedback (90%), further explanations and details (7%), appreciation and support (2%) and some non-relevant replies, usually spam (1%). Peer-to-peer feedback is more frequent than teacher-learner feedback, especially once a few comments have been posted or after some time since publishing the "lesson". Very often, remarks and feedback come from other ELT accounts, usually less influential ones, who try this way to make themselves known and increase the number of their followers. The most influential accounts usually only provide feedback for the first few comments and then move on. Accounts related to businesses frequently use feedback to promote their services and direct followers towards their courses or other external resources. Overall, in the accounts analysed, peer interaction seems to be more productive and often more useful and to the point than interaction with the teachers.

4.3. Methodology and modes of presentation

All posts are multimodal ensembles, combining verbal and non-verbal elements. This is not a novelty in ELT per se. However, *Instagram* ELT exploits the integration of multiple semiotic resources to maximise input noticing. Moreover, the platform's technical features shape the input presentation by forcing creators to design short and effective "bite-sized language teaching" materials (Aslan 2024).

Static posts exploit different font combinations, colour patterns and the juxtaposition of verbal and non-verbal elements to draw learners' attention to what each post wants to teach. In this sense, they do not differ much from flashcards and other traditional multimodal strategies adopted in the language classroom or printed materials. Some posts focusing on vocabulary items, expressions or idioms pair the lexical label with a picture of the referent or with an image representative of its meaning or the context in which it would be used. The major difference between traditional printed materials or flashcards and *Instagram* posts is that the latter usually present one concept at a time, out of context and simultaneously provides information about different aspects (e.g. pronunciation, spelling, meaning, phonetic transcription, an example of its usage and referent's images). Each piece of knowledge becomes, therefore, easily noticeable and accessible and each post is a micro-lesson in itself. The most interesting aspect of static ELT posts seems to be the interaction with captions. They often enter into a complementary relationship with the visual component and guide the follower along a specific learning path by asking them to use the rule or expressions in the comments, find out more about it by following a link or providing more examples of similar words, phrases or contexts of use. Many static posts, moreover, are slides containing multiple choice questions or "language challenges" organised as a series of slides in a carrousel, with the correct answers provided in the caption or as feedback to comments.

As can be expected, reels are the type of posts that exploit the platform's affordances the most. Videos by explainers, enactors and popularisers share some common features, but they also have unique modes of presentation. *Instagram* teachers can be on camera or use voice-over. Most non-institutional accounts belong to micro-celebrity teachers who are the protagonists of their reels. Institutional accounts only feature a few reels with actual people, generally enactors or popularisers. In line with Aslan's (2024) findings, most reels were found to exploit transduction and transformation: meaning construal is the result of the presentation of concepts through different alternative semiotic resources encoding the same type of information (e.g. written and spoken language, gestures and images) and changes within the same mode (e.g. using prose and poetry to present the same idea).

Explainers' reels exist in very diverse formats. In their most basic form, the *Instagram* teacher is on camera and talks directly to followers. As mentioned in section 4.2, some reproduce a traditional teaching setting, as in Figure 15.

Figure 15
@jayfujiwara teaching form of lexical verbs after auxiliaries.

However, even in these cases, they can be quite creative, inventing songs and even posting new reels that join the original video and followers' videos, as in Figure 16. Here, @jayfujiwara is teaching the use of auxiliaries and a follower is playing his soundtrack for the original reel.[5]

---

[5] @jayfujiwara's original reel: https://tinyurl.com/jayfujiwara1 (2.4.2024) and the second reel with the follower's soundtrack video https://tinyurl.com/jayfujiwara2 (2.4.2024).
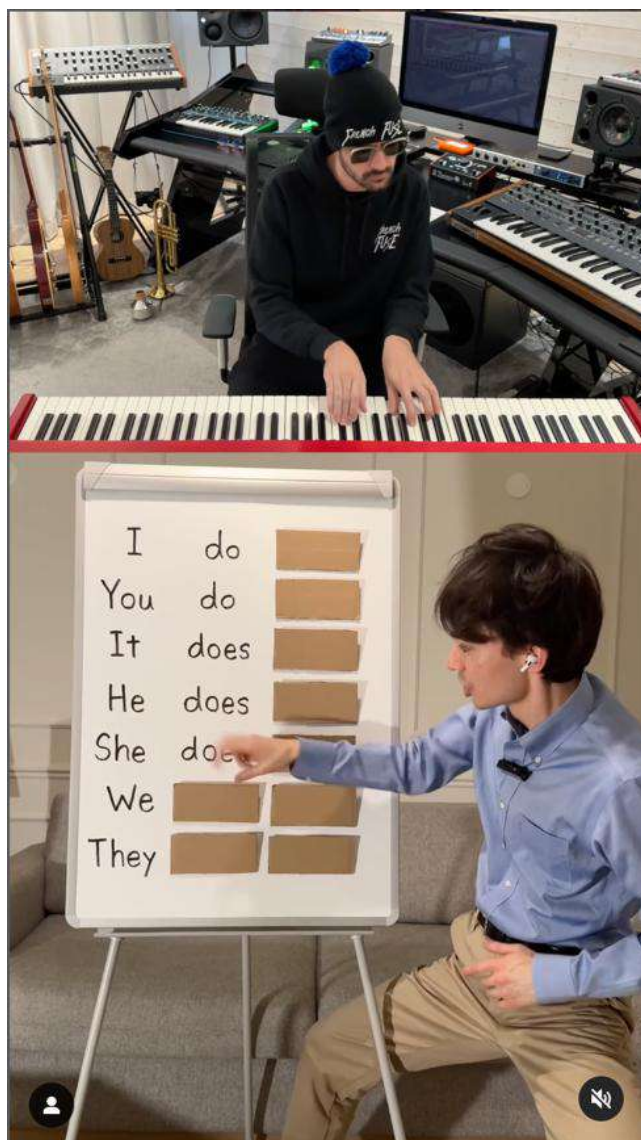
Figure 16
@jayfujiwara and @frenchfusemusic's reel.

Besides the video recording of the teacher speaking and often using iconic gestures to increase conceptual accessibility (Bonsignori, Cappelli 2020), reels may include graphic elements and realia to highlight important concepts or enhance and integrate the speaker's words and explanations. This is especially common with vocabulary lessons, where a combination of spoken and written language is often associated with images of the referents. Other graphic elements are also used, such as a flashing stop sign to mark wrong usage and phonetic transcriptions of specific sounds (either actual IPA symbols or folk linguistics transcriptions), as in the examples shown in Figure 17.
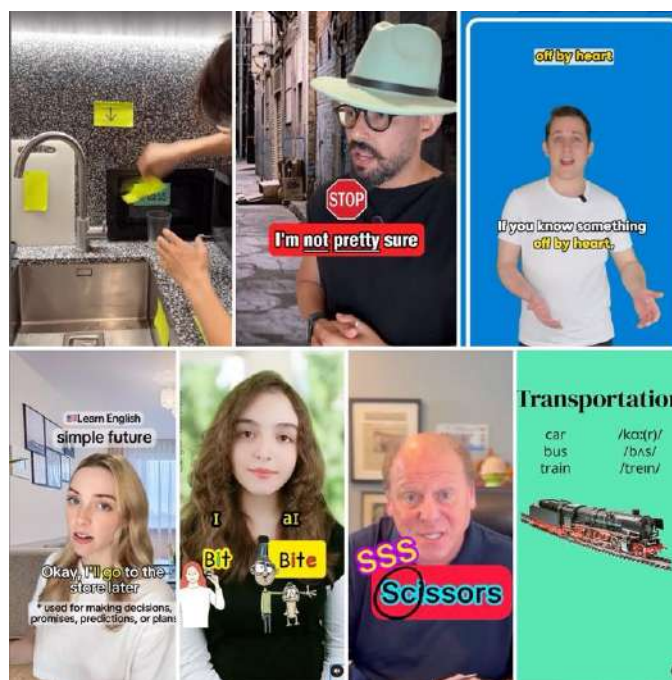
Figure 17
Use of multiple semiotic resources and transduction in explainers' reels (top, left to right:
@jayfujiwara, @englishwithnab, @englishwithrhys; bottom, left to right:
@englishteacherclaire, @englishwithronia, @pronunciationworkshop,
@learningenglishwithoxford).

These videos and static posts reproduce the Presentation – Practice –
Production (PPP) methodology, although only in a partial form. The
presentation phase is predominant. Posts including quizzes or polls in the
visual component or captions can be equated to the practice stage of a lesson:
learners try their hands at the newly acquired rules or expressions.
Occasionally, production is prompted as an invitation to write something in the
comments or to post videos with attempts to use what has just been learnt.
However, this is rare and what is generally lacking is the possibility of
receiving thorough and constructive feedback, especially for spoken
production. Some reels simulate practice/production by allowing some time for
followers to complete a sentence or pronounce a word and then the *Instagram*
teacher provides the "solution". Alternatively, the teacher plays a part and then
the video allows the follower some time to read the following turn.[6] These
cases represent an attempt to create what Aslan (2024, p. 7) defines as "a
simulation-based pattern of participation framework". Institutional accounts
post more content meant to stimulate production than non-institutional
accounts, probably because speaking and writing skills play an important part

---

[6] Examples can be found by following these links: https://tinyurl.com/cambridgeenglish1 (2.4.2024);
https://tinyurl.com/pearsonlanguages1 (2.4.2024); https://tinyurl.com/cambridgeenglish2 (2.4.2024);
https://tinyurl.com/tofluency1 (2.4.2024); https://tinyurl.com/ashbritishenglish3 (2.4.2024).

in the international certificate exams they are associated with. This might indicate that they are not meant as an independent source of informal EFL but as a supplement to traditional preparation courses. Some posts stimulate meta-linguistic or strategic reflection on the learning process (see Figure 18). However, although the caption reads, "Are you a grammar enthusiast? Or perhaps you're an avid 📔 word collector? 😊 Maybe you're a bit of both – but which is more important when you're trying to learn and use a new language? ✍️ Join the debate!", no feedback is provided in the comments.

Figure 18
Post by @bbclearningenglish meant to prompt written argumentative text production.

Enactors act as if in a specific situation or occasionally simulate dialogue with other people by "reduplicating" their image. The dialogues may aim to teach certain words or structures or show what is appropriate. Like explainers' videos, all reels have subtitles and the relevant elements are usually highlighted in different colours or font sizes to ensure that learners understand the aim of the micro-lesson. They are often humorous. Body language plays an especially important role in these short sketches.[7] Enactors' reels are difficult to classify according to the traditional teaching models because followers are expected to learn by example. They represent an inductive, usage-based form of teaching, where rules must be inferred from examples of language in action. There is no space for practice and production is only occasionally prompted by captions.

---

[7] Some examples of enactor's reels are the following: https://tinyurl.com/englishteacherclaire2 (2.4.2024); https://tinyurl.com/rebeccasenglishhub1 (2.4.2024); https://tinyurl.com/englishlingo1 (2.4.2024); https://tinyurl.com/bbclearningenglish2 (2.4.2024).

Overall, they share the transductive nature of explainers' videos and their use of resemiotisation, as verbal to visual modes are intertwined or used as alternative but parallel means to establish "a form-meaning relationship between the linguistic unit and its visual representation" (Aslan 2024, p.8).

Popularisers' reels are uncommon and, in our sample, only available on institutional ELT accounts, except for a post about Britain by @ash_britishenglish. They are, nevertheless, quite interesting because they are primarily meant to disseminate cultural knowledge and, in a way, build the brand popularity of English and the Anglophone culture. Popularisation videos often involve intertextuality (non-formal, connecting specific teaching content with other culturally relevant texts). Thus, a video about Shakespearean expressions still in use today provides the opportunity to learn about the great playwright,[8] while another video discusses video games and the subtitles allow followers to learn relevant vocabulary and practice listening comprehension.[9] Other reels offer a narration of the plot of famous literary works such as Shakespeare's most famous plays.[10] These reels, like others by explainers that include clips of famous movies or TV series exemplifying the use of certain expressions, present contexts in which language is used more authentically and which can, therefore, enhance the learning process significantly and support motivation.

5. Concluding remarks

Formal instruction of language follows a structured syllabus, usually relies on a specific methodology (e.g. Presentation, Practice, Production – PPP) and reliable reference material (e.g. textbooks, grammar books) and aims at developing the four skills in a balanced way. Moreover, it involves interaction with teachers and peers at all stages of the teaching/learning process. Feedback and formal assessment are typical measures of learners' progress. In other words, "traditional" settings offer a structured learning environment with teacher guidance. Learning English via *Instagram* differs from this model in many ways, the most evident being the opportunity for individuals to explore content based on personal interests whenever they wish and for as long as they like and the lack of a structured learning path design (non-formal, sequences of lessons aimed at gradually building learners' proficiency). Each post constitutes a micro-lesson and its graphically pleasant or entertaining nature has the potential to capture the attention of learners effectively, as surveys seem to indicate.

---

[8] https://tinyurl.com/learningenglishwithoxford (2.4.2024).
[9] https://tinyurl.com/britishcouncilenglishoinline3 (2.4.2024).
[10] https://tinyurl.com/britishcouncilenglishonline4 (2.4.2024).

In terms of contents and topics, *Instagram* EFL materials are predominantly dedicated to vocabulary and grammar, that is, to knowledge of useful expressions, lexical labels and rules, presented through an approach resembling the PPP model, but in which opportunities for production are widely lacking (unless followers join pay courses or one-to-one tutoring).

Many posts are more reminiscent of situational rather than communicative language teaching methodologies. As for participants, even when competent, engaging and reliable, *Instagram* teachers frequently fail to consistently provide the most crucial element in the educational relationship, which is feedback. Although feedback can be provided in the comments, it tends to be lacking in most of our data. In contrast, peer-to-peer feedback appears to be the most valuable form of assessment available on the platform and often the only one. It should be noted that, since only native speakers of English were included in this analysis, no mistakes could be found in the content proposed. Nevertheless, some explanations in non-institutional accounts were quite naïve and more about folk linguistics than the result of meta-linguistic awareness.[11] For this reason, although a valuable and interesting platform for learning vocabulary, useful expressions, improving listening and reading abilities and learning (or rather revising) grammar rules, the platform does not appear to provide sufficient support for a balanced development of the four skills. Rather, it seems to be an ideal addition to traditional foreign language instruction or a good space for expanding and maintaining skills acquired more traditionally.

The modes of presentation of EFL content are the most interesting feature of *Instagram* ELT since teachers creatively exploit the platform's affordances, enhancing input noticeability and potentially keeping learners' attention. Moreover, by featuring authentic material and offering opportunities to connect with native speakers of English and other users worldwide, they raise interest in the Anglophone lingua-culture and increase motivation to learn English. Institutional accounts seem to be especially invested in promoting the cultural side of ELT, which is the major difference we could observe between the two types of accounts. Overall, institutional accounts appear to adopt a more traditional approach to lexical and morpho-syntactic content presentation and to avoid folk linguistics in favour of well-researched explanations and illustrations. They also provide more opportunities for global skill development, awareness of language learning strategies and knowledge of the English-speaking world. Our investigation did not, however, reveal any significant difference in the reliability and accuracy of the teaching imparted by the institutional and non-institutional accounts included in the sample.

---

[11] An example is the explanation of the present perfect vs. simple past usage in American English: https://tinyurl.com/englishwithjenna (2.4.2024).

The present study has the limitations of most qualitative studies carried out on small data sets. Nevertheless, it can offer an operational framework for further research into non-formal EFL learning and social media. Each of the identified components should be further investigated qualitatively and quantitatively to understand better the input offered via this extramural exposure to English. While the present study provides insights into the generic features of *Instagram* ELT accounts, it is important to acknowledge that we were unable to thoroughly discuss the linguistic input provided in greater detail (e.g. what morpho-syntactic aspects are most commonly presented, the ratio of academic vs informal vocabulary, which semantic fields are privileged, etc.), since it would have exceeded the limits of the present article. Despite this, we believe that such an overview would yield significant contributions to the field. Future research should, therefore, focus on the nature and quality of the input offered on the platform, as such investigation could provide a more comprehensive understanding of the implications for students' learning outcomes. Moreover, future empirical studies could focus on assessing the impact of different types of *Instagram* material. The analysis presented in this study will hopefully offer a frame of reference for isolating the role of multiple aspects and variables in English language learning through *Instagram*.

Bionotes:

Gloria Cappelli is Associate Professor of English Language and Linguistics at the University of Pisa. Her research focuses on semantics and pragmatics, ESP and multimodal discourse and second and foreign language acquisition in learners with and without dyslexia. She has published in national and international journals and collected volumes. She has authored and edited books on English verbs of cognitive attitude, lexical semantics, tourism communication and the impact of dyslexia on language skills.

Nicoletta Simi is a Research Fellow (RTDa) at the Department of Philology, Literature and Linguistics at the University of Pisa. Before her current post, she was a Postdoctoral Research Fellow at the Department of Psychology of the University of Tübingen, Germany. She holds a PhD in English linguistics from the University of Pisa. Her interests involve the development of a comprehensive understanding of key phenomena at stake in reading comprehension processes. She gives further attention to the underlying general processes such as cognitive processing speed, working memory and motivational / anxiety aspects of language use.

Authors' addresses: gloria.cappelli@unipi.it; nicoletta.simi@unipi.it

References

Alotaibi M.S. 2023, *Exploring the Impact of Using Instagram to Develop Saudi EFL Learners' Vocabulary*, in "Journal Corner of Education, Linguistics, and Literature" 2 [4], pp. 280-294.

Aslan E. 2024, *Bite-sized language teaching in the digital wild: Relational pedagogy and micro-celebrity English teachers on Instagram*, in "System" 121, 103238, n.p.

Barrot J.S. 2022, *Social media as a language learning environment: a systematic review of the literature (2008-2019)*, in "Computer assisted language learning" 35 [9], pp. 2534-2562.

Bonsignori V. and Cappelli G. 2020, *Specialized and culture-bound knowledge dissemination through spoken tourism discourse: multimodal strategies in guided tours and documentaries*, in "Lingue e Linguaggi" 40, pp. 213-239.

Canani M. and Zulli T. 2022, *ELT and Social Media: Integrating TikTok into Class Practice*, in Abbamonte L. and Cavaliere F. (eds.), *Updating Discourse/s on Method/s*, "mediAzioni" 34, pp. A165-A183.

Crompton H. and Burke D. 2018, *The use of mobile learning in higher education: A systematic review*, in "Computers & Education" 123, pp. 53-64.

Dos Santos L.M. 2022, *Online learning after the COVID-19 pandemic: Learners' motivations*, in "Frontiers in Education" 7, 879091, n.p.

Dressman M. 2020, *Introduction*, in Dressman M. and Sadler R.W. (eds.), *The Handbook of Informal Language Learning*, Wiley Blackwell, Hoboken, NJ, pp. 1-12.

Erarslan A. 2019, *Instagram as an education platform for EFL learners*, in "TOJET: Turkish Online Journal of Education and Technology" 18, pp. 54-69.

Gonulal T. 2019, *The use of Instagram as a mobile-assisted language learning tool*, in "Contemporary Educational Technology" 10 [3], pp. 309-323.

Guo X. and Lee J.S. 2023, *A systematic review of Informal Digital Learning of English: An ecological systems theory perspective*, in "System", 103097, n.p.

Junior, R. C. G. 2020, *Instanarratives: Stories of foreign language learning on Instagram*, "System" 94, 102330, n.p.

Kamasak R., Özbilgin M., Atay D. and Kar A. 2021, *The effectiveness of mobile-assisted language learning (MALL): A review of the extant literature*, in Moura A.S., Cordeiro M.N.D.S. and Reis P. (eds.), *Handbook of research on determining the reliability of online assessment and distance learning*, IG Gobal, London, pp. 194-212.

Karakaya K. and Bozkurt A. 2022, *Mobile-assisted language learning (MALL) research trends and patterns through bibliometric analysis: Empowering language learners through ubiquitous educational technologies*, in "System" 110, 102925, n.p.

Kern R. 2021, *Twenty-five years of digital literacies in CALL*, in "Language Learning and Technology" 25 [3], pp. 132-150.

Kress G. 2010, *Multimodality: A Social Semiotic Approach to Contemporary Communication*, Routledge, London.

Kusyk M. 2023, *Does informal mean implicit?*, in Toffoli D., Sockett G. and Kusyk M. (eds.), *Language learning and leisure: Informal language learning in the digital age*, Walter de Gruyter, Berlin, pp. 43-68.

Lai C., Hu X. and Lyu B. 2018, *Understanding the nature of learners' out-of-class language learning experience with technology*, in "Computer Assisted Language Learning" 31 [1-2], pp. 114-143.

Lailiyah M. and Setiyaningsih L. 2020, *Students' perception of online communication language learning through Instagram*, in "EnJourMe (English Journal of Merdeka): Culture, Language, and Teaching of English" 5 [2], pp. 188-195.

Lee J.S. 2022, *Informal digital learning of English: Research to practice*, Routledge, New York, NY.

Lee Y.J. 2023, *Language learning affordances of Instagram and TikTok*, in "Innovation in Language Learning and Teaching" 17 [2], pp. 408-423.

Lei X., Fathi J., Noorbakhs S. and Rahimi, M. 2022, *The impact of mobile-assisted language learning on English as a foreign language learners' vocabulary learning attitudes and self-regulatory capacity*, in "Frontiers in psychology" 13, 872922, n.p.

Lier L.V. 2007, *Action-based teaching, autonomy and identity*, in "International Journal of Innovation in Language Learning and Teaching" 1 [1], pp. 46-65.

Mahmud M., Ammade S., Halim A. and Amin F.H. 2022, *Students' Voices of the Use of Facebook and Instagram in Teaching English in the University Context*, in "International Journal of Language Education" 6 [2], 113-127.

Manning J. 2014, *Definition and classes of social media*, in Harvey K. (ed.), *Encyclopedia of social media and politics*, Sage, Thousand Oaks, CA, pp. 1158-1162.

Meier A. and Johnson B.K. 2022, *Social comparison and envy on social media: A critical review*, in "Current opinion in psychology" 45, 101302, n.p.

Meirbekov A., Nyshanova S., Meiirbekov A., Kazykhankyzy L., Burayeva Z. and Abzhekenova B. 2024, *Digitisation of English language education: Instagram and TikTok online educational blogs and courses vs. traditional academic education. How to increase student motivation?* in "Education and Information Technologies" 29, pp. 13635-13662.

Metruk R. 2022, *Smartphone English language learning challenges: a systematic literature review*, in "Sage Open" 12 [1], pp.1-15.

Min T.S. and Hashim H. 2022, *Boosting students' motivation in learning descriptive writing through Instagram*, in "Creative Education" 13 [3], pp. 913-928.

Misnawati M., Yusriadi Y. and Tahir S.Z.B. 2022, *MALL in learning English through social networking tools: students' perceptions on Instagram feed-based task and peer feedback*, in "Computer-Assisted Language Learning Electronic Journal" 23 [2], pp. 198-216.

Nasution A.K.P. 2023, *Instagram in English Language Learning: A Systematic Literature Review*, in "JLLLT - Journal of Linguistics, Literature, and Language Teaching" 3 [1], pp. 33-52.

Reinders H., Lai C. and Sundqvist P. (eds.) 2022, *The Routledge handbook of language learning and teaching beyond the classroom*, Routledge, London.

Reinhardt J. 2019, *social media in second and foreign language teaching and learning: Blogs, wikis, and social networking*, in "Language Teaching" 52 [1], pp. 1-39.

Reinhardt J. 2020, *Metaphors for social media-enhanced foreign language teaching and learning*, in "Foreign Language Annals" 53 [2], pp. 234-242.

Richards, J. C. 2015, *The changing face of language learning: Learning beyond the classroom*, "RELC Journal", 46 [1], pp. 5-22.

Rezaie F. and Chalak A. 2021, *A SWOT analysis of Instagram English teaching pages*, in "Journal of Foreign Language Teaching and Translation Studies" 6 [3], pp. 25-44.

Schmidt R. 1995, *Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning*. In Schmidt R. (ed.), *Attention and Awareness in Foreign Language Learning*, University of Hawai'i Press, Honolulu, pp. 1-63.

Simsek T. 2023, *Social Media and Literacy Education: An Evaluation of Turkish Teaching Accounts on Instagram*, in "International Journal of Education and Literacy Studies" 11 [1], pp. 139-148.

Stockwell G. 2022, *Mobile assisted language learning: Concepts, contexts and challenges*, Cambridge University Press, Cambridge.

Sundqvist P. 2024, *Extramural English as an individual difference variable in L2 research: Methodology matters*, in "Annual Review of Applied Linguistics" 44, pp. 1-13.

Suryanti Tambunan A.R., Sari W.S. and Rasmitadila R. 2022, *Students' Perceptions of Social Networking Sites for English Language Learning: A Study in an Indonesian Higher Education Context*, in "World Journal of English Language" 12 [8], pp. 86-90.

Teng C., Heydarnejad T., Hasan M.K., Omar A. and Sarabani L. 2022, *Mobile assisted language learning in learning English through social networking tools: An account of Instagram feed-based tasks on learning grammar and attitude among English as a foreign language learners*, in "Frontiers in Psychology" 13. https://doi.org/10.3389/fpsyg.2022.1012004 (2.4.2024).

Toffoli D., Sockett G. and Kusyk M. (eds.) 2023, *Language learning and leisure: Informal language learning in the digital age*, Walter de Gruyter, Berlin.

Tommaso L. 2024, *A genre-oriented analysis of TikTok instructional discourse*, in "Translation and Translanguaging in Multilingual Contexts" 10 [1], 6-27.

Vizcaíno-Verdú A. and Abidin C. 2023, *TeachTok: Teachers of TikTok, micro-celebrification, and fun learning communities*, in "Teaching and Teacher Education" 123, 103978, n.p.

Wagner K. 2021, *Using Instagram for language learning*, in "Second language research & practice" 2 [1], pp. 154-164.

Zheng C., Liang J.C., Li M. and Tsai CC. 2018, *The relationship between English language learners' motivation and online self-regulation: A structural equation modelling approach*, in "System" 76, pp. 144-157.

Annexes

Institutional accounts			
Instagram account	N. of followers	*Instagram* account	N. of followers
@bbclearningenglish	4.4 M	@britishcouncil	231 K
@cambridgeenglish	747 K	@toefl_official	205 K
@takeieltsofficial	276 K	@britishcouncilenglishonline	196 K
@learnenglishwithoxford	267 K	@pearsonlanguages	106 K
Non-institutional accounts			
Instagram account	N. of followers	*Instagram* account	N. of followers
@jayfujiwara	7.3 M	@englishwithshei	512 K
@english.ingeneral	4 M	@englishenjoyed	503 K
@english.grammar.vocab	3.5 M	@english_squad__	488 K
@englishwithrahel	1.7 M	@idiom.land	439 K
@antonios.english	1.6 M	@englishwithus4	425 K
@englishwithnab	1.6 M	@activeenglish_programs	311 K
@phrasalidiomatic	1.6 M	@omg.classes	277 K
@teacherprix	1.6 M	@pronunciationwithemma	250 K
@daniellalewicz	1.2 M	@english.with.arezou	210 K
@englishwithrhys	1.2 M	@johnplusenglish	189 K
@english_veronika	1.1 M	@english_pro_tips	174 K
@ielts.lessson	1.1 M	@englishwithadriana	162 K
@ZenfluentMaria	1.1 M	@englisharound	150 K
@stage_door_johnny	998 K	@english_with_manuela	144 K
@englisteacherclaire	711 K	@englishwithjenna	114 K
@englishwithronia	785 K	@englishwithlizziejane	113 K
@shawenglishonline	691 K	@english_helenrox	109 K
@abraham.piper	682 K	@inenglishwithlove	98.8 K
@pronunciationworkshop	673 K	@pronunciation.studio	86.7 K
@fluentjoy_app	666 K	@etjenglish	85.5 K
@english_vocabulary	662 K	@tofluency	81.7 K
@hadar.accentsway	656 K	@listeningtime.english	71.6 K
@ash_britishenglish	649 K	@english_al_fresco	44.6 K
@rebeccas_english_hub	610 K	@lingojet_english	21 K
@english_lingo	525 K	@lola_speak	16.7 K

Table 2
Instagram accounts followed between January and March 2024.

Institutional accounts			
Instagram account	N. of followers	*Instagram* account	N. of followers
@bbclearningenglish	4.4 M	@learnenglishwithoxford	268 K
@cambridgeenglish	747 K	@britishcouncilenglishonline	196 K
@takeieltsofficial	276 K	@pearsonlanguages	106 K
Non-institutional accounts			
Instagram account	N. of followers	*Instagram* account	N. of followers
@jayfujiwara	7.3 M	@englishenjoyed	503 K
@englishwithrahel	1.7 M	@idiom.land	439 K
@englishwithnab	1.6 M	@pronunciationwithemma	250 K
@englishwithrhys	1.2 M	@johnplusenglish	189 K
@englisteacherclaire	711 K	@english_pro_tips	174 K
@englishwithronia	785 K	@englishwithadriana	162 K
@pronunciationworkshop	673 K	@englishwithjenna	114 K
@english_vocabulary	662 K	@englishwithlizziejane	113 K
@ash_britishenglish	649 K	@pronunciation.studio	86.7 K
@rebeccas_english_hub	610 K	@etjenglish	85.5 K
@english_lingo	525 K		

Table 3
Instagram accounts analysed qualitatively.

SOCIAL MEDIA
AND ONLINE INFORMAL LEARNING OF ENGLISH
University language students'
preferences, attitudes and perceptions

ELENA MANCA
UNIVERSITÀ DEL SALENTO

Abstract – The emergence of Web 2.0 technologies has significantly influenced informal language learning, offering learners greater control and autonomy over their educational journey. Within this context, social media and networking platforms have become integral to the informal learning landscape, fostering collaborative learning environments and facilitating knowledge sharing among users. Despite their initial design not being education-centric, these platforms have become prominent elements in modern e-learning environments, operating outside the control of traditional educational institutions. Building upon these trends, this research paper investigates the preferences, perceptions and attitudes of Italian university language students, particularly those at the University of Salento, towards the role of social media in learning English. Using a questionnaire inspired by the IECoL tool (Pavesi *et al.* 2023), administered to 213 language students, the current study seeks to explore the extent to which students use social media to enhance their English language proficiency. The findings aim to contribute insights into the evolving dynamics of informal language learning in the digital age and into the implications for language education practices.

Keywords: social media platforms; language learning; informal learning; perceptions; language education practices.

1. Introduction

Informal language is a field of research that has received significant attention in recent years. It has been defined using a variety of labels, such as, to name but a few, Extramural English (Sundqvist, Sylvén 2016), Informal Digital Learning of English (Lee 2022), Online Informal Learning of English (Sockett 2014; Toffoli, Socket 2015), Language Learning Beyond the Classroom (Reinders, Benson 2017). All these terms refer to how individuals learn and acquire by exposing themselves and interacting with leisure or meaning-based activities (Kusyk 2023, p. 37). In this paper, the label Online Informal Learning of English (OILE) by Toffoli and Socket (2015, p. 1) is adopted, as it specifically refers to language development through online activities such as social networking, streaming and/or downloading television series or films,

listening to music on demand and web browsing. Our interest here is on social media, that is to say on those digital platforms and websites that allow users to create, share, and engage with various forms of content, including text, images, videos, and links, while participating in social networking. These platforms support communication, collaboration, and the development of communities, enabling individuals, groups, and organisations to connect and exchange information in both real-time and asynchronous formats. Prominent examples include Facebook, X (formerly Twitter), Instagram, and TikTok, where users can post updates, follow others, interact through comments, and engage in discussions.

OILE differs from learning through OERs (Open Educational Resources) because the learning that occurs is mostly unintentional, as the primary goal of the activities is not to learn a language. Instead, individuals perceive themselves as engaging in recreational activities, with language learning happening as a side effect. In this context, it is the learner, rather than the instructor or course developer, who selects the materials to use, and these materials were not originally created for language learning purposes (Toffoli, Socket 2015, p.1). In OILE, incidental learning plays a primary role, as there is no deliberate intention to learn; rather, learning occurs as a by-product. Learners are often unaware of the learning process and its outcomes. While OILE includes both explicit and implicit elements—meaning that the intention to learn is not entirely absent—there is consistently a lack of conscious awareness of the learning taking place (Kusyk 2023, p. 44).

Over the last decades, the capabilities of Web 2.0 have massively contributed to the enhancing of the learning experience, and contact with foreign languages outside the classroom has become commonplace (Toffoli *et al.* 2023). This is also shown by several recent studies investigating the correlation between self-directed, informal second-language learning and advancements in technology (Toffoli, Socket 2015; Richards 2015; Dressman, Sadler 2020; Toffoli *et al.* 2023). Indeed, websites and applications that make use of user-generated content for end users contribute to language learning by empowering students and fostering greater control over their educational journey (Bielawska 2015). Toffoli and Sockett (2015) conducted a survey on the teacher's perceptions of and beliefs about online informal language learning and practices, and among the positive sides of OILE, teachers reported heightened curiosity about English and greater motivation to learn the language, often driven by the students' sense of understanding. They observed increased confidence and a general comfort with spoken English. For some students, English had become a regular part of their daily routines.

In the realm of social media, online informal learners are free to choose their own contents, are not subject to evaluation and the leisure activities they engage in are low in anxiety and high in motivation (Sockett 2014).

Lingue e Linguaggi

Social media and online informal learning of English. University language students' preferences, attitudes and perceptions

221

Furthermore, in online collaboration and communication, the sharing of knowledge becomes inevitable. Unbeknownst to users, this interactive dynamic fosters an environment conducive to informal learning, occurring subtly and indirectly through these digital interactions (Sockett 2014; Yaşar, Karadeniz 2011).

While not originally designed with education in mind, these digital affordances have become prominent elements within the modern e-learning environment. Unlike official educational applications, these tools are recognised as a significant component of the digital landscape for many learners, distinctly operating beyond the control of educational institutions (Selwyn 2007, p. 3). More specifically, social media foster collaborative and participatory roles and allow users to be engaged learners, rather than passive receivers, and active co-producers of contents.

Based on the ever more prominent role of social media platforms in language learning, this paper aims to investigate Italian university language students' preferences, perceptions and attitudes regarding the role of social media usage in learning English. The study particularly examines students' perceptions of the benefits of using social media platforms in English and their attitudes, that is to say, adapting Smith 1971, the enduring structuring of beliefs regarding social media usage and language learning, influencing their specific responses. To achieve this, a questionnaire consisting of 19 items, primarily multiple-choice, was administered to a sample of 213 language students from the University of Salento (Lecce, Italy). This analysis was inspired by the findings of the IECoL questionnaire, a tool used in an ongoing large-scale national (PRIN) project on the informalisation of English language learning among university students in Italy (Pavesi *et al.* 2023). A section of this questionnaire focused on the Internet in English as a form of input, revealing that language students from the four universities involved (Catania, Salento, Pavia, and Pisa) exhibited higher exposure to English language content on the Internet compared to students involved in the project but from other degree courses. For this reason, this paper sets out to explore whether and how a sample of University of Salento language students use social media to improve their level of competence of the English language.

The remainder of this paper is organised as follows: Section 2 provides an overview of previous studies analysing social media usage in autonomous language learning; Section 3 describes the participants in the current study, the administered questionnaire, and its findings; Section 4 offers concluding remarks and outlines the attitudes and perceptions of University of Salento language students regarding social media usage in learning English.

2. Previous studies

To the best of the author's knowledge, there is limited scholarly attention devoted to exploring the specific connection between social media and autonomous learning. The majority of existing studies predominantly centre on the integration of technology within formal educational settings, concentrate on the intersection of the Internet and informal learning, or investigate informal learning through digital media other than social media. What these studies have in common is their focus on learners' autonomy, which, in turn, leads to the development of successful strategies in the process of language learning. An illustrative example is presented by Dabbagh and Kitsantas (2011, p. 5) who focus on Personal Learning Environments (PLEs). PLEs are multidimensional spaces encompassing social media tools and enabling students to acquire competence or knowledge. Whether the tool facilitates interaction with peers on a class project or involves seeking online examples and suggestions for approaching a project, students use social media to cultivate informal learning communities centred on course topics. Personal Learning Environments are, therefore, transformed from an individual learning space into a communal learning environment. These collaborative endeavours actively involve students in self-regulation processes, encouraging self-monitoring, and prompt students to recognise and employ essential strategies for tackling more formal learning tasks (Dabbagh, Kitsantas 2011, p. 6). Although PLEs involve the input of instructors, the potential for cultivating learner autonomy is a positive aspect. When learners encounter challenges in initiating their informal language learning, the instructor can play a role in fostering their self-reliance and their ability to deal with problems and unusual or new situations autonomously. This study shows that autonomy and collaboration play pivotal roles in language learning, potentially fostering a sense of responsibility among learners for their own educational journey. This is where social media come into play, being characterised primarily by autonomous, collaborative and interactive use.

The advantages of employing social networking sites for foreign language learning among students was studied by Bicen *et al.* (2015) at Near East University, in Northern Cyprus. The research engaged a total of 85 undergraduate students, comprising 58 male students and 27 female students, who participated through the completion of a questionnaire. Among the available options, which required rating on the Likert scale, students conveyed their agreement with the following:

1. Social networks play a useful impact in improving foreign language;
2. I follow foreign language teachers through social networks to improve the target language;

Social media and online informal learning of English. University language students' preferences, attitudes and perceptions

223

3. I follow online pages on social networking sites to learn foreign language;

4. I use chat tools via social networking sites to improve foreign language;

5. I try to understand news in the foreign language while reading them on social networks;

6. I try to enhance the language ability by communicating with foreign people via social networks.

What is interesting about this study is that students expressed disagreement on the options "I try to answer tests in the foreign language to evaluate my level in the target language", "I force myself to learn the language by writing sentences on social networking groups", "Audiovisuals are used via social networks to practice foreign language", and "I feel confident about making posts in foreign language on social networks". This may indicate that students feel more confident using social media to enhance their receptive skills rather than their productive ones.

AbuSa'aleek (2015) explored university students' perceptions towards learning English in a Facebook context at Qassim University in the Kingdom of Saudi Arabia. The survey employed a questionnaire comprising two sections. Section A gathered demographic information about the students, including their level in the college and language proficiency. Section B included items aimed at gathering insights into the impact of Facebook as an online English language learning environment and exploring students' perceptions of learning English within the Facebook context. Results showed that students primarily interacted on Facebook using a combination of Arabic and English, or exclusively in English, but never in Arabic alone. Furthermore, EFL students hold the belief that Facebook, serving as an online learning environment, facilitates, supports, and motivates their English language learning. Consequently, it plays a role in assisting EFL students throughout the process of enhancing their English language skills and helps them overcome their language mistakes and learn new words. According to AbuSa'aleek (2015, p. 69), the favourable attitudes of students towards using Facebook as an online English language-learning environment could be attributed to the fact that Computer-Mediated Communication (CMC) allows EFL learners to access the Facebook platform at their convenience and from any location, thus confirming the importance of autonomy as a learning motivation. Additionally, Facebook encourages self-reliance through a student-centred approach, typically fostering improvements in language proficiency and overall growth. Consequently, students can acquire knowledge and skills from English speakers and participate in authentic asynchronous and synchronous interactions. EFL students also indicated that reading English materials on Facebook enhanced their confidence in communicating using the English language and their motivation to learn this foreign language. As Kabilan *et al.* (2010) state, the technologies underpinning

Facebook and the distinctive features of the platform have the capability to involve students in meaningful language-related activities, despite their initial purpose for joining Facebook being socialisation.

Ismail and Shafie (2018) conducted a research study involving students from three public universities in Selangor and one in Negeri Sembilan (Malaysia), with a sample size of 120 subjects. The study involved surveying the participants through a questionnaire, and 30 responses from each of the four universities were selected for analysis.

A specific section of the questionnaire focused on informal English language learning through Social Network Sites (SNS). The options with the highest scores related to ways of learning English informally through SNS were the following:

1. Save interesting quotes/images posted on SNS in English;
2. Pay attention to the sentence patterns and accent while watching videos on SNS;
3. Try to understand news in English language while reading them on SNS;
4. Watch English language videos posted on SNS to improve my English language.

Another section of the questionnaire focused on Perceived English Language Skills Learned Informally from the SNS. There were six categories listed (four basic English language skills, plus vocabulary and grammar), and the skill that had the highest mean score was the listening skill followed by vocabulary. The third section focused on Perceived Receptive and Productive Skills Learned through SNS, and the receptive skills (listening and reading skills) achieved the highest score. The authors concluded that informal learning through Social Network Sites (SNS) has the potential to complement formal learning and that SNS serve as the ideal platform for students to use and enhance their English language skills, even in the absence of direct monitoring from their instructors.

This view is emphasised by Godwin-Jones (2018, p. 18), who states that the co-adaptive nature of informal online language learning is notably pronounced in participatory Web 2.0 activities. Engaging in reading and writing within online communities significantly contributes to the development of a second language (L2). This ongoing process leads to the evolution of resources over time, transforming websites, social network services, discussion forums, and other affinity sites from static entities into dynamic and evolving platforms.

Murphy Odo (2020, p. 428) suggests that online fan communities serve as non-language oriented spaces which offer L2 learners an opportunity to engage with individuals globally, establishing connections based on shared interests like their preferred pop singers or fan fiction. This interaction can contribute to the cultivation of their identity as proficient users of the second

language and to learning autonomy. Illustratively, Stan Twitter serves as a prime example of these online fan communities. This community consists of Twitter users who express their opinions and enthusiasm about celebrities and related topics, displaying the dynamics of global interactions centred on shared interests in the realm of pop culture. Malik and Haidar (2021) analysed Twitter communities of K-Pop fandom and specifically focused on how memetic discourse is learnt and reused by non-native speakers of English and how interaction takes place. Some community members interviewed by the two scholars maintained that the platform enhances their proficiency in the English language. They engage in online learning using search engines, encountering memetic discourse on Twitter fandom communities, interacting with others, and internally assimilating memetic language patterns and sequences. These learning processes are incidental to the community's explicit objectives because members collaboratively support, promote, and express their shared love for the pop stars they follow (Malik, Haidar, p. 375).

Harmaini and Nanda (2023) focused on incidental learning on Instagram by considering a sample of 52 secondary school students from a vocational school in Padang (Indonesia). Students were administered a questionnaire which aimed to ascertain students' perception on the influence of Instagram use on their writing skills, vocabulary, grammar, and schematic learning. Results confirmed that students perceive Instagram as a supplementary tool for formal English learning. Most of them, when writing on Instagram, try to use appropriate vocabulary, grammar and patterns, thus showing that this social is not only an entertainment tool but also a learning platform.

2.1. The IECoL questionnaire

Prior to delving into the analysis of the data gathered for the current study, it is necessary to provide a concise overview of the IECoL questionnaire (Pavesi *et al.* 2023). This is essential because the analysis presented in this paper is derived from a survey conducted across four Italian universities, where the IECoL questionnaire was administered, namely the universities of Catania, Pavia, Pisa, and Salento.

The IECoL questionnaire (Pavesi *et al.* 2023, pp. 78-79), written in Italian, is structured into three main sections. Initially, it collects information regarding participants' language background, encompassing details such as the number of languages known, the age at which English instruction commenced in school, and involvement in extramural English language courses. The subsequent section of the questionnaire delves into behavioural inquiries, concentrating on exposure to diverse input types. This part is further segmented into subsections that cover a range of media: films, TV series, and programmes; YouTube videos; video games; songs and lyrics; Internet activities, which are subdivided into activities such as reading and posting on

social networks, blogs, and forums, as well as using various online platforms and websites. Additionally, it explores participants' engagement levels, frequency, duration of exposure per session, modalities of access, and motivating factors across each type of input.

The third section expands upon behavioural and attitudinal inquiries, as it collects demographic information from participants, as well as additional insights into their language background and activities beyond the conventional language learning setting.

The questionnaire developed for the current study narrows the scope of the IECoL questionnaire by focusing solely on social media use and exposure and aims to conduct a comprehensive analysis of the attitudes of a sample of University of Salento language students towards using social media in English. To achieve this, it combines some items from the IECoL with those described in the surveys outlined in section 3.1. Additionally, it introduces new questions to ascertain students' perceptions regarding the advantages of using social media in English within their L2 learning process.

3. Case study: participants and questionnaire

Data were collected at the University of Salento, a middle-sized university located in southern Italy. All the students involved in the project attend the BA course in Language Mediation Theories and Strategies (*Scienza e Tecnica della Mediazione Linguistica*). The decision to exclusively select language students stems from two primary reasons: (1) their increased motivation to use social media in English as a means to enhance their language proficiency, and (2) findings from a survey conducted across four Italian universities (Pavesi *et al.* 2023), which indicated that a significant proportion of respondents who acknowledged using the Internet in English were language students. Consequently, they emerged as the ideal target for this new survey.

The students who completed the questionnaire were 213 and were distributed across the three years of the BA course as follows: 93 from the 1st year, 64 from the 2nd year, and 56 from the 3rd year.

The questionnaire was in Italian, it was anonymous, and included 19 items focusing on:
- languages in which social media platforms are used;
- reasons for using or not using social media in English;
- frequently accessed topics and pages;
- types of activities carried out on social media;
- length of exposure;
- perceptions of the benefits of using social media in English;

-   use of social media platforms to improve the English language.

All the questions except the last one were multiple-choice, typically featuring the option '*Altro*' (Other), allowing respondents to specify additional choices not listed.

Prior to the administration of the questionnaire, students were informed about the objectives of the study, for which the questionnaire served as a tool, as well as the guaranteed anonymity of their responses.

3.1. Results from the questionnaire

At the beginning of the questionnaire, students were asked to indicate which social media platform they use, regardless of the language in which they use it. More than one option could be indicated. The question listed a series of social media platforms but students could also add other platforms by typing their names in the option '*Altro*' (Other). The most used revealed Instagram (94.9%), followed by TikTok (76.7%), Pinterest (51.2%), Twitter (30.8%), and Facebook (11.6%). A portion of the respondents (9.9%) added YouTube, which, similarly to socials, has recently introduced a real-time counter for video likes and views, facilitating active tracking of a video's popularity while watching. Additionally, YouTube features a comment section for each video, allowing users to engage in interactions.

This general question regarding social media usage placed at the outset of the questionnaire aimed to verify whether students actually use these platforms and to enhance the interpretation of the questionnaire's findings concerning usage and activities. This initial question was followed by a specific question on the language in which social media are used. Italian and English were listed as options, but respondents could also add other languages by using the empty option '*Altro*' (Other). Italian was selected by 78% of respondents while the English language seemed to be preferred as this option was chosen by 87.4% of students. Other languages were indicated, such as Spanish (8%), French (3.9%), German (2.9%), and Japanese (1.9%). These percentages seem to validate the decision to exclusively target foreign language students for this survey, leading to a more comprehensive insight into university language students' behaviours concerning social media usage and English language acquisition.

The questionnaire also included a question aimed at understanding the reasons why students do not use social media platforms in English. The question *Se NON usi i social in lingua inglese, puoi indicarci il perché?* [If you do NOT use social media in English, can you tell us why?] included the following options:

-   *Ci ho provato, ma non riesco a capire il contenuto dei post o dei reel* [I tried but I'm not able to understand the content of posts or of reels]

- *Per me i social sono intrattenimento e quindi li uso solo in italiano* [I consider socials as entertainment and, for this reason, I use them only in Italian]
- *Le pagine che mi interessano sono solo in italiano* [The pages I'm interested in are only in Italian]
- *Altro* [Other]

This question received responses from 27 students, comprising approximately 13% of total respondents. The second and third response options [I consider social media as entertainment and, for this reason, I use it only in Italian - The pages I'm interested in are only in Italian] were the most selected, with a selection rate of 32.4% and 20.6%, respectively. These two percentages, when considered together, may suggest that, for some students, the choice of language on social media platforms is driven more by established habits than by their level of English proficiency. Notably, only 11.8% opted for the first choice [I tried but I'm not able to understand the content of posts or of reels]. As for the '*Altro*' [Other] option, three students stated that they do not use social media platforms in English because they prefer other foreign languages, while another one admitted that s/he had never thought of setting English as the primary language of socials.

The next question focused on students' reasons for using socials in English, and the options listed were those proposed by the IECoL for a similar type of question to which three more options were added (the last three in the following list): *Se usi i social in lingua inglese, per quale motivo lo fai?* [If you use socials in English, for what reason do you do so?]

- *Per svago e intrattenimento* [For leisure and entertainment]
- *Per socializzare con persone straniere* [To socialise with foreign people]
- *Per acquisire familiarità con altre culture* [To get acquainted with other cultures]
- *Per studio universitario* [For university study]
- *Per imparare un vocabolario di uso specialistico* [To learn a specialised vocabulary]
- *Per imparare la lingua inglese* [To learn the English language]
- *Per usare la lingua inglese* [To use the English language]
- *Per accedere a informazioni in generale* [To access information]
- *Per accedere a informazioni in lingua inglese* [To access information in English]
- *Perché i social in lingua inglese offrono più intrattenimento e informazioni rispetto a quelli in lingua italiana* [Because socials in English offer more entertainment and information than those in Italian]

Social media and online informal learning of English. University language students' preferences, attitudes and perceptions

229

The last three options [To access information; To access information in English; Because socials in English offer more entertainment and information than those in Italian] were added to ascertain whether the choice of using social media in English could depend more on the amount and type of information available on socials in English rather than on the language in which contents are available. Figure 1 summarises the results for this question.
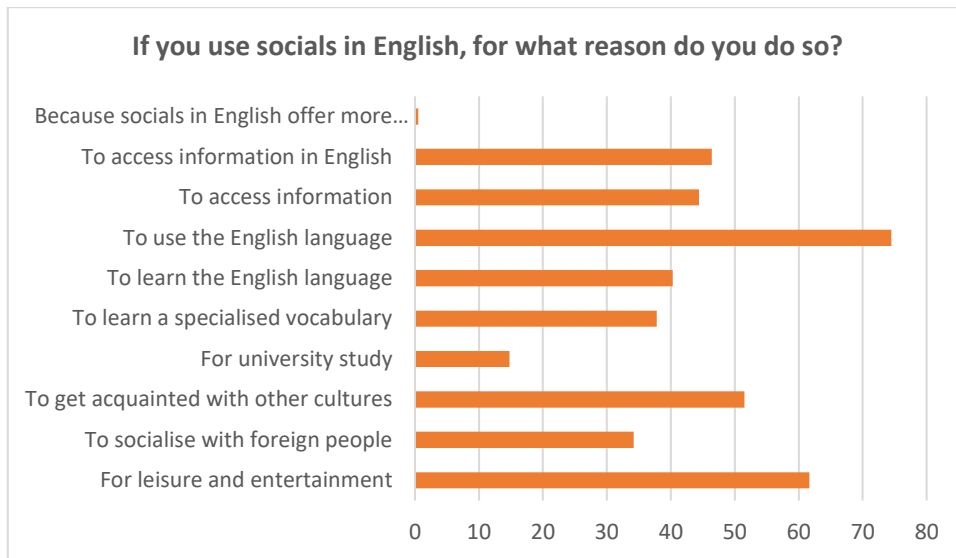
Figure 1
Reasons for using social media in English, with percentages.

As shown in Figure 1, the most selected option was 'To use the English language' (74.6%), indicating that university language students view social media as a viable virtual platform for practicing the language/s they are learning. The selection ranking as second in the list was 'For leisure and entertainment' (61.4%), followed by 'To get acquainted with other cultures' (51.8%), 'To access information in English' (46.2%), 'To access information' (44.2%). These percentages suggest that social media platforms are perceived as rich and multi-functional environments, enabling users to engage with foreign languages and cultures, as well as to be entertained and informed. The percentage associated to the selection 'To access information in English' reveals interesting insights into language students' attitudes on socials. They appear to be interested in exploring certain topics in English, possibly because they trust information in English more or aim to improve their language skills and expand their vocabulary.

To check which topics are accessed most frequently by students on social media, they were asked *Quali tematiche approfondisci più spesso tramite i social in lingua inglese?* [Which topics do you most often explore through social media in English?]. The options for this question were taken from the IECoL questionnaire, which included the same list of topics for a

question regarding the reasons for visiting blogs and forums. Figure 2 summarises the percentages for each selected option.
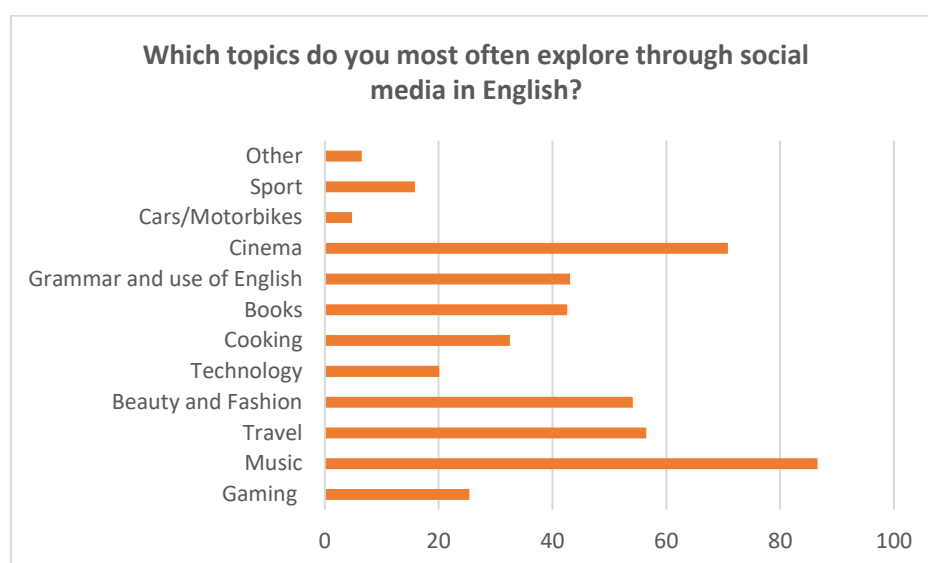
Figure 2
Topics more often explored on social media in English, with percentages.

Music and cinema rank at the top of the list with 86.6% and 70.8%, respectively. Travel is the third most selected option with a percentage of 56.5%, followed by Beauty and Fashion (54.1%), Grammar and use of English (43.1%), Books (42.6%), Cooking (32.5%), and Gaming (25.4%). The remaining topics (Technology, Sports, Cars/Motorbikes) have percentages ranging from 20% downwards. The relatively low percentage of the topic 'Grammar and use of English' may suggest that students primarily use social media to practice English in areas that they like or find interesting. Additionally, they use social media more traditionally as an easily accessible language-learning platform.

The question *Quali attività svolgi prevalentemente sui social in lingua inglese?* [Which activities do you primarily carry out on social media in English?] aims to explore attitudes and behaviours of language students in a more detailed manner. The options available were *Leggo contenuti* [I read posts], chosen by the majority of respondents and having a percentage of 95.7%, *Leggo e commento post* [I read and comment posts], with a percentage of preference of 21.2%, *Studio* [I study], selected by 24% of students. In the option *Altro* [Other], 12.7% of students indicated the action *Guardo video* [I watch videos]. The type of involvement on social media seems to be more receptive than productive or interactional, contrarily to the options selected for the question regarding the reasons for using social media in English. Indeed, the most frequently chosen was 'To use the English language' (74.5%) while the option 'To learn the English language' had a lower percentage (40.3%).

The next two questions are behavioural questions addressing both frequency and length of exposure to English through social media in order to quantify the hours of contact with English (Socket 2014; Pavesi *et al.* 2023, p. 76). The options listed in these questions are taken from similar questions of the IECoL referring to the length of exposure to different types of media, including social platforms. The first of these two questions focused on the time spent on social media reading posts in English (*Con che frequenza giornaliera leggi post in lingua inglese sui social?* How often do you read posts in English on social media each day?). Options were 'Never' (0%), 'Less than 30 minutes' (16.3%), 'Between 30 minutes and 1 hour' (19.1%), 'About 1 hour' (12.9%), 'From 1 to 2 hours' (22.5%), 'More than 2 hours' (29.2%). These figures suggest that students spend a high amount of time reading posts on socials. These percentages dramatically change in the second of the behavioural questions, focusing on the amount of time spent writing and commenting on posts. In fact, in response to the question 'How often do you write posts and comments in English on social media each day?' almost half of respondents (44.5%) answered selecting the option Never, 36.4% selected 'Less than half an hour', 8.1% chose 'Between 30 minutes and 1 hour', 5.3% selected 'About 1 hour'. The percentage of the options 'Between 1 and 2 hours' and 'More than 2 hours' was 2.9%. Most of the language students involved in the survey appear not to fully exploit the interactivity of this type of communication. Social media seem to be viewed more as a huge and varied repository of content rather than an interactive social environment where users can engage in synchronous and asynchronous conversations. This tendency to avoid interacting with other people on social media may be partly explained by the answers to another question, which focuses more specifically on the activity of writing posts, which reads: *Quando chatti o commenti un post in inglese, cerchi di evitare di fare errori?* [When you chat or comment on a post in English, do you try to avoid making mistakes?]. Interestingly, 49.5% of respondents answered 'Yes, I write and rewrite the post until it is correct', 42.2% selected the option 'I try to avoid making mistakes but interacting is what interests me most', and 6.1% of them chose the option 'I'm afraid of making mistakes and I usually write only a couple of words in the post'. Only 2% of respondents answered by selecting the alternative 'No, I primarily think about the content I'm communicating'. Based on these figures, we may hypothesise that students refrain from writing and commenting on posts because they lack confidence in their productive skills. Although they like to interact, the fear of making mistakes seems to act as an inhibiting factor.

The last group of questions focuses on students' perceptions of the benefits of using social media for language learning. The first of these perception questions is *Pensi che usare i social in lingua inglese possa aiutare a migliorare le tue competenze nella lingua?* [Do you think that using social media in English can help you improve your language skills?]. More than half

of respondents (56.5%) answered *Molto* [A lot], 40.5% selected the option *Abbastanza* [Enough], 2.8% answered *Poco* [Little], and only 1% chose the alternative *Per niente* [In no way]. Nobody selected the option *Non so* [I don't know] and this may indicate that all the respondents had a clear idea of the influence of social media on language learning. To better understand what kinds of activities, aside from reading content and comments, are viewed by students as important to improve their language skills, the following question was included: *Se usi i social per migliorare il tuo inglese, cosa fai esattamente?* [If you use social media to improve your English, what do you do exactly?]. Options and percentages are summarised in the graph in Figure 3.
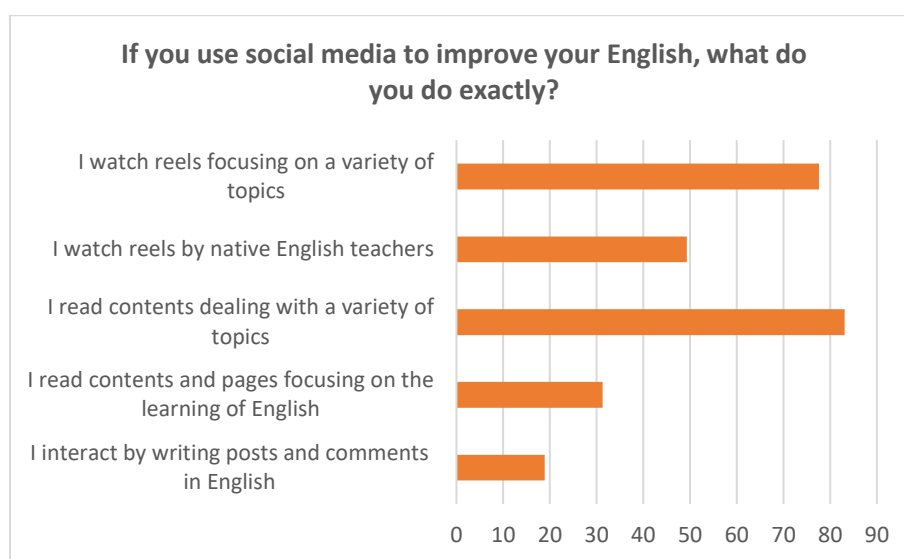
Figure 3
Activities carried out on social media to improve English, with percentages.

Percentages confirm what had already emerged in the previous questions, namely that language students primarily read posts (83.1%) and/or watch reels (77.6%) dealing with a variety of topics. This may indicate that students perceive they improve their English primarily by entertaining themselves with topics they like or are interested in, through reading and watching. The option 'I watch reels by native English teachers' was selected by 49.3% of respondents while the alternative 'I read contents and pages focusing on the learning of English' was chosen by 31.3%. We hypothesise that students may prefer watching videos to reading posts, probably because, in this way, they may also improve their listening and pronunciation skills. As expected, only 18.9% of respondents selected the option 'I interact by writing posts and comments in English' thus confirming their reluctance to use productive skills.

        To delve deeper into what students do on social media, they were asked four questions focusing on activities they consciously engage in to improve their English. The first of these is *Che cosa fai quando leggi o ascolti delle*

Social media and online informal learning of English. University language students' preferences, attitudes and perceptions

233

espressioni o parole che non conosci? [What do you do when you read or listen to expressions or words you don't know?]. Options and percentages of choice are summarised in Figure 4.
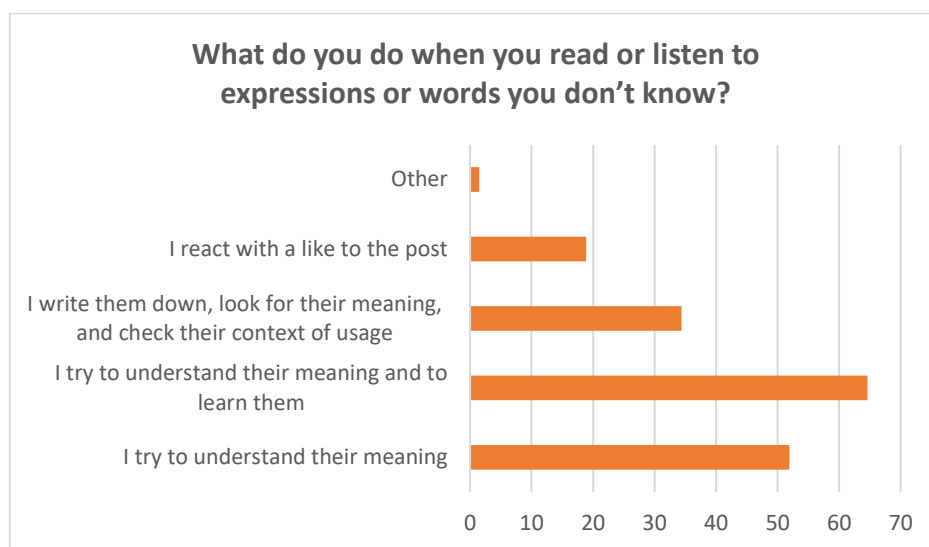
Figure 4
Activities related to words and expressions unfamiliar to students, with percentages.

When students encounter words and expressions they do not know they primarily try to understand their meaning and learn them (64.6%). Many students (51.9%) try to understand their meaning only, and 34.4% of them write down unfamiliar words and expressions, look for their meaning and check their context of usage. Some of them (18.9%) like the post, and, as some other language students explained to the author on a separate occasion, they usually get back to it later with the aim of checking the meaning of the words they do not know. The preferred choices illustrate the students' willingness to acquire new vocabulary while reading posts and watching videos on social platforms.

Another question was related to quotes and memes and reads *Ti capita di salvare o di fare lo screenshot a post contenenti citazioni o meme in inglese?* [Do you ever save or take screenshots of posts containing English quotes or memes?]. This question was included because, according to Ismail and Shafie (2018, p. 220), when quotes are deemed interesting enough to be saved, informal learning may occur. Almost all of the students (95.3%) replied 'Yes' to this question, thereby revealing an interesting behaviour on social media that should be taken into account in projects aiming to integrate language learning and new technologies.

The questionnaire also focuses on students' activities related to the usage of words. It does so by asking the question *Ti capita di riflettere su parole ed espressioni che incontri nei post perché ti interessa capire come vengono utilizzate?* [Do you reflect on words and expressions you come across in posts

because you're interested in understanding how they're used?]. Students could choose between two options, *Sì* [Yes] or *No* [No], and 96.7% of them selected the 'Yes' option, thus revealing a conscious usage of social media for learning purposes.

A similar question focused on the pronunciation of words: *Ti capita di riflettere sulla pronuncia delle parole inglesi quando guardi i reel e i video?* [Do you reflect on the pronunciation of English words when you watch reels and videos?]. Once again, the majority of respondents selected the 'Yes' option, confirming that content available on social media is considered a viable tool for improving their English fluency.

The last three questions focused on students' perception of the benefits of social media on their English. The first of these aimed to ascertain the degree of influence of social media on students' language skills: *Secondo te, quanto ha influito l'uso dei social in inglese sul tuo livello di competenza della lingua?* [In your opinion, how much has the use of social media in English influenced your language proficiency level?]. Students had six options available: *Moltissimo* (Very much) *Molto* (Much), *Abbastanza* (Enough), *Poco* (Little), *Non so* (I don't know). Most of them answered very positively, and almost all agreed on the importance of using social media in the process of learning the English language. Figure 5 provides a summary of the percentages related to this question.
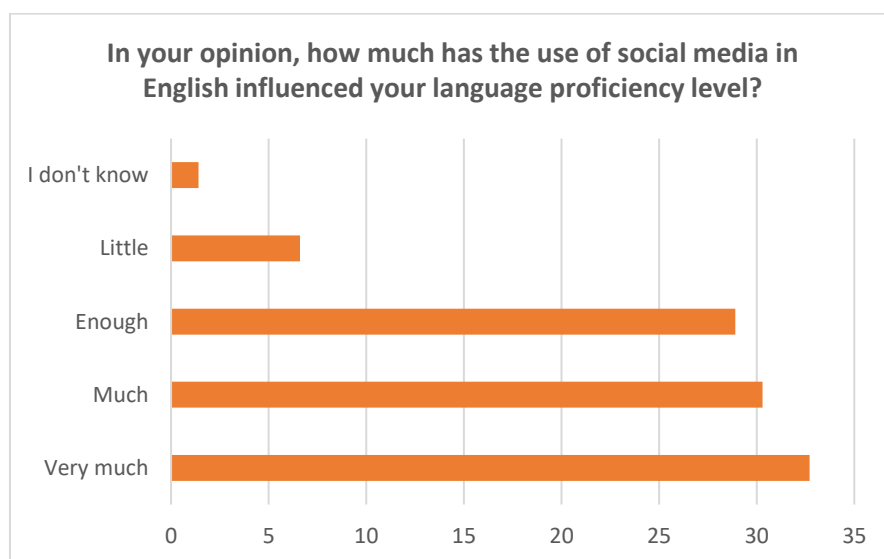
Figure 5
Students' perceptions on the level of influence of social media usage on their English, with percentages.

Next, students were asked their opinion on using social media in formal school and university teaching. To the question *Secondo te, sarebbe utile usare i social come uno strumento per la didattica scolastica e universitaria?* [In your opinion, would it be useful to use social media as a tool for school and

university teaching?], 91.6% answered 'Yes' and 8.4% disagreed answering 'No'. These percentages suggest that students consider social media as a valuable tool for language teaching. Reasons could be many: they may perceive the benefits of exposure to authentic examples of language use, or they could be encouraged to interact more and to engage in conversation with their peers, particularly if topics are selected based on students' interests. Additionally, this integration could make learning more dynamic and in line with young people's needs and likes.

The last item of the questionnaire asked students to indicate which media (not only social media) available on the Internet have mostly contributed to improving their language skills and fluency. YouTube was the most mentioned (26.5%), followed by TikTok (25.5%), Films and TV Series (21.3%), and Instagram (19%). Twitter was selected in 8% of cases while songs and music were mentioned with a percentage of 9.7%. Video games were mentioned only by 5% of respondents.

4. Some concluding remarks

The vast majority of the University of Salento language students involved in this survey actively engage with social media platforms, using both Italian and English languages.

Predominantly, they favour platforms such as Instagram, TikTok, Pinterest, alongside more traditional platforms like Twitter and Facebook. Their motivations for social media use include practicing English language skills, seeking entertainment, fostering cross-cultural interactions, and accessing information in English.

Within this sample, common themes explored through social media encompass cinema, music, travel, beauty, and fashion. When navigating social media pages, students predominantly read posts and comments, and a significant portion invests one to more than two hours in these activities. However, engagement in productive activities such as posting and commenting is less frequent. When they do contribute, students meticulously write their text, often revising to ensure accuracy or avoiding errors altogether.

Participants in the survey express a strong belief in the language-enhancing potential of English-language social media usage. They find value in reading diverse content and watching reels spanning various subjects. When encountering unfamiliar expressions, the majority endeavour to understand their meaning and incorporate them into their vocabulary. Furthermore, they actively collect English quotes and memes for future use, reflecting on words and expressions encountered in posts. Pronunciation is also a point of interest, as students strive to comprehend and emulate correct usage.

Almost unanimously, respondents perceive social media's profound impact on their language proficiency. Many advocate for its integration into informal educational settings at schools and universities, recognising its role in language learning and cultural exchange.

However, these findings should also be interpreted considering that the data collected from the survey may have been influenced by the context in which it was conducted—specifically, during a university lecture—and by the individual administering the survey, namely the lecturer. Respondents may have selected certain answers to align with the expectations of the survey organiser, despite assurances of anonymity. Nonetheless, the implications derived from this study remain significant across various domains.

Primarily, in the realms of language learning strategies and language awareness, social media emerge as a tool capable of enhancing students' sensitivity to linguistic nuances and overall language proficiency. This underscores the potential for educators to integrate social media platforms into language learning curricula, thereby bolstering students' capabilities in mastering a second language. Moreover, there is a pressing need to emphasise productive activities on social media platforms, alongside devising novel strategies to encourage active participation among students.

Furthermore, the study sheds light on the role of social media platforms such as Instagram, TikTok, and Pinterest in fostering cross-cultural interactions among students. These platforms serve as conduits for promoting cultural exchange and understanding, particularly pertinent in the context of our increasingly globalised world. The steadfast belief exhibited by participants in the language-enhancing capabilities of social media underscores the imperative of recognising these platforms as invaluable educational resources. Consequently, educational institutions should explore avenues for integrating social media into both formal and informal language learning settings, thereby enriching students' overall learning experiences.

In conclusion, the findings of this study underscore the multifaceted role that social media play in shaping the language learning and cultural experiences of University of Salento language students. Recognising and comprehending these dynamics can provide invaluable insights, guiding the formulation of effective educational practices and strategies aimed at harnessing the full potential of social media for language learning and cultural exchange.

Bionote: Elena Manca is an Associate Professor of English Linguistics, Language and Translation at the University of Salento (Lecce). Her areas of research are Corpus Linguistics, specialised languages, cross-cultural studies, translation studies. She is also interested in the use of technology in language teaching and learning, and has developed language-teaching materials for familiarising students with Corpus Linguistics and with the notion of extended units of meaning and functionally complete units of meaning which are at the basis of the interplay between language and meaning. She has extensively investigated the linguistic, visual and cultural features of tourism discourse by cross-comparing several multimodal tourist texts in English and in Italian produced by a variety of cultures. She has also elaborated the Cultural Communication Grammar framework for the classification of cultural features through the analysis of language and visuals.

Author's address: elena.manca@unisalento.it

References

AbuSa'aleek A.O. 2015, *Students' perceptions of English language learning in the Facebook context*, in "Teaching English with Technology" 15 [4], pp. 60-75.

Bicen H., Sadikoglu S. and Sadikoglu G. 2015, *The Impact of Social Networks on Undergraduate Students Learning Foreign Language*, in "Procedia - Social and Behavioral Sciences" 186, pp. 1045-1049.

Bielawska E.D. 2015, *Technology in ESP Pedagogy*, in "English for Specific Purposes World" 16 [48], pp. 1-23.

Dabbagh N. and Kitsantas A. 2011, *Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning*, in "Internet and Higher Education" 15, pp. 3-8.

Dressman S. and Sadler R.W. 2020, *The handbook of informal language learning*, Wiley Blackwell, Hoboken.

Godwin-Jones R. 2018, *Chasing the butterfly effect: Informal language learning online as a complex system*, in "Language Learning & Technology" 22 [2], pp. 8-27.

Harmaini F. and Nanda D.N. 2023, *Incidental writing learning on Instagram in informal setting*, in "KnE Social Sciences" 8 [13], pp. 277-284.

Ismail I.S. and Shafie N.H. 2018, *English informal language learning through social networking sites among Malaysian university students*, in "Asian Journal of University Education" 15 [3], pp. 211-225.

Kabilan M.K., Ahmad N. and Abidin M.J.Z. 2010, *Facebook. An online environment for learning of English in institutions of higher education*, in "The Internet and Higher Education" 13 [4], pp. 179-187.

Kusyk M. 2023, *Does informal mean implicit? Exploring the nature of informal second language learning*, in Toffoli D., Sockett G. and Kusyk M. 2023, *Learning and leisure. Informal language learning in the digital age*, De Gruyter, Berlin/Boston, pp. 37-51.

Lee J.S. 2022, *Informal digital learning of English. Research to practice*, Routledge, New York, NY.

Malik Z. and Haidar S. 2021, *English language learning and social media: Schematic learning on Kpop Stan twitter*, in "E-Learning and Digital Media" 18 [4], pp. 361-382.

Murphy Odo D. 2020, *Connecting Informal and Formal Language Learning*, in Dressman M. and Sadler R.W. (Eds.), *The Handbook of Informal Language Learning*, 1st Edition, John Wiley & Sons Ltd, Hoboken, pp. 421-438.

Pavesi M., Ghia E. and Ciabattoni T. 2023, *Profiling informal experience with English. An instrument for data collection*, in "Lingue e Linguaggi" 59, pp. 303-332.

Reinders H. and Benson P. 2017, *Research agenda: Language learning beyond the classroom*, in "Language Teaching" 50 [4], pp. 561-578.

Richards J.C. 2015, *The changing face of language learning: learning beyond the classroom*, in "RELC Journal" 46 [1], pp. 5-22.

Selwyn N. 2007, *Web 2.0 applications as alternative environments for informal learning- a critical review*, paper presented at *CERI-KERIS international expert meeting on ICT and educational performance*, Organization for Economic Cooperation and Development, South Korea. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=976e3d244e3b09f084af95115a8d04f8ff569ab7 (6.12.2024).

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave Macmillan, London.

Sundqvist P. and Sylvén L.K. 2016, *Extramural English in teaching and learning. From theory and research to practice*, Palgrave Macmillan, Basingstoke, London.

Social media and online informal learning of English. University language students' preferences, attitudes and perceptions

239

Toffoli D. and Sockett G. 2015, *University teachers' perceptions of Online Informal Learning of English (OILE)*, in "Computer Assisted Language Learning" 28 [1], pp.7-21.

Toffoli D., Sockett G. and Kusyk M. 2023, *Learning and leisure. Informal language learning in the digital age*, De Gruyter, Berlin/Boston.

Yaşar Ö. and Karadeniz S. 2011, *The power of social media in informal learning*, in Méndez-Vilas A. (Ed.), *Education in a technological world. Communicating current and emerging research and technological efforts*, Formatix Research Center, Badajoz, pp. 531-538.

VIDEO GAME IN-GAME DIALOGUE
An analysis of lexical coverage

FRANCESCA BIANCHI
UNIVERSITÀ DEL SALENTO

Abstract – The impact of extramural encounters with English in video games on L2 English learning is high. However, linguistic analysis of the language of the scripted dialogue in video games is still scarce. The current study investigates the lexical coverage of the lines of non-player characters in a small corpus of nine games from three different genres. The analyses show a lexical coverage comparable to that of British and American films. Furthermore, our data suggest that lexical coverage differences largely depend on individual games rather than genre, although setting may play some role.

Keywords: scripted language; performed language; pop-culture; non-player characters; text coverage; lexical profile.

1. Introduction

The impact of extramural encounters with English in video games on L2 English learning is high (e.g. Sundqvist, Wikström 2015). Several empirical studies have shown positive correlation between frequency and or length of engagement with video games and language skills in the foreign language and vocabulary knowledge in particular (e.g. De Wilde *et al.* 2019; De Wilde, Eyckmans 2017; Kuppens 2010; Sylvén, Sundqvist 2012; Sundqvist, Wikström 2015; Rød, Calafato 2024, in Calafato, Clausen 2024).[1]

Indeed, as "action-and-goal-directed preparations for, and simulations of, embodied experience" (Gee 2008, p. 203), video games are a powerful means for situated learning (Gee 2006, 2008). Furthermore, narrative structures provide thematically coherent linguistic as well as graphic contextualisation cues, which supports language awareness and language learning, and in particular form-meaning-function associations. Thus, vocabulary is more readily learned and retained for a longer time when encountered in narrative texts, rather than in non-narrative ones (Reinhardt 2018). This is indirectly confirmed by analyses of specific game genres. In fact, results from a few studies suggest that different game profiles may facilitate or hinder language acquisition. For instance, driving games, which lack a

---

[1] Studies contradicting this finding also exist (e.g. Peters 2018; Muñoz 2020).

narrative structure and involve controlling a vehicle while navigating obstacles rather than dialogic situations, have been found to negatively correlate with vocabulary learning (Calafato, Clausen, 2024). More broadly, high interactivity (e.g. actively playing the game) has been shown to impede vocabulary acquisition compared to lower levels of interactivity (e.g. watching the game; deHaan *et al*. 2010). This suggests that slower-paced games might be more conducive to vocabulary learning than fast-paced ones.

Despite all this, the attention posed by linguists to the study of the scripted dialogue in video games, which creates the narrative structure of the game (Domsch 2017), is scarce (Schmitt 2019), especially if compared to other narrative pop cultural multimedia genres, such as films or TV series. Most research on video games, in fact, has been carried out in fields other than linguistics, on the visual features of video games, or on game paratext.

The current study investigates scripted dialogue in video games, i.e. the lines of non-player characters (NPCs), in a corpus of nine titles belonging to three genres and analyses it in terms of lexical coverage. The study was inspired by a similar study by Rodgers and Heidt (2021) and other studies on text coverage. Our study, however, differs from Rodgers and Heidt's in several ways. First of all, these authors conflated in a single corpus written and spoken texts, game dialogue and paratext, which makes it difficult to compare results with other works on lexical coverage. On the other hand, our corpus exclusively comprises the spoken lines of NPCs. Furthermore, their corpus is not fully balanced in terms of game types, while ours includes three games per type. Finally, we shall also consider game setting as a potentially interesting variable for lexical coverage.

The aims of the current study are to 1) analyse the lexical coverage of the scripted language used by NPCs in the game; 2) assess whether game type and/or game setting have an impact on lexical coverage; 3) verify Rodgers and Heidt's (2021, p. 222) hypothesis that "the language that one would tend to find in video games may be expected to be more similar to the type found in movies and television (based on the aural mode of the language)".

The remainder of this paper is organised as follows: Section 2 outlines the conceptual background of this study, introducing video games (2.1) and linguistic studies of video game discourse, with a special focus on video game lexis (2.2), the notions of text-coverage and lexical coverage (2.3), and the results of previous studies on lexical coverage (2.4). Section 3 presents the materials and methods used in the current analysis. Section 4 reports the results of our analyses. Finally, Section 5 compares our results to previous ones on similar text types, adding comments and ideas for future research.

Video game dialogue as input for language learning. An analysis of lexical coverage

243

2. Conceptual background

2.1 Video games

"Video game" is an umbrella term for an enormous amount of different types of games (Winget 2011), all played electronically, to the point that scholars find it difficult to agree on a single definition, or even on what such a definition should be based on.[2] Our interest in video games resides in their being widespread multimedia pop-cultural artefacts. Like all other pop-cultural products, they "have a commercial, entertainment-related purpose, are (mass-)mediated, fall within the mainstream, and represent largely fictional and scripted content" (Werner 2022, p. 1).

An infinite number of classifications of video games exist, based on platform, age of the players, type of interaction, setting, actions performed by the players, etc., but none of them is free of problems: borders are blurred, games cross boundaries, and categories and subcategories eventually overlap (Ensslin 2012, p. 42). Despite this, in keeping with previous research (Rodgers, Heidt 2021), an attempt will be made in the current study to focus on specific genres.

For genres, we shall use Adams' (2010) classification. In particular, games from three genres will be considered: action games; action-adventure games; and role-playing games. As Adams (2010, p. 392) explains "[a]n action game is one in which the majority of challenges presented are tests of the player's physical skills and coordination. [...] In such games the player doesn't have time for strategy or planning". This category includes a range of subtypes, such as shooters and fighting games.[3] In this study, we shall consider action games that include some form or narration and dialogue between NPCs – the latter presenting and explaining the fictional world of the game to the player step by step (see Domsch 2017). On the other hand, "[t]he action-adventure is a hybrid genre, combining features from both action games and adventure games. To play them well requires a fair amount of physical skill, but they also offer a storyline, numerous characters, an inventory system, dialog, and other features of adventure games" (Adams 2010, p. 398). Finally, in role-playing games "the object [...] is to experience a series of adventures in an imaginary world, through an avatar character or a small group of characters whose skills and powers grow as time goes on" (Adams 2010, p. 453). As a consequence of structure and aims, these three genres differ in terms of the relevance that

---

[2] See for example Adams (2010, Chapter 1) for a technical definition of what a game is and what is not and for the difference between game and video game. But see also Bergonse (2017) and Arjoranta (2019) for a heated debate on the (im)possibility of outlining a definition of video game.

[3] Other researchers, including Lee *et al*. (2014), classify shooters and fighting games as separate categories.

dialogue has within the playing experience. As Domsch (2017, p. 265) points out, generally, in-game dialogue is "less prominent in first person shooters [included in this research in the adventure category], more elaborate in action adventures, prevalent in role-playing games."

Our study will also consider the setting of the video games, [4] hypothesising that the setting of the action may influence the topics of character talk. Actually, NPCs' talk performs the primary diegetic functions of describing the fictional world of the game and supporting the progression of the narrative, alongside the ludic function of giving instructions to the player (Domsch 2017). Settings are rarely classified by scholars, but amply cited by game producers, reviewers, bloggers and players (e.g. the Wikipedia page 'Category: Video games by setting' lists an enormous number of possible settings).[5] To the best of our knowledge, the only academic paper attempting a classification of settings is Lee *et al.* (2014). Based on the then existing labels from scholarly, commercial, and popular sources, they examined the subject field of video game genres and divided it into fundamental categories each representing an essential characteristic of division of the subject field, which they call 'facets'. They identify several facets, including Setting and Mood/Affect. [6] They define setting as "the surroundings or environment (spatial or temporal) in which the game takes place" (Lee *at al.* 2014, p. 134), and Mood/Affect as "the pervading atmosphere or tone of the video game which evokes or recalls a certain emotion or state of mind" (Lee *at al.* 2014, p. 135). For each facet, foci, i.e. indexing terms, are listed. Their list of foci under Setting include categories such as Cyberpunk, Futuristic, Gothic, Historic, Medieval, Modern, Renaissance, and Steampunk, while under Mood/Affect they list Horror, Humorous, Dark, and Peaceful. By the authors' admission, this scheme is incomplete and provisional and has been devised to be updated and revised upon feedback from gamers. Their considering Horror as a Mood/Affect feature rather than a type of setting is interesting, but debatable from a linguistics perspective. Although it is definitely true that horror images are meant to scare, horror products are characterised by specific types of characters, such as werewolves, vampires, zombies, demons and other monsters – physical entities inhabiting the fictional world of the game. These entities inhabit the game's spatial environment and their presence impacts on the topics of character talk as they will be named, described, talked about or event talked to by NPCs. For this reason, we prefer to consider horror as a type

---

[4] "Games can have identical settings and yet belong to different genres, so a medieval role-playing game belongs to a different genre than a medieval war game" (Adams 2010, p. 390).

[5] https://en.wikipedia.org/wiki/Category:Video_games_by_setting (8.7.2024).

[6] The other facets are Gameplay, Style, Purpose, Target audience, Presentation, Artistic style, Temporal aspect, Point-of-view, Theme, and Type of editing.

Video game dialogue as input for language learning. An analysis of lexical coverage

245

of setting. We are aware however that our classification, too, may be subject to future revisions and improvements.

2.2 Linguistic studies of video game discourse

Video games have been studied from a variety of perspectives, including the analysis of player-game relationship and the socio-cultural role of video games (e.g. Fernández-Vara 2014), the rhetoric of video games (Bogost 2007; Paul 2012), the study of video game characters (e.g. Fizek 2012), and video games as forms of communication (Gee 2014) to mention but a few.

Indeed, their complex nature lends itself to several forms of investigation. A few studies have analysed interaction between their various semiotic layers. Most of these studies however focus on images, written text and other graphic elements, disregarding character dialogue (e.g. Stamenković, Jaćević 2019; Stamenković 2023). An exception is represented by Stamenković *et al.* (2016), who analysed the famous 1998 action-adventure video game, *Metal Gear Solid*. In their multimodal analysis – which considers dialogue among in-game characters alongside other semiotic resources – the authors observed a network of discursive strategies aimed at persuading players to refrain from killing enemies. These include the use of both emotional and rational argumentation, expressive language, literary language (poetic imagery, symbols and metaphors) alongside referential discourse aimed at situating the game's narrative within a specific political, military and cultural paradigm.

A substantial amount of text linguistics research has focused on video game paratext and the language of players interacting online in Massive Multiplayer Online Role Playing Games (MMORPGs). These two areas have, actually, been the most productive for linguistic studies to date. This is certainly due to the indisputable importance that paratext plays in drawing a player's attention to a specific video game, constructing the player-game relationship and building and bonding the player community (Ensslin 2012). However, the fact that paratext is much easier to retrieve than in-game text may have also contributed to this focus. One of the first significant studies in this respect was carried out by Ensslin (2012), who created a corpus including several types of paratext (video game magazine articles, threads in gamer forums and chats, and live conversations during gameplay) and used it to discuss the way gamers and developers view and talk about games, gaming and game players. Her analysis considers semantic, pragmatic, metaphorical, textual and ideological aspects of gaming discourse. More recent studies on video game paratext include Álvarez-Bolado Sánchez and Álvarez de Mon (2019) on video game reviews, Balteiro (2019) on forums, Gledhill (2019) on tutorials and walkthroughs, Campos-Pardillos (2019) on end-user agreements, as well as Hancock (2019) on manuals. Studies on player chats and interactions

during gaming include Ensslin and Finnegan (2019), Rudge (2019), Graham and Dutt (2019), Kiourti (2019) and Iaia (2021).

The academic works considering the scripted dialogue of NPCs are few and they generally have specific slants: Iaia (2016, Chapter 7) analysed the scripted lines of non-English NCPs speaking in English and compared their ways of speaking to real ELF variations as described in the literature, while Heritage (2020, 2022) and Rennick *et al.* (2023) focused on gender representation.

A broader scope of analysis is pursued by Dixon, who compiled a corpus of the scripted in-game language of four highly popular titles of single-player offline role-playing games (Dixon 2022, further refined in Dixon 2024). The corpus includes both spoken and written scripted text in the game. The texts were divided into registers; in particular, the spoken texts were divided into Immersive Speech (speech where the player is not prompted to respond) and Interactive Speech (speech instances involving the player to respond). The various registers were compared to real-life discourse using Biber's (1988) Involved versus Information Production dimension (MDA Dimension 1) (Dixon 2022). Both Immersive Speech and Interactive Speech resulted towards the positive end of the involved-informational continuum, though with mean dimension scores (23.61 and 29.8, respectively) that are slightly lower than that of face-to-face conversation (35.3). The author interpreted these results declaring that "[t]he linguistic environments of these games represent language exposure that compares quite similarly to real-world language use situations" (Dixon 2022, p. 183). Dixon's results are not surprising, considering that in-game spoken text is a form of scripted text (or 'performed' text) produced for asynchronous fruition by an extradiegetic audience (the player). As Werner (2021) observes, these specific production and reception circumstances are markedly different from those of informal conversation and can only produce a linguistic outcome which is generally slightly more formal than unscripted dialogue. More specifically, although performed text may often simulate dialogue and be used to present (pseudo-)intimacy between stage participants, scripted language is a form of second-orality, as it is written-to-be-spoken. This implies that scripted texts have been preplanned and carefully edited for specific purposes and to achieve a specific effect. The result is a form of language that lacks those conversational features which are directly connected to real-time processing constraints. Furthermore, the presence of an extradiegetic audience requires performed text to be more explicit that spontaneous conversation (see Werner 2022, p. 8-9, among others).

Finally, a few studies have engaged in analyses of the vocabulary that appears in video games from an applied linguistics perspective with a view to language learning. Rodgers and Heidt (2021) analysed lexical coverage in a corpus of ten video games, discussing its distribution across game types. This

study – which inspired the current one – is more thoroughly presented in Section 2.4. Its primary finding is the indication that the lexicon in video games might be difficult for beginner or intermediate learners. This inspired Heidt *et al.* (2023) to create a pedagogical word list of gaming vocabulary that could be used by creators of game-related material or teachers to scaffold learners prior to their engaging with video games. In their study, the researchers distinguished diegetic vocabulary, i.e. words that pertain to the narrative level, from ludic vocabulary, i.e. words that are key to the manipulation of the game. Taking advantage of Rodgers and Heidt's (2021) game corpus and employing analytical methods that include keyness calculation, lexical frequency profiling (see Section 2.3), manual reading of concordance lines and manual classification of items, they produced three lists of gaming vocabulary and identified three classes of ludic vocabulary: Game System, Game Property, and Game Information. Their study demonstrates the existence of a distinctive vocabulary associated with games, shows that several game words can function in both diegetic and ludic contexts,[7] and also suggests that the kind of vocabulary found in games is genre-dependent.

2.3. Text coverage and Lexical coverage as notions

When talking about text comprehension, the terms 'text coverage' and 'lexical coverage' are frequently used as synonyms. However, they are more like two sides of the same coin.

Text coverage is the percentage of words in a text that the reader should know in order to gain some understanding of its contents. In an empirical experiment where a fictional text was manipulated by inserting varying percentages of non-existent words, Hu and Nation (2000) found a predictable correlation between the percentage of unknown words and the participants' degree of understanding. In their experiment, when knowledge of the tokens in the text was lower than 90%, no participant showed understanding. When it was 90% and 95%, a few managed to understand the text, but not many. For adequate comprehension of their narrative text, most participants needed to know as much as 98% of the words. However, while Liu and Nation (1985) suggested that 95% coverage may represent the percentage of known words necessary to correctly guess words in context, Laufer (1989) suggested that knowing 95% of text coverage is enough to understand a text. For these reasons, the 95% and 98% thresholds are generally considered as reference

---

[7] As the authors point out, the distinction between ludic and diegetic language is not clear cut as both the words and their context contribute to attributing an expression to one or the other category. Furthermore, while this distinction is certainly interesting form a theoretical and research perspective, its role in text comprehension and vocabulary acquisition is still unclear (Heidt *et al.* 2023, p. 156).

points for discussions of lexical coverage (e.g. Webb, Rodgers 2009a, 2009b; Rodgers, Heidt 2021).

Lexical coverage, by contrast, is a measure of the vocabulary size readers need in order to understand a text without external support. The first attempts at calculating vocabulary size date back to the end of the 19th century, but it was the advent of electronic corpora that really contributed to advances in this area.[8] As Nation and Anthony (2017, p. 360) point out, three ways of measuring lexical coverage have been devised, respectively based on the count of words (Meara, Buxton 1987; Meara, Jones 1988), lemmas (Leech *et al.* 2001) or word families (Bauer, Nation 1993). However, the word-families model is the one that has most productively been used in L2 research. For this reason, this paper will resort to Bauer and Nation's technique and related software to directly compare our results to previous studies using the same metrics and techniques.

In Bauer and Nation's (1993, p. 253) own words,

> "a word family consists of a base word and all its derived and inflected forms that can be understood by a learner without having to learn each form separately. […] The important principle behind the idea of a word family is that once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort. Clearly, the meaning of the base in the derived word must be closely related to the meaning of the base when it stands alone or occurs in other derived forms, for example, *hard* and *hardly* would not be members of the same word family."

These authors base the creation of word families on a seven-layer ranking of affixes, which is meant to mirror ease of learning for readers at different levels of morphological awareness. The inclusion of an affix at a specific level depends on the following criteria (Bauer, Nation 1993, pp. 255-256), frequency/likelihood being calculated against a large general corpus:

- Frequency, i.e. "the number of words in which the affix occurs";
- Productivity, i.e. "the likelihood that the affix will be used to form new words";
- Predictability, i.e. "the degree of predictability of the meaning of the affix";
- Regularity of the written/spoken form of the base;
- Regularity of the spelling/pronunciation of the affix;
- Regularity of function, i.e. "the degree to which the affix attaches to a base of known form-class and produces a word of known form-class".

The border between layers – though arbitrary – is nevertheless a useful way for standardising research on vocabulary size. Indeed, their approach has been

---

[8] For a comprehensive review of receptive and productive vocabulary size measures see Nation and Anthony (2017).

Video game dialogue as input for language learning. An analysis of lexical coverage

249

adopted in several studies of lexical frequency profiling (reviewed in Section 2.5), also thanks to the support of two software tools for quick and easy measure of lexical coverage – *Range* (Nation, Heatley 2002) and the more recent *AntWordProfiler* (Anthony 2022) – that have facilitated testing their measure on several different types of written and spoken texts and discussing their results with reference to L2 learning.

2.4. Studies on lexical coverage

Not unexpectedly, it has been observed that lexical coverage varies with text type.[9] A study by Nation on a range of written and spoken text types (Nation 2006) showed that from newspapers the most frequent 4,000 and 8,000 word families are respectively necessary to reach 95.39% and 98.31% text coverage; with a novel like *Lady Chatterly's Lover* similar thresholds could be reached with the most frequent 4,000 and 9,000 word families, respectively; while with a graded reader 2,000 and 3,000 word families are sufficient to reach 96.75% and 98.86%. In spontaneous conversation, 96.03% text coverage is reached with the 3,000 most frequent word families, but over 9,000 word families are required to reach the 98% threshold. On the other hand, in a movie like *Shrek*, 4,000 and 7,000 word families are needed for 96.74% and 98.08% text coverage, respectively. In a wider study on 318 American and English films, Webb and Rodgers (2009a) observed 95.76% coverage with a vocabulary of the most frequent 3,000 word families and 98.15% coverage with a vocabulary of 6,000 word families. If American and British films are considered separately, then the figures change slightly, but not substantially, with 3,000 word families and 6,000 families to reach, respectively, 95.76% and 98.14% for American films, and 3,000 and 7,000 word families to reach 95.50% and 98.29% for British ones. Larger differences could be observed between film genres, in particular to reach 98% coverage. While 3,000 word families were sufficient to reach 95% coverage in eight of the eleven film genres considered – war, animated and action films requiring knowledge of 4,000 word families for the same target –, the vocabulary needed to reach 98% coverage ranged from 5,000 to 10,000 word families – the 'least demanding' types of films being horror, crime and drama (5,000 word families), and the 'most demanding' being war (9,000 word families) and animated films (10,000 word families). The same authors have also investigated both American and British television genres, namely news, drama, situation comedy, older programmes, children's programmes, and science fiction (Webb, Rodgers 2009b). Considering the whole corpus – which comprised eighty-eight television programmes – the 95% threshold was reached with the most frequent 3,000 word families while the 98% threshold with the most frequent 7,000 word

---

[9] Text type is not the only variable influencing lexical coverage; other factors include length and homogeneity of the text (Hu, Nation 2000, p. 409).

families. Variation in coverage was observed between genres, with TV news being the most demanding, but above all between episodes of shows of the same genre. From 2,000 to 4,000 word families were required to reach the vocabulary size necessary to gain 95% coverage, while from 5,000 to 9,000 to gain 98% coverage.

Finally, to the best of our knowledge only two studies exist on the lexical coverage of the language of video games. Chien (2019) examined YouTube videos about *Minecraft*, a famous sand-box video game. They considered videos from three channels. One of the channels features spontaneous interaction between two American players filmed while playing *Minecraft* together. The other two channels feature monologic situations where the channel owner talks while playing or creating tutorials informing listeners how to play. Their analysis, performed with *AntWordProfiler* and the British National Corpus (BNC) as reference corpus, showed that 95% coverage is reached with 4,000 word families for monologic videos and as many as 6,000 for dialogic videos, while to reach 98% coverage more than 14,000 word families are necessary.

By contrast, Rodgers and Heidt (2021) focused on language within video games. These authors created a corpus of ten video games. The corpus comprises four role-playing games, two action-adventure games, two graphic adventure games and two life simulation games. The corpus included scripted in-game dialogues between characters, but also the texts of user interfaces, menus, in-game books, narration, tutorials, etc. (Rodgers, Heidt 2021, p. 219) and was analysed with *AntWordProfiler*. Considering the whole corpus, the first 5,000 word families proved necessary to reach 95% lexical coverage and the first 10,000 word families to reach 98% coverage. Furthermore, the authors noticed individual differences between video games, but no consistent differences between game genres.

Methodologically, in all the studies above, the coverage percentage mentioned is reached by adding marginal words – i.e. exclamations, interjections and the like – and proper nouns to the percentage figures, based on the assumption that these can be recognised and learnt with very little effort (Nation 2006). Furthermore, Nation (2006) and Webb and Rodgers (2009a, 2009b) took advantage of the *Range* programme and the BNC as reference corpus, while Rodgers and Heidt (2021) used *AntWordProfiler* and both the BNC and the Corpus of Contemporary American English (COCA). *Range* and *AntWordProfiler* respectively consider 14,000 and 24,000 word families and classify unrecognised words in a separate group called *Not in the lists*. Webb and Rodgers (2009a) observed the presence of many fairly common words in the *Not in the list* category, which suggests that their results – and by extension all counts with *Range* – could perhaps be rather conservative.

3. Materials and methods

Video game dialogue as input for language learning. An analysis of lexical coverage

251

This paper analyses the scripted language used by characters of nine video games belonging to three different types. In this respect, this dataset represents a relatively balanced subset of a larger corpus of video games that is currently under construction and was conceived as a pilot corpus for testing the available analytical methods and tools. The *Video game Pilot Corpus* (ViPiCo) includes three action-adventure games (AAG), comprising the AAG sub-corpus, three action games (AG), comprising the AG sub-corpus, and three role-playing games (RPG), comprising the RPG sub-corpus. The details of the pilot corpus are illustrated in Table 1; the codes in the Type-Setting column are detailed in Table 2.·

File name	Type-Setting	Title	Year	Developers	Main story length (hours)[10]	Token count
1	AA-Ho	*Resident Evil Village*[11]	2021	Capcom	10	10,324
2	AA-Pa	*The Last of Us*[12]	2013	Naughty Dog	15	27,301
3	AA-F/PA	*Devil May Cry V*[13]	2019	Capcom	11	10,385
		AAG sub-corpus			**36**	**48,010**
4	A-SF	*Dead Space 2*[14]	2011	Visceral Games	9	10,167
5	A-Hi	*Mafia Definitive Edition*[15]	2020	Hangar 13	10.5	32,928
6	A-SF	*Halo Infinite*[16]	2021	343 Industries	11.5	14,535
		AG sub-corpus			**31**	**57,630**
7	RP-F	*Bloodborne*[17]	2015	From Software	33	12,429
8	RP-F	*Chrono Trigger*[18]	1995	Square	23	15,753
9	RP-F	*Pokémon Black*[19]	2010	Game Freak	32	33,836
		RPG sub-corpus			**88**	**62,018**
		ViPiCo				**167,658**

Table 1
Details of the games comprising the corpus.

Code	Type	Setting
AA-Ho	Action-Adventure	Horror
AA-PA		Post-apocalyptic

[10] Data in this column were obtained from the *HowLongToBeat* website. https://howlongtobeat.com/ (8.7.2024).

[11] https://www.residentevil.com/village/us/ (8.7.2024); https://en.wikipedia.org/wiki/Resident_Evil_Village (8.7.2024).

[12] https://www.playstation.com/en-gb/games/the-last-of-us-part-i/?country-selector=true (8.7.2024).; https://en.wikipedia.org/wiki/The_Last_of_Us (8.7.2024).

[13] https://www.devilmaycry.com/5/ (8.7.2024); https://en.wikipedia.org/wiki/Devil_May_Cry_5 (8.7.2024).

[14] https://www.ea.com/games/dead-space/dead-space-2 (8.7.2024); https://en.wikipedia.org/wiki/Dead_Space_2 (8.7.2024).

[15] https://mafiagame.com/mafia/ (8.7.2024); https://en.wikipedia.org/wiki/Mafia:_Definitive_Edition (8.7.2024).

[16] https://www.halowaypoint.com/ (8.7.2024); https://en.wikipedia.org/wiki/Halo_Infinite (8.7.2024).

[17] https://www.fromsoftware.jp/ww/detail.html?csm=094 (8.7.2024); https://en.wikipedia.org/wiki/Bloodborne (8.7.2024).

[18] https://www.jp.square-enix.com/chronotrigger/en/ (8.7.2024); https://en.wikipedia.org/wiki/Chrono_Trigger (8.7.2024).

[19] https://www.pokemon.com/us/pokemon-video-games/pokemon-black-version-and-pokemon-white-version (8.7.2024); https://en.wikipedia.org/wiki/Pok%C3%A9mon_Black_and_White (8.7.2024).

AA-F/PA		Fantasy/Post-apocalyptic
A-SF		Science fiction
A-Hi	Action	Historical
RP-F	Role-Playing	Fantasy

Table 2
Type-setting codes.

Unfortunately, the ViPiCo could not be balanced in terms of settings, nor in terms of token counts. The year of publication was also not considered as a separate variable. These certainly represent limitations of the current study, and this will be considered in the discussion of results and in prospective research.

For each game, the scripted language was retrieved from the Internet, from fans' transcripts (a method also used by Heritage 2020). More specifically, the following search string was used in the Google search engine: *name_of_game game script*. The scripts were then copied and pasted into text files. To make sure that the corpus exclusively included the characters' lines, the scripts were manually purged from comments and descriptions of action – which luckily always appeared in either square or round brackets – and of characters' names when these were used to mark the speaker. The scripts were also checked for evident typos, and the latter were fixed. Furthermore, for comparability with previous studies based on Nation's framework and tools (see Section 2.4), the spelling of contractions, connected speech, and hyphenated words was adjusted to conform to that used in the reference lists available with the software. All typos and variants were identified through manual scanning of word lists and fixed using the search-and-replace feature in MS Word.

In keeping with Rodgers and Heidt (2021, p. 220), the analyses were carried out using *AntWordProfiler* (Anthony 2022), the most recent software tool specifically created for the identification of lexical coverage using the word-families model (Nation, Anthony 2017), and the BNC and COCA word lists were employed as terms of comparison. Furthermore, coverage was calculated cumulating percentages and including proper nouns and marginal words; items in the *Not in the list* category were manually checked for proper nouns and marginal words to add to the cumulative counts. Indeed, in all files, the off-list category included some percentage of unusual proper nouns (e.g. Gbraakon; Geryon)[20] and imaginatively-spelt marginal words (e.g. *ooof*; *whewww*). These calculations are reported in the tables under the tag *Cum. Cov. %*.

During the manual analysis of the off-list items, it was observed that in one game (the shooter *Mafia Definitive Edition*), the *Not in the list* category also included common Italian words (e.g. *mamma*, *nonna*, *impara*). A check of

---

[20] Gbraakon is one of several Banished warships, which fought at the Battle of Zeta Halo, in the *Halo Infinite* video game; Geryon is a character in the Devil May Cry series, named after a giant of the Greek mythology.

Lingue e Linguaggi

Video game dialogue as input for language learning. An analysis of lexical coverage

253

the full transcript of the game showed that all lines in Italian were accompanied by English subtitles – appearing in parentheses right next to the Italian line. Since the English subtitles are included in the corpus and automatically analysed in terms of word families, the Italian words were disregarded.

4. Findings

Table 3 presents percentage token counts and percentage cumulative coverage including proper nouns and marginal words (*Cum. cov. %*) of the whole corpus, as well as its three sub-corpora. Percentage token counts are reported exactly as extracted by the software. Manual calculations of the proper nouns and marginal words in the off-list category – used for the cumulative coverage % measure – are reported in the comment to the table.

Percentage token counts show analogous trends in all the (sub-)corpora: 83.27%-87.12% of the running words belong to the most frequent 1,000 families, while the second set of 1,000 word families accounts for only 3.44%-4.96% of the running words; token counts gradually decrease as the word families increase. These results are in keeping with the token-count trends observed in all previous studies.

Considering measure *Cum. Cov. %*, 95% text coverage is reached with only 3,000 word families plus proper names and marginal words considering ViPiCo as a whole, as well as in the AG and in the RPG sub-corpora, while it requires as few as 2,000 word families in the AAG sub-corpus. On the other hand, greater variation is visible across the four corpora when it comes to reaching 98% coverage, with as few as 5,000 word families plus proper names and marginal words for the AAG sub-corpus, 7,000 for the RPG one and as many as 8,000 for the AG one. In the entire corpus, 7,000 word families plus proper names and marginal words are needed to reach this threshold.

The *Not in the list* category included quite a few proper nouns and some marginal words. The percentage values of proper nouns were the following: 0.24%, 0.59%, 2.77% and 1.3% in the AAG, AG, RPG and ViPiCo (sub-)corpora, respectively. The percentage values of marginal words were as follows: 0.06%, 0.01%, 0.15% and 0.08% in the AAG, AG, RPG and ViPiCo (sub-)corpora, respectively.

Word families	ViPiCo		AAG sub-corpus		AG sub-corpus		RPG sub-corpus	
	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %

1.000	85.46	89.87	88.29	92.57	85.44	89.44	83.27	88.33
2.000	4.08	93.95	3.12	**95.69**	3.93	93.37	4.96	93.29
3.000	1.61	**95.55**	1.27	96.97	1.70	**95.06**	1.79	**95.08**
4.000	1.04	96.59	0.69	97.65	1.13	96.20	1.22	96.30
5.000	0.82	97.41	0.59	**98.24**	0.69	96.89	1.12	97.42
6.000	0.48	97.89	0.32	98.56	0.57	97.46	0.50	97.92
7.000	0.27	**98.16**	0.15	98.71	0.40	97.86	0.24	**98.17**
8.000	0.24	98.39	0.21	98.92	0.25	**98.11**	0.25	98.41
9.000	0.18	98.57	0.19	99.11	0.13	98.24	0.21	98.62
10.000	0.10	98.68	0.06	99.17	0.08	98.32	0.16	98.79
11.000	0.10	98.78	0.05	99.22	0.13	98.45	0.12	98.90
12.000	0.07	98.85	0.05	99.28	0.09	98.54	0.07	98.97
13.000	0.07	98.92	0.03	99.31	0.06	98.59	0.10	99.07
14.000	0.03	98.95	0.02	99.33	0.03	98.63	0.04	99.12
15.000	0.04	98.99	0.04	99.37	0.03	98.66	0.04	99.16
16.000	0.01	99.00	0.01	99.38	0.01	98.67	0.01	99.17
17.000	0.01	99.01	0.01	99.39	0.02	98.69	0.01	99.18
18.000	0.03	99.04	0.00	99.40	0.06	98.75	0.02	99.20
19.000	0.02	99.06	0.00	99.40	0.02	98.76	0.03	99.23
20.000	0.02	99.08	0.03	99.42	0.00	98.77	0.04	99.26
21.000	0.01	99.09	0.02	99.45	0.01	98.78	0.01	99.27
22.000	0.00	99.1	0.01	99.45	0.00	98.79	0.00	99.27
23.000	0.01	99.1	0.01	99.46	0.01	98.79	0.01	99.28
24.000	0.10	99.2	0.00	99.46	0.21	99.01	0.06	99.34
25.000	0.00	99.2	0.00	99.46	0.01	99.01	0.00	99.34
proper nouns	1.56		1.84		2.39		0.56	
marginal words	1.54		2.14		1.00		1.58	
transparent compounds	0.30		0.36		0.37		0.17	
acronyms	0.07		0.03		0.08		0.09	
ignored	0.00		0.00		0.00		0.00	
not_in_lists	1.74		0.44		1.14		3.32	
TOTAL	100.00		100.00		100.00		100.00	

Table 3
Token count (%) and cumulative coverage (%) of ViPiCo and its three sub-corpora.

Word families	Resident Evil Village		The Last of Us		Devil May Cry V	
	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %
1.000	86.32	90.45	87.78	93.75	82.20	87.35
2.000	4.26	94.72	3.16	**96.43**	4.68	92.03
3.000	1.96	**96.67**	1.27	97.45	3.27	**95.29**
4.000	0.89	97.56	0.95	97.95	1.46	96.75
5.000	0.53	**98.10**	1.05	**98.39**	0.75	97.5
6.000	0.40	98.49	0.43	98.64	0.35	97.85
7.000	0.25	98.74	0.18	98.73	0.46	**98.31**
8.000	0.21	98.96	0.11	98.97	0.20	98.51
9.000	0.07	99.03	0.22	99.21	0.26	98.77
10.000	0.06	99.08	0.16	99.23	0.12	98.88
11.000	0.05	99.13	0.06	99.28	0.11	98.99
12.000	0.08	99.21	0.09	99.31	0.07	99.06
13.000	0.03	99.24	0.04	99.34	0.10	99.16
14.000	0.00	99.24	0.02	99.36	0.01	99.17
15.000	0.06	99.30	0.03	99.40	0.02	99.19
16.000	0.03	99.33	0.02	99.41	0	99.19
17.000	0.02	99.35	0.01	99.42	0.01	99.2
18.000	0.00	99.35	0.01	99.42	0	99.2
19.000	0.00	99.35	0.00	99.42	0	99.2
20.000	0.10	99.44	0.03	99.42	0	99.2
21.000	0.00	99.44	0.11	99.42	0	99.2
22.000	0.00	99.44	0.03	99.42	0.01	99.21
23.000	0.00	99.44	0.00	99.42	0	99.21
24.000	0.01	99.45	0.00	99.44	0	99.21
25.000	0.00	99.45	0.00	99.44	0	99.21
proper names	2.35		1.26		2.73	
marginal words	1.44		1.61		1.09	
transparent compounds	0.26		0.30		0.35	
acronyms	0.01		0.05		0.08	
ignored	0.00		0.00		0	
not_in_lists	0.61		1.03		1.68	
TOTAL	100		100		100	

Table 4

Token count (%) and cumulative coverage (%) of the games in the AAG sub-corpus.

Lingue e
Linguaggi

Word families	*Dead Space 2*		*Mafia Definitive Edition*		*Halo Infinite*	
	Token count %	**Cum. Cov. %**	**Token count %**	**Cum. Cov. %**	**Token count %**	**Cum. Cov. %**
1,000	82.20	87.35	87.3	91.75	83.5	85.67
2,000	4.68	92.03	3.26	**95.01**	4.91	90.58
3,000	3.27	**95.29**	0.75	95.76	2.73	93.31
4,000	1.46	96.75	0.77	96.53	1.73	**95.04**
5,000	0.75	97.50	0.57	97.10	0.93	95.97
6,000	0.35	97.85	0.46	97.56	0.99	96.96
7,000	0.46	**98.31**	0.23	97.79	0.72	97.68
8,000	0.20	98.51	0.25	**98.04**	0.28	97.97
9,000	0.26	98.77	0.07	98.11	0.17	**98.14**
10,000	0.12	98.88	0.03	98.14	0.17	98.31
11,000	0.11	98.99	0.11	98.25	0.20	98.51
12,000	0.07	99.06	0.06	98.31	0.16	98.67
13,000	0.10	99.16	0.04	98.35	0.08	98.74
14,000	0.01	99.17	0.06	98.41	0.00	98.74
15,000	0.02	99.19	0.05	98.46	0.01	98.75
16,000	0.00	99.19	0.01	98.47	0.01	98.76
17,000	0.01	99.20	0.01	98.48	0.05	98.81
18,000	0.00	99.20	0.04	98.52	0.12	98.94
19,000	0.00	99.20	0.03	98.55	0.01	98.94
20,000	0.00	99.20	0.01	98.56	0.00	98.94
21,000	0.00	99.20	0.01	98.56	0.03	98.98
22,000	0.01	99.21	0.00	98.57	0.00	98.98
23,000	0.00	99.21	0.01	98.57	0.02	99.00
24,000	0.00	99.21	0.37	98.95	0.00	99.00
25,000	0.00	99.21	0.01	98.96	0.00	99.00
proper names	2.73		2.93		0.92	
marginal words	1.09		1.16		0.58	
transparent compounds	0.35		0.38		0.36	
acronyms	0.08		0.09		0.07	
ignored	0.00		0.00		0.00	
not_in_lists	1.68		0.93		1.24	
TOTAL	100		100		100	

Table 5

Token count (%) and cumulative coverage (%) of the games in the AG sub-corpus.

Word families	Bloodborne		Chrono Trigger		Pokémon Black	
	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %	Token count %	Cum. Cov. %
1.000	84.61	87.98	80.51	86.76	84.06	89.17
2.000	4.01	91.99	5.41	92.17	5.10	94.28
3.000	1.49	93.47	1.85	94.02	1.86	**96.14**
4.000	2.44	**95.91**	1.34	**95.36**	0.72	96.87
5.000	0.84	96.76	1.07	96.43	1.24	**98.11**
6.000	0.54	97.30	0.60	97.03	0.45	98.55
7.000	0.43	97.72	0.39	97.43	0.10	98.66
8.000	0.31	**98.03**	0.33	97.76	0.19	98.84
9.000	0.29	98.32	0.22	97.98	0.17	99.02
10.000	0.20	98.52	0.18	**98.16**	0.14	99.16
11.000	0.36	98.88	0.05	98.21	0.06	99.22
12.000	0.12	99.00	0.09	98.30	0.04	99.26
13.000	0.05	99.05	0.30	98.59	0.03	99.29
14.000	0.04	99.09	0.07	98.66	0.03	99.32
15.000	0.05	99.14	0.05	98.71	0.04	99.36
16.000	0.01	99.15	0.03	98.74	0.01	99.37
17.000	0.02	99.16	0.00	98.74	0.01	99.38
18.000	0.02	99.18	0.02	98.76	0.02	99.40
19.000	0.01	99.19	0.08	98.84	0.01	99.41
20.000	0.08	99.27	0.08	98.91	0.00	99.41
21.000	0.01	99.27	0.01	98.92	0.00	99.41
22.000	0.01	99.28	0.00	98.92	0.00	99.41
23.000	0.00	99.28	0.02	98.94	0.01	99.42
24.000	0.00	99.28	0.23	99.17	0.01	99.43
25.000	0.00	99.28	0.00	99.17	0.00	99.43
proper names	0.23		1.20		0.38	
marginal words	2.46		1.45		1.31	
transparent compounds	0.10		0.22		0.18	
acronyms	0.08		0.11		0.09	
ignored	0.00		0.00		0.00	
not_in_lists	1.21		4.10		3.72	
TOTAL	100		100		100	

Table 6

Token count (%) and cumulative coverage (%) of the games in the RPG sub-corpus.

Within the sub-corpora ample internal variation can be observed (Tables 4-6). Considering the *Cum. Cov. %* measure, in the AAG sub-corpus, the number of word families needed to reach 95% coverage ranges from 2,000 in *The Last of Us*, to 3,000 in *Resident Evil Village* and *Devil May Cry V*, while the number needed to reach 98% coverage amounts to 5,000 in *Resident Evil Village* and *The Last of Us* and as many as 6,000 in *Devil May Cry V*. In the AG sub-corpus,

the number of word families needed to reach 95% coverage goes from 2,000 in *Mafia Definitive Edition*, to 3,000 in *Dead Space 2*, to 4,000 in *Halo Infinite*, while the number needed to reach 98% coverage amounts to 6,000 in *Dead Space 2*, 8,000 in *Mafia Definitive Edition* and as many as 9,000 in *Halo Infinite*. Finally, in the RPG sub-corpus, both *Bloodborne* and *Chrono Trigger* require 4,000 word families to reach 95% coverage; they however differ as for the 98% coverage, needing, respectively, 8,000 and 10,000 word families. On the other hand, *Pokémon Black* requires as few as 3,000 and 5,000 word families to reach the first and second thresholds, respectively.

If we group the video games by setting (Table 7), no clearer picture emerges. This is largely a consequence of the fact that only two settings out of six are represented by more than one game in the PiViCo.

Setting type	Game title	95% coverage, reached at	98% coverage, reached at
Historical	*Mafia Definitive Edition*	2,000	8,000
Horror	*Resident Evil Village*	3,000	5,000
Post-Apocalyptic	*The Last of Us*	2,000	5,000
Fantasy/Post-Apocalyptic	*Devil May Cry V*	3,000	7,000
Fantasy	**All F games in the ViPiCo**	3,000	7,000
	Bloodborne	4,000	8,000
	Chrono Trigger	4,000	10,000
	Pokémon Black	3,000	5,000
Science Fiction	**All SF games in the ViPiCo**	3,000	5,000
	Halo Infinite	4,000	9,000
	Dead Space 2	3,000	7,000

Table 7
Summary of lexical coverage by setting.

The only two categories including more than one video game are Fantasy and Science Fiction, and within each of them considerable variation can be observed. Unfortunately, here the other categories are all represented by a single game each. However, we could try and argue that the settings fully grounded in the world as it has really existed at some point in time – i.e. the Historical and the Post-Apocalyptic settings – require both as few as 2,000 world families plus proper nouns and marginal words to reach 95% coverage, while the games that are fully or partially set in non-existing worlds – i.e. all the other games in the corpus – require 3,000-4,000 word families plus proper nouns and marginal words for the same threshold. This is the minimum number of word families needed to be able to guess the meaning of unknown words in context. For a better understanding of the contents of the dialogues, on the other hand, 98% lexical coverage is needed, and in this respect, there seems to be little correlation between number of word families and setting.

However, Table 7 presents an interesting fact: in the two categories comprising more than one game, the number of word families plus proper nouns and marginal words needed to reach 95% and 98% coverage is

Video game dialogue as input for language learning. An analysis of lexical coverage

259

systematically lower – and in the case of the 98% threshold also considerably lower – when we look at the whole set of games sharing the same type of setting (Fantasy – All, and Science Fiction – All), compared to the individual games that comprise the group. This suggests that the games in each of these setting categories share several word families. No such clear-cut picture can be observed when grouping games by type (Table 8).

Game type	Game title	95% coverage, reached at	98% coverage, reached at
Action-Adventure	**AAG sub-corpus**	2,000	5,000
	Resident Evil Village	3,000	5,000
	The Last of Us	2,000	5,000
	Devil May Cry V	3,000	7,000
Action	**AG sub-corpus**	3,000	8,000
	Dead Space 2	3,000	7,000
	Mafia Definitive Edition	2,000	8,000
	Halo Infinite	4,000	9,000
Role-Playing	**RPG sub-corpus**	3,000	7,000
	Bloodborne	4,000	8,000
	Chrono Trigger	4,000	10,000
	Pokémon Black	3,000	5,000

Table 8
Summary of lexical coverage by type.

This suggests that setting could be more relevant than game type in determining the words used in the dialogues, and that it may be easier for a learner to understand game dialogues if, when choosing a new game to play, they stick to the same type of setting. However, further investigation with a much larger and more balanced dataset is needed to confirm this hypothesis.

5. Discussion and conclusions

Our first aim in this study was to analyse the lexical coverage of the scripted language used by NPCs in the game. Our analyses show that, in order to understand the NPCs lines of the games in our corpus as a whole, a player should be familiar with the most frequent 3,000 word families to be able to guess unknown words in context (95% coverage) and as many as 7,000 word families to fully understand the dialogue (98% coverage). If we consider sub-corpora divided by genre, 95% coverage is reached with 3,000 word families in the AG and RPG sub-corpora, and as few as 2,000 word families in the AAG sub-corpus, while 98% coverage with as few as 5,000 word families for the AAG sub-corpus, 7,000 for the RPG one and as many as 8,000 for the AG one. However, like Rodgers and Heidt (2021), we observed ample internal variation in each (sub)corpus.

*L*ingue e
*L*inguaggi

Another aim was to assess whether game type and/or game setting have an impact on lexical coverage. Our data thus support Rodgers and Heidt's (2021) observation that lexical coverage differences largely depend on individual games rather than genre. Their findings were based on four game types, two of which also appear among the types analysed in this paper, but their corpus was not balanced for game type, and most types were represented by only two games. In our corpus, which is balanced by type, despite clear variation within each genre, dialogues in Action games as a whole would appear to be lexically slightly more challenging than in Role-Playing games, which in turn would seem to be more demanding than Action-Adventure game dialogues. This result suggests not to totally discard the possibility of genre-dependent differences and to investigate further using a much larger corpus. Indeed, the pedagogical lists of gaming vocabulary produced by Heidt *et al.* (2023) show that at least some kind of vocabulary found in games is genre-dependent.

Furthermore, our tentative analysis of games divided by setting suggests that setting does probably play a role in determining the words used in the dialogues. In particular, from a learner perspective, it suggests that it may be easier for a learner to guess unknown words when the game is grounded in reality and to understand game dialogues if, when choosing a new game to play, they continue with the same type of setting. Consequently, since many games come in series, with sequels and prequels it would be possible to hypothesise that to play titles from the same series – if they repeat the same setting, which is not always the case (e.g. the *Indiana Jones* series) – may represent an excellent way to scaffold language acquisition. This however should be verified on individual titles, as research on other types of scripted dialogue, namely telecinematic products, has observed ample lexical variation between episodes of the same series (Webb, Rodgers 2009b). It could also be interesting to investigate whether specific genre-setting configurations prove easier or more difficult for learners, as it has been observed that in films of specific genres-settings, such as horror and crime, dialogue is less lexically demanding than in others such as war (Webb, Rodgers 2009a). In any case, a corpus representing a much wider number of video games for each genre and setting and a wider variety of genres including more clear-cut ones (e.g. strategy games; simulation games; visual novels) is necessary to confirm or refine these hypotheses.

The third aim of our study was to see whether the lexical profile of in-game dialogue is somehow similar to that of other aural scripted dialogic discourse, such as movies and television. In this respect, our corpus, exclusively comprising in-game dialogue, reaches 95% and 98% lexical coverage with 3,000 and 7,000 word families respectively. These values are much lower than those identified by Rodgers and Heidt (2021) in a video game

Video game dialogue as input for language learning. An analysis of lexical coverage

261

corpus conflating both written and spoken text (5,000 and 10,000 word families for 95% and 98% coverage respectively). At the same time, our data are largely in line with those observed by Webb and Rodgers (2009a) on British and American films, where 95% coverage was reached with the most frequent 3,000 word families, and 98% coverage with 6,000-7,000 word families. This confirms the hypothesis of Rodgers and Heidt (2021, p. 222) that in-game dialogue resembles that of other spoken media. It also supports their hypothesis that the higher lexical demands observed in their analysis are likely to result from having included in the corpus user interfaces, tutorials, equipment inventories and the like. In fact, such types of text – performing a ludic function – include low frequency vocabulary such as *autosave*, *gameplay*, *overwrite*, *melee*, or *recoil* (Heidt *et al.* 2023).

The ViPiCo is definitely very small, limited in terms of number and genre of video games, and unbalanced in terms of settings, but we believe it nevertheless offered useful data. Besides the above conclusions, our data lead us to make a few more comments. First of all, the ViPiCo seems to challenge Domsch's (2017) hypothesis about the prominence of dialogue in the three genres considered (see Section 2.1). Ample variation can be observed between games within and across genres. Furthermore, if we consider main story length and calculate the average number of tokens per hour of gameplay, the genre with the greatest amount of dialogue appears to be the action-adventure one, with an average of 1,859 words uttered by NPCs per hour, immediately followed by the action genre, with 1,334 words per hour, and last comes the role-play genre, with as few as 705 words per hour of gameplay. The year when the game was published does not help in explaining this. Our best hypothesis is that the amount of dialogue (calculated in number of words) is determined by the multimodal construction of individual games. Indeed, the role of individual games has repeatedly come to the fore throughout this research, which is certainly a direct consequence of the complex and ever-evolving nature of video games. This should not however let scholars refrain from looking at groups of games sharing at least one feature in the search for common trends.

Furthermore, in order to better understand the nature of the language gamers are exposed to with in-game dialogue, it is important to go beyond lexical profiling and engage in semantic, syntactic and pragmatic analyses of NPC discourse, on its own and compared to other forms of spoken, dialogic discourse. In particular, corpus-linguistic investigations taking advantage of POS and semantic taggers could provide interesting insight into the topics game players encounter, which could then inform pedagogical uses of off-the-shelf video games. Something has been done in this direction (see Section 2.2), but several angles remain largely unexplored.

Finally, an area that deserves further research is that of video game paratext in relation to its capability to support preliminary acquisition of fundamental vocabulary prior to the actual gameplay, and to allow players with lower levels of knowledge of English to better engage with the game itself. Forums and game reviews are of course relevant, but the most interesting type of paratext could be walkthrough videos, and in particular monologic ones, as they have been observed to have simpler lexical profiles than dialogic ones (Chien 2019).

Bionote: Francesca Bianchi holds a PhD in English linguistics from Lancaster University (UK). She is associate professor of English linguistics and translation at the University of Salento (Lecce, Italy). Her research interests include corpus linguistics as an analytical method for the study of text types and forms of discourse, multimedia texts and audiovisual translation, and the use of technology in the teaching and learning of English, translation and interpreting.

Author's address: francesca.bianchi@unisalento.it

References

Adams E. 2010, *Fundamentals of Game Design*, 2nd ed., New Riders, Berclay, CA.

Álvarez-Bolado Sánchez C. and Álvarez de Mon I. 2019, *Video games: A Lexical Approach*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 13-38.

Anthony L. 2022, *AntWordProfiler*, version 2.0.1, [Computer Software]. Waseda University, Tokyo. http://www.laurenceanthony.net/software

Arjoranta J. 2019, *How do define games and why we need to*, in "The Computer Games Journal" 8, pp. 109-120.

Balteiro I. 2019, *Lexical and Morphological Devices in Gamer Language in Fora*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, p. 39-57.

Bauer L. and Nation P. 1993, *Word families*, in "International Journal of Lexicography" 6 [4], pp. 253-279.

Bergonse R. 2017, *Fifty Years on, What exactly is a video game? An essentialistic definitional approach*, in "The Computer Games Journal" 6 [4], pp. 239-255.

Bogost I. 2007, *Persuasive Games. The Expressive Power of Video games*, The MIT Press, Cambridge, MA.

Calafato R. and Clausen T. 2024, *Vocabulary learning strategies in extramural English gaming and their relationship with vocabulary knowledge*, in "Computer Assisted Language Learning", pp. 1-19. https://www.tandfonline.com/doi/full/10.1080/09588221.2024.2328023 (15.7.2024).

Campos-Pardillos M.A. 2019, *End-user Agreements in Video games: Plain English at Work in an Ideal Setting*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 116-135.

Chien Y.C. 2019, *The Language of Massively Multiplayer Online Gamers: A Study of Vocabulary in Minecraft Gameplay*, in "TESL-EJ" 23 [3], pp. 1-16. https://tesl-ej.org/pdf/ej91/int.pdf (8.7.2024).

De Wilde V., Brysbaert M. and Eyckmans J. 2019, *Learning English through out-of-school expo-sure. Which levels of language proficiency are attained and which types of input are important?*, in "Bilingualism" 23 [1], pp. 171-185.

De Wilde V. and Eyckmans J. 2017, *Game on! Young learners' incidental language learning of English prior to instruction*, in "Studies in Second Language Learning and Teaching" 7 [4], pp. 673-694.

deHaan J., Reed W.M. and Kuwada K. 2010, *The effect of interactivity with a music video game on second language vocabulary recall*, in "Language Learning & Technology" 14 [2], pp. 74-94.

Dixon D.H. 2022, *The Language in Digital Games: Register variation in virtual and real-world contexts*, PhD dissertation, Northern Arizona University, available from ProQuest Dissertations & Theses.

Dixon D.H. 2024, *Introducing the Single Player Offline Game Corpus (SPOC): A corpus of seven registers from digital role-playing games*, in "Corpora" 19 [1], pp. 107-122.

Domsch S. 2017, *Dialogue in video games*, in Mildorf J. and Thomas B. (Eds), *Dialogue across media*, John Benjamins, Amsterdam, pp. 251-270.

Ensslin A. 2012, *The Language of Gaming*, Palgrave Macmillan, London.

Ensslin A. and Finnegan J. 2019, *Bad Language and Bro-up Cooperation in Co-sit Gaming*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 139-156.

Fernández-Vara C. 2014, *Introduction to Game Analysis*, Routledge, London/New York.

Fizek S. 2012. *Pivoting the Player. A methodological toolkit for player character research in offline role-playing games*. ProQuest Dissertations Publishing, Bangor University, Bangor, UK. https://www.proquest.com/openview/385e0788aac018f08ac6021469e42fc1/1?pq-origsite=gscholar&cbl=51922 (12.7.2023).

Gee J.P. 2006, *Are video games good for learning?*, in "Nordic Journal of Digital Literacy" 1 [3], pp. 172-183.

Gee J.P. 2008, *Learning theory, video games and popular culture*, in Drotner K. and Livingstone S. (Eds.), *The International Handbook of Children, Media and Culture*, SAGE Publications, London, pp. 196-213.

Gee J.P. 2014, *Unified Discourse Analysis: Language, Reality, Virtual Worlds and Video Games*, Routledge, London/New York.

Gledhill C. 2019, *Phraseology and Lexico-grammatical Patterns in Two Emergent Paragame Genres: Video game Tutorials and Walkthroughs*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 58-86.

Graham S.L. and Dutt S. 2019, *"Watch the Potty Mouth": Negotiating Impoliteness in Online Gaming*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 201-223.

Hancock M. 2019, *Playing It By the Book: Instructing and Constructing the Player in the Video game Manual Paratext*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 288-303.

Heidt J., Pinchbeck G.G. and Rodgers M.P.H. 2023, *The Good Gaming (GG) list: Key vocabulary in video games*, in Reynolds B.L. (Ed.), *Vocabulary Learning in the Wild*, Springer, Singapore, pp. 143-161.

Heritage F. 2020, *Applying corpus linguistics to video game data: Exploring the representation of gender in video games at a lexical level*, in "*Game studies*" 20 [3], n.p.. https://gamestudies.org/2003/articles/heritage_frazer (19.6.2023).

Heritage F. 2022, *Magical women: Representations of female characters in the Witcher video game series*, in "Discourse, Context & Media" 49, 100627, pp. 1-9.

Hu M. and Nation P. 2000, *Vocabulary density and reading comprehension*, in "Reading in a Foreign Language" 13 [1], pp. 403-430.

Iaia P. 2016, *Analysing English as a Lingua Franca in Video Games. Linguistic Features, Experiential and Functional Dimensions of Online and Scripted Interactions*, Peter Lang, Berlin.

Iaia P. 2021, *Towards a 'COOPING' Model for the Investigation of Gamers' Online Conversations in English*, in "Iperstoria" 17, pp. 222-240.

Kiourti E. 2019, *"Shut the Fuck up Re! Plant the Bomb Fast!": Reconstructing Language and Identity in First-person Shooter Games*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 157-177.

Kuppens A.H. (2010), *Incidental foreign language acquisition from media exposure*, in "Learning, Media and Technology" 35 [1], pp. 65-85.

Laufer B. 1989, *What percentage of text lexis is essential for comprehension?*, in Lauren C., Nordman M. (Eds.), *Special language: From humans thinking to thinking machines*, Multilingual Matters Clevedon, England, pp. 316-323.

Lee J.H., Karlova N., Clarke R.I., Thornton K. and Perti A. 2014, *Facet Analysis of Video Game Genres,* Paper presented at iConference, Berlin, Germany.

Leech G., Rayson P. and Wilson A. 2001, *Word Frequencies in Written and Spoken English*, Longman, Harlow.

Liu N.A. and Nation P. 1985, *Factors affecting guessing vocabulary in context*, "RELC Journal" 16, pp. 33-42.

Meara P. and Buxton B. 1987, *An alternative to multiple choice vocabulary tests*, in "Language Testing" 4 [2], pp. 142-151.

Meara P. and Jones G. 1988, *Vocabulary size as a placement indicator*, in Grunwell P. (Ed.), *Applied Linguistics in Society*, CILT, London, pp. 80-87.

Muñoz C. 2020, *Boys like games and girls like movies. Age and gender differences in out-of-school contact with English*, in "Revista Española de Lingüística Aplicada" 33 [1], pp. 172-202.Nation P. 2006, "How large a vocabulary is needed for reading and listening?", *Canadian Modern Language Review* 63 [1], pp. 59-82.

Nation P. and Anthony L. 2017, *Measuring vocabulary size*, in Hinkel E. (Ed.), *Handbook of Research in Second Language Teaching and Learning*, vol. III, Routledge, London/New York, pp. 355-368.

Nation P. and Heatley A. 2002, *Range: A program for the analysis of vocabulary in texts*, software. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-analysis-programs (1.4.2023).

Paul C.A. 2012, *Wordplay and the Discourse of Video Games. Analysing Words, Design and Play*, Routledge, New York.

Peters E. 2018, *The effect of out-of-school exposure to English language media on learners' vocabulary knowledge*, in "ITL – International Journal of Applied Linguistics" 169 [1], pp. 142-167.

Reinhardt J. 2018, *Gameful second and foreign language teaching and learning: Theory, research, and practice*, Springer, New York.

Rød A.J., and Calafato R. 2024, *Exploring extramural English: Impacts, integration and future directions*, in Skulstad A.S. (Ed.), *Current issues in English teaching*, Fagbokforlaget, pp. 117-140.

Rodgers M.P.H. and Heidt J. 2021, *Levelling up comprehensible input and vocabulary learning. The lexical profile of video games*, in Werner V., Tegge F. (Eds.), *Pop Culture in Language Education. Theory, research, practice*, Routledge, London, pp. 215-227.

Rudge L.A. 2019, *"I Cut It and I ... Well Now What?": (Un)Collaborative Language in Timed Puzzle Games*, in Ensslin A., Balteiro I. (Eds.), *Approaches to Video game Discourse. Lexis, interaction, textuality*, Bloomsbury, New York, pp. 178-200.

Schmitt N. 2019, *Understanding vocabulary acquisition, instruction, and assessment: A research agenda*, in "Language Teaching" 52 [2], pp. 261-274.

Stamenković D. 2023, *The stylistic journey of a video game. A diachronic approach to multimodality in the Football Manager series*, in Schubert C. and Verner W. (Eds.), *Stylistic Approaches to Pop Culture*, Routledge, New York, pp.227-246.

Stamenković D. and Jaćević M. 2019, *Video Games and Multimodality: Exploring Interfaces and Analyzing Video Game Screens Using the GeM Model*, in Wildfeuer J. Pflaeging J., Bateman J., Seizov O. and Tseng C.-I (Eds.), *Multimodality. Disciplinary Thoughts and the Challenge of Diversity*, De Gruyter, Berlin, pp.277-294.

Stamenković D., Jaćević M. and Wildfeuer J, 2016, *The persuasive aims of Metal Gear Solid: A discourse theoretical approach to the study of argumentation in video games*, in "Discourse, Context and Media" 15, pp. 11-23.

Sundqvist P. and Wikström P. 2015, *Out-of-school digital gameplay and in-school L2 English vocabulary outcomes*, in "System" 51, pp. 65-76.

Sylvén L.K. and Sundqvist P. 2012, *Gaming as extramural English L2 learning and L2 proficiency among young learners*, in "ReCALL" 24 [3], pp. 302-321.

Webb S. and Rodgers M.P.H. 2009a, *The lexical coverage of movies*, in "Applied Linguistics" 30 [4], pp. 407-427.

Webb S. and Rodgers M.P.H. 2009b, *Vocabulary demands of television programs*, in "Language Learning" 59 [2], pp. 335-366.

Werner V. 2021, *Text-linguistic analysis of performed language: Revisiting and remodeling Koch & Oesterreicher*, in "Linguistics" 59 [3], pp. 541-575.

Werner V. 2022, *Pop Cultural Linguistics*, in *Oxford Research Encyclopedia of Linguistics*, Oxford University Press, online. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-999 (8.7.2024).

Winget M.A. 2011, *Video game preservation and massively multiplayer online role-playing games: A review of the literature*, in "Journal of the American Society for Information Science and Technology" 62 [10], pp. 1869-1883.

COMPLEXITY MATTERS IN TV DIALOGUE AS LANGUAGE INPUT
A corpus-based description of clausal and phrasal patterns[1]

MAICOL FORMENTELLI[1], RAFFAELE ZAGO[2]
[1]UNIVERSITÀ DI PAVIA, [2]UNIVERSITÀ DI CATANIA

Abstract – The spread of digital devices and new media has led to a proliferation of English-language audiovisual products and to a gradual shift in learning English as a second language (L2), which is increasingly occurring outside of the classroom. In fact, the extensive informal contact with English audiovisual dialogue as part of L2 learner-users' everyday leisure activities has been found to be potentially conducive to incidental learning and language acquisition (Kusyk 2020; Sockett 2014). This calls for a description of English-language media, the major source of language input to which L2 viewers are exposed (Pavesi, Ghia 2020). The present study aims to investigate grammatical complexity in fictional TV dialogue by drawing on the *Sydney Corpus of Television Dialogue* (SydTV; Bednarek 2018a). A register-functional (RF) approach (Biber 1988) is adopted for the quantitative and qualitative assessments of clausal and phrasal complexity features in TV series, with a focus on finite/nonfinite subordinate clauses and noun phrase premodification. The data are first interpreted in relation to the diegetic and extradiegetic functions served by complexity features onscreen and are then compared to previous corpus-based findings regarding the grammatical complexity of spontaneous face-to-face conversations (Biber 2015; Biber *et al.* 2021). The results show that TV dialogue closely approximates casual conversation in terms of the main patterns of phrasal and clausal complexity. Register-specific functions emerge in accordance with the audience-oriented narrative dimension of telecinematic products, the striving for realism and the expression of characters' stances. Such referential and communicative functions may increase the accessibility of TV narratives for L2 viewers, often in tandem with visuals in a multimodal fashion. Overall, by reproducing the complexity of conversational exchanges, TV dialogue qualifies as a rich, reliable source of input that learner-viewers can readily use as a model of spoken English.

Keywords: Informal learning; TV dialogue; grammatical complexity; subordinate clauses; phrasal premodification.

---

[1] Maicol Formentelli authored the following sections/subsections: 1. Introduction; 2. Complexity matters; 5.1. Clausal complexity; 5.2. The functions of clausal complexity features in TV dialogue. Raffaele Zago authored the following sections/subsections: 4. Data and methodology; 5.3. Phrasal complexity; 5.4. The functions of phrasal complexity features in TV dialogue; 6. Concluding remarks. Maicol Formentelli and Raffaele Zago co-authored the following sections/subsections: 3. Complexity in telecinematic dialogue; 5.5. Comparison with spontaneous conversation.

1. Introduction

The spread of digital devices and new media has led to a proliferation of English-language audiovisual products being distributed across the globe, which has enabled learner-users of English as a second language (L2) to have extensive informal contact as part of their everyday leisure activities. This phenomenon has also been accompanied by a gradual shift in learning L2 English, which is increasingly occurring outside of the language classroom, since prolonged exposure to rich, authentic input such as audiovisual dialogue (Pavesi 2015) is potentially conducive to incidental learning and language acquisition (Arnbjörnsdóttir, Ingvarsdóttir 2018; Kusyk 2020; Pavesi, Ghia 2020; Sockett 2014). The description of audiovisual dialogue thus becomes essential to better understand the type of language that L2 English learner-users encounter while watching audiovisual products.

The present study attempts to accomplish this task by adopting a register-functional (RF) approach (Biber 1988) with a focus on the grammatical complexity of onscreen language, a major dimension of register variation that captures the structural and functional sophistication of texts at the clausal and phrasal levels (Biber *et al.* 2022). The study is a follow up of a recent investigation of the grammatical complexity of Anglophone film dialogue (Formentelli *et al.* forthcoming – see Section 2 below), which inspired the application of the methodology to the study of fictional TV dialogue.

TV series and sitcoms in English are watched by vast international audiences and are amongst the most appreciated types of popular culture (cf. Werner 2018). At present, fictional TV programmes are easily available in large quantities and in different genres via traditional cable TV, as well as via video-on-demand streaming platforms. L2 learner-users are exposed to copious amounts of spoken English input when watching TV series; such input is chosen autonomously, according to one's preferences and for extended periods given the delivery of these audiovisual products as episodes and seasons.

By drawing on a corpus of English-language TV series, we aim to provide quantitative and qualitative descriptions of clausal and phrasal complexity features occurring in TV dialogue, including the main types of finite and nonfinite dependent clauses and patterns of noun phrase premodification, which will be interpreted in relation to the communicative functions they serve in the register. As a second step in the analysis, the distribution of complexity features in TV dialogue will be discussed in comparison to corpus-based findings regarding the complexity of spontaneous face-to-face conversation to assess how and the extent to which the two registers of spoken language compare. The goal is to account for the frequency

*L*ingue e
*L*inguaggi

and types of grammatical complexity features in TV dialogue, as these key language features may play a role in L2 development.

2. Complexity matters

The study of complexity has a long tradition in linguistics, and several definitions, theoretical frameworks and analytical approaches have been proposed for this multifaceted concept (see Arends 2001; Biber 1992; Biber *et al*. 2022, 2023; Bulté, Housen 2012; Miestamo *et al*. 2008; Nichols 2009; Pallotti 2015, amongst others). By adopting a strict linguistic focus, scholars have differentiated between complexity and the concept of processing difficulty – more cognitive in nature; complexity has been defined quantitatively in terms of the number of grammatical distinctions in a language system (Szmrecsanyi 2015) or the number of different components in a sentence/utterance and the interconnections amongst these components (Bulté, Housen 2012; Pallotti 2015). More linguistic material is often regarded as a sign of complexity: The longer the expression, the higher the degree of complexity. Similarly, clauses and phrases featuring the embedding of optional constituents are generally considered to be more complex compared to basic structures that only contain obligatory elements (Biber *et al*. 2023).

Along these lines, research in applied linguistics has employed omnibus measures to predict L2 proficiency and development by assessing complexity through a combination of multiple structural aspects computed as single quantitative variables, such as the mean length of sentences/utterances, the average number of clauses per sentence/utterance and coordination index (Bulté, Housen 2012, 2014; Norris, Ortega 2009). This view is based on the theoretical assumption that complexity is a unitary concept that is realised through patterns that are common to all texts and registers.

The RF approach to complexity is slightly different, as it aims to capture the internal variability of grammatical complexity strategies associated with changing communicative functions and production circumstances across registers (Biber *et al.* 2022). In RF research, grammatical complexity is conceived as a multidimensional construct that is manifested in the co-occurrence of phrasal and clausal features that "pattern together in texts and vary in systematic ways across registers" (Biber *et al.* 2023, p. 5). These quantitative dimensions reflect the rate of occurrence of groupings of phrasal/clausal elements and allow for a more detailed linguistic description of the system of complexity features of English, as they account for variation in structural types and syntactic functions within and across texts (Biber *et al.* 2020).

Numerous RF corpus-based studies have revealed systematic differences in the manifestation of grammatical complexity in spoken and written registers, prototypically represented by spontaneous conversation and academic writing

*L*ingue e *L*inguaggi

(see Biber *et al.* 2022). The complexity of speech is mainly due to the structural elaboration of the utterance and the addition of dependent clauses as clause constituents (adverbial clauses and verb-controlled complement clauses), in line with the property of intricacy of spoken language that Halliday (1989) described. Conversely, written texts mainly express grammatical complexity through the embedding of constituents at the phrasal level; that is, pre- and postmodification in noun phrases by means of attributive adjectives, premodifying nouns and postmodifying prepositional phrases (Biber *et al.* 2020, p. 8).

These distinctive realisations of grammatical complexity have been linked to the specific production circumstances and communicative functions of spoken and written registers. The real-time, unplanned production that is typical of most conversational exchanges favours the frequent use of structurally elaborate dependent clauses, which are also used to express personal stances and attitudinal meanings, as well as to establish common ground amongst interlocutors, which are central aspects of the interpersonal and involved character of face-to-face interactions (Biber 2015). Conversely, the written mode allows for careful planning, revision and editing, thus enabling the use of features of both clausal and phrasal complexity (Biber *et al.* 2022, p. 463). Phrasal complexity is particularly pervasive in written texts that have informational communicative purposes, as phrasal modifiers enable users to compress the informational content into expressions that are denser and more concise.

Based on these observations, we adopted the RF approach to complexity to address Biber *et al.*'s (2022, p. 481) claim that "one major priority for future [...] research is more detailed studies of complexity characteristics in a wide range of spoken registers". In this regard, a very recent study has approached this task by assessing the linguistic expression of grammatical complexity in a corpus of Anglophone film dialogue (Formentelli *et al.* forthcoming). In the present contribution, we follow the same theoretical perspective and methodology to focus on another major type of audiovisual dialogue, namely fictional TV series, a register that also originates as written language, is devised to simulate spontaneous conversation and is ultimately processed aurally as spoken language.

3. Complexity in telecinematic dialogue

An examination of Bednarek and Zago's (2024) latest updated bibliography of linguistic research on fictional TV series and films reveals that telecinematic dialogue has been the subject of numerous studies that have explored a range of linguistic phenomena using various methods and from different perspectives. However, despite the wealth of studies in this area, little attention has been paid to assessing the complexity of the language used on the screen thus far.

The lexical complexity of film and TV dialogue has been described in terms of lexical density, frequency, variation and sophistication (Formentelli

2014; Jones 2017; Scheffler *et al.* 2020; Webb, Rodgers 2009a, 2009b). In particular, Webb and Rodgers (2009a, 2009b) found that the knowledge of the 3,000 most frequent English word families provided learner-viewers with 95% of the necessary lexical coverage to understand British and American films and TV programmes in various genres, while the knowledge of 6,000 to 7,000 word families was needed to attain 98% coverage. Similarly, by analysing a corpus of British animated TV series for preschool children, Scheffler *et al.* (2020) found that the 2,000 most common words in the spoken component of the *British National Corpus* covered more than 80% of the words in the dialogues in these series. These findings show a link between lexical complexity and language proficiency, and suggest that telecinematic dialogue can be an appropriate source of comprehensible L2 input for young learners (Scheffler *et al.* 2020) and adults (Webb, Rodgers 2009a, 2009b), potentially leading to incidental vocabulary learning.

Onscreen complexity has also been examined by Zago (2019), who illustrated how different combinations of lexicogrammatical features frequently occur in American film dialogue with register- and genre-specific functions. For instance, activity verbs, prepositions and progressive aspect are used to create the urgent, action-oriented quality that is typical of crime films. In particular, cinematic speech is complex in the sense that it follows two parallel paths: It constantly has to advance the storyline while simultaneously simulating real speech. Zago cited utterance launchers (such as *the thing is*) as an example of the latter type of complexity; in film dialogue, these linguistic features mimic casual conversation while simultaneously functioning as discursive cues of a character's intention to speak.

The language of contemporary American TV programmes has also been described by applying a multidimensional analysis framework (Biber 1988) to a corpus of 31 different TV registers, including films, drama series, sitcoms and soap operas (Berber Sardinha, Veirano Pinto 2019). The authors identified clusters of lexicogrammatical features that combined along four dimensions of register variation, namely 'Exposition and discussion versus Simplified interaction' (Dimension 1), 'Simulated conversation' (Dimension 2), 'Recount' (Dimension 3) and 'Engaging presentation' (Dimension 4). Dimensions 2 and 3 are particularly relevant for fictional TV programmes, as they reflect the striving for realism and the narrative needs that are typical of the genre, and confirm that "television is essentially an oral, involved, stance marked form of spoken language, close in some ways to but not exactly like face-to-face encounters" (Berber Sardinha, Veirano Pinto 2019, p. 26; see also Bednarek 2010; Quaglio 2009). Even though complexity was not mentioned explicitly in their study, the interplay of numerous syntactic constructions at the phrasal and clausal levels testifies to the structural elaboration and grammatical intricacy of telecinematic language.

More recently, Formentelli *et al.* (forthcoming) applied the RF approach to the study of the grammatical complexity of film language, as expressed in the occurrence of finite and nonfinite dependent clauses. By drawing on a corpus of 34 transcribed Anglophone films, the distribution of clausal complexity features was first assessed in relation to the situational characteristics and communicative functions of the filmic register and were then compared to the findings of previous research on complexity in spontaneous conversation; the goal was to describe film dialogue as relevant language input for L2 acquisition. The results corroborate the close similarity between film dialogue and spontaneous spoken language observed in previous research (Forchini 2012) by extending the evidence to clausal complexity. At the same time, the results also point to distinctive patterns of complexity in films linked to register-specific communicative functions and medium-related constraints. The authors conclude that film dialogue is a useful source for L2 English learners inside and outside of the classroom; these learners can expose themselves to realistic language patterns that would only otherwise be experienced when engaging in face-to-face exchanges (Pavesi 2012, 2015). In the present contribution, we adopted the same theoretical framework and methodology and applied it to the investigation of clausal and phrasal complexity in a corpus of TV series, as described in the following section.

4. Data and methodology

The present analysis of grammatical complexity was conducted based on *The Sydney Corpus of Television Dialogue* (Bednarek 2018a; henceforth SydTV). SydTV is a specialised and POS-tagged corpus of contemporary American TV dialogue (276,899 words; cf. Bednarek 2018b). It consists of one full episode from the first season of 66 American TV series that were broadcast between 2000 and 2012, a period that "was characterised by the global rise of American TV series […] and has been labelled the new 'golden age of television'" (Bednarek 2018a, p. 82). The corpus contains a balanced combination of episodes occurring towards the beginning, the middle and the end of the respective seasons and consists of both quality series – based on Emmy or Golden Globe award nominations or wins, and mainstream series. The corpus is also balanced in terms of genres; approximately half of the series in SydTV can be labelled 'comedies' and the other half 'dramas', including hybrid combinations such as action/drama or comedy/crime.

The patterns of grammatical complexity considered in this study were selected following Biber *et al.*'s (2022, pp. 13-16) taxonomy of grammatical complexity features in English. We examined both clausal and phrasal complexity patterns. Clauses were categorised according to their syntactic

Complexity matters in TV dialogue as language input. A corpus-based description
of clausal and phrasal patterns

273

functions, namely complement clauses, adverbial clauses or relative clauses,
as well as according to their finiteness, that is, finite versus nonfinite clauses.
With regard to phrasal complexity, we examined both adjectival
premodification with up to three attributive adjectives and noun
premodification with up to three nouns, including genitives.

In the first stage of the analysis, which was conducted semi-
automatically using the online version of SydTV[2], specific POS-tag queries
were employed to identify all the occurrences of the aforementioned clausal
and phrasal complexity features. For instance, the POS-tag sequence *_NN*
_NN was used to capture the noun phrases that were premodified by one
noun (*phone call*, *baseball fan*, *photo album*, *wine list*, *lunch break*, etc.). The
occurrences thus obtained were checked manually to exclude tagging errors
and duplicates and were then counted. Finally, illustrative examples of all the
structural and syntactic patterns were selected from the concordance lists and
were subjected to a qualitative analysis in context with a view to understanding
the main register-specific functions that they served in SydTV. Samples of the
data were first coded and interpreted individually by the two authors and were
then discussed jointly to reach the necessary degree of agreement.

In the second stage of the analysis, the quantitative counts obtained from
SydTV were compared to those reported for conversation in Biber *et al.* (2021)
and Biber (2015) in order to assess the degree of similarity between TV dialogue
and spontaneous, casual speech at the level of grammatical complexity.

5. The clausal and phrasal complexity of American TV dialogue

The main empirical findings are reported and discussed from quantitative,
qualitative and comparative perspectives in this section. In particular,
Subsections 5.1. and 5.2. first address the frequency of clausal complexity
features in SydTV followed by the register-specific functions they serve, while
Subsections 5.3. and 5.4. address phrasal complexity in the same manner.
Finally, Subsection 5.5. presents a comparison of the frequency and types of
clausal and phrasal complexity features that occurred in SydTV and in casual
conversation.

5.1. Clausal complexity

The investigation of the patterns of subordination identified in SydTV provided
a quantitative insight into the clausal complexity structures that are more
frequent in American TV series and are more likely to be encountered by

---

[2] https://cqpw-prod.vip.sydney.edu.au/CQPweb/index.php (15.3.2024).

English L2 learner-users in their informal contact with the language. Figure 1 presents the overall picture of the distribution of clauses across the three macrocategories of complement clauses, adverbial clauses and relative clauses in their finite and nonfinite realisations. Complement clauses are by far the most common subordinative structures in TV dialogue, accounting for slightly more than 60% of the occurrences, and are almost equally divided into finite and nonfinite clauses (1,252 versus 1,095 per 100,000 words). Adverbial clauses follow with 27% of occurrences and a preponderance of finite over nonfinite clauses (958 versus 87 per 100,000 words). The least frequent category is that of relative clauses, accounting for approximately 13% of the subordinative clauses, mainly with finite rather than nonfinite verbs (404 versus 81 per 100,000 words).

The frequencies that are summarised in Figure 1 indicate that grammatical complexity at the clausal level is typically realised through verbal attachment rather than through noun modification. Moreover, the preponderance of finite over nonfinite clauses shows a preference for explicit constructions to encode the meaning relations of propositions, thus producing discourse that is more accessible and easier for viewers to process, in line with the tendency to reduce vagueness that has been observed in TV dialogue (Quaglio 2009).
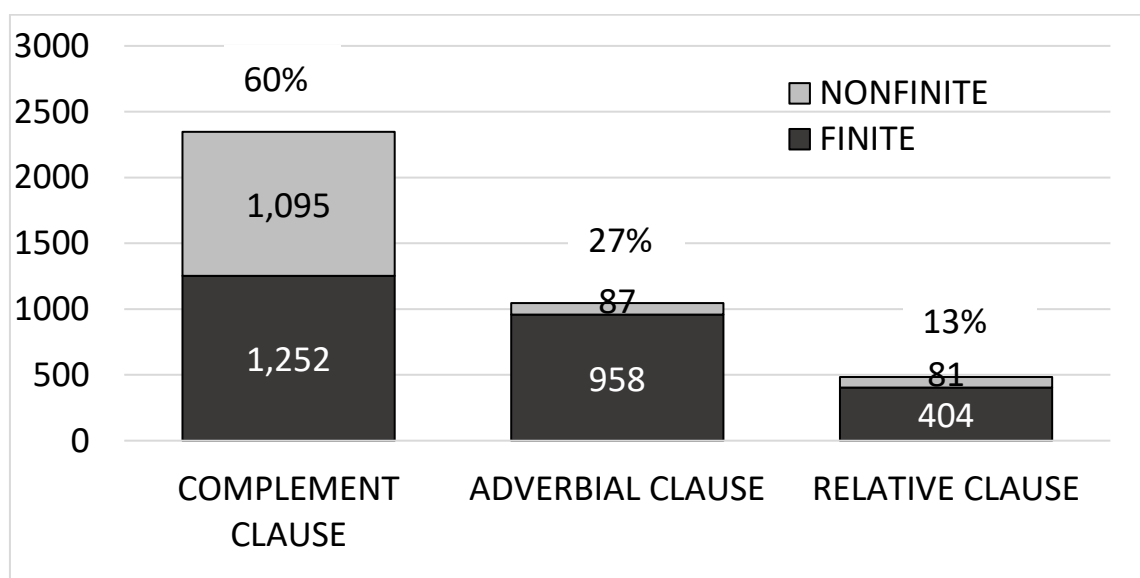
Figure 1
Finite and non-finite subordinate clauses in SydTV (frequency per 100,000 words).

The quantitative profile of the clausal complexity of TV dialogue is complemented by the frequencies of the structural/syntactic subtypes included in the three macrocategories of clauses, as shown in Figure 2.

Complexity matters in TV dialogue as language input. A corpus-based description
of clausal and phrasal patterns
275
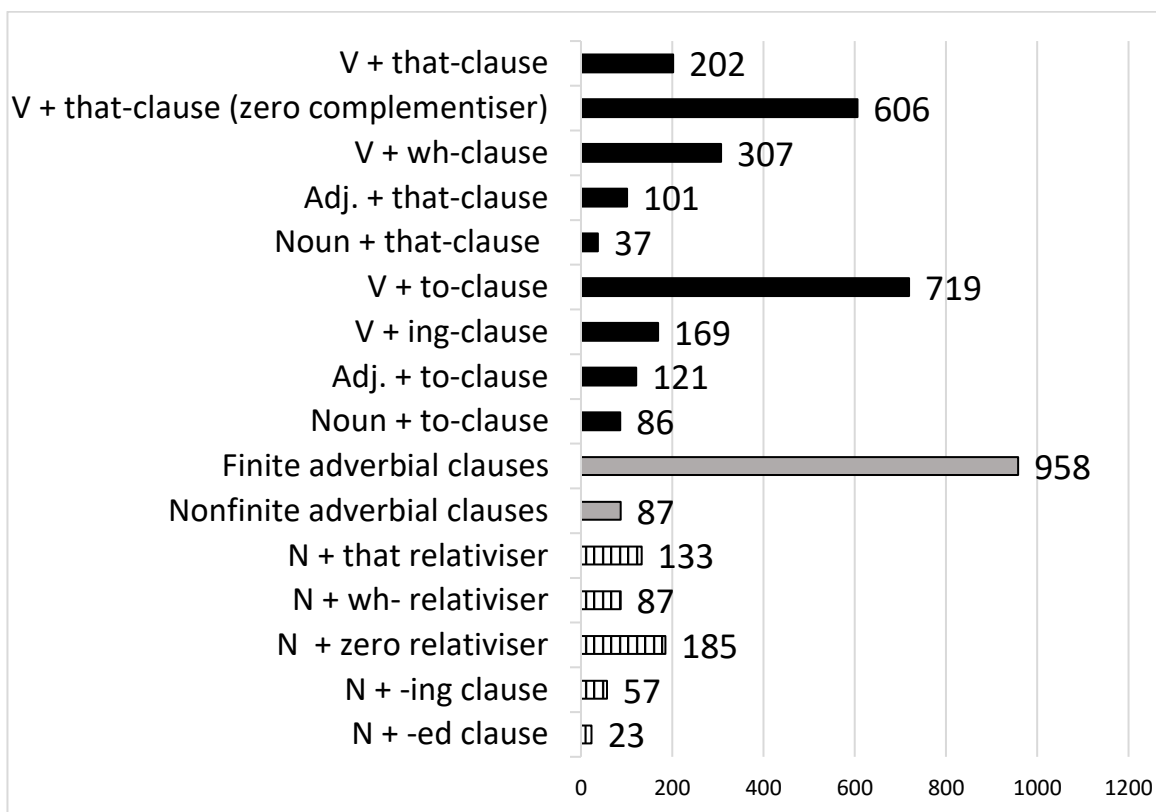
Figure 2
Structural/syntactic subtypes of clauses in SydTV (frequency per 100,000 words).

The category of complement clauses encompasses the largest number of structural/syntactic subtypes, including clauses controlled by verbs, adjectives and nouns. In SydTV, the vast majority of both finite and nonfinite complement clauses are introduced by verbs of cognition and communication (*think*, *say*, *know*, *find*, *believe*), while the frequency of adjective-controlled (*sure*, *glad*, *afraid*) and noun-controlled (*fact*, *idea*, *chance*) complement clauses is much lower. These distributional patterns reflect those identified for the register of spontaneous conversation (Biber *et al.* 2021, p. 668). Verb-controlled *that*-clauses (808 occurrences per 100,000 words) are particularly frequent, and are commonly used to report the speaker's or a third party's speech, thoughts, attitudes or emotions (see Section 5.2 below). In TV series, *that*-clauses are preferably realised with the omission of the *that* complementiser to simulate the online production of spontaneous speech and its involved and interpersonal nature (Biber *et al.* 2021, p. 674). Moreover, *that*-clauses with zero complementiser are associated with reduced discourse complexity at the structural level (Biber *et al.* 2022, pp. 54-55) and promote the economy of speech required by the storytelling constraints of TV dialogue.

Postpredicate infinitive clauses controlled by verbs (V + *to*-clause) are the second most frequent type of complement clauses found in SydTV, with 719 occurrences per 100,000 words. In spontaneous conversation, these

clauses are introduced by verbs expressing the speaker's personal desires, intentions and efforts (*want*, *like*, *try*) and, to a more limited extent, aspectual verbs (*begin*, *continue*) and verbs of probability (*seem*). The high frequency of *to*-clauses in TV series is not surprising, as these verbal combinations may play key roles at the narrative level, adding dynamicity to the dialogue and advancing the plot (see Section 5.2 below).

Another relevant type of complement clause in the data is that of *wh*-clauses introduced by verbs of cognition (*know*, *wonder*), communication (*tell*, *say*) and perception (*see*). These constructions are strategic for the contextualisation of an event by conveying information about the participants, the spatial and temporal setting, the reasons and manners in the form of indirect questions, exclamations and nominal clauses. This finding confirms the strategic narrative function of *wh*-complement clauses that was recently identified in Anglophone film dialogue (Formentelli *et al*. forthcoming).

The category of finite adverbial clauses comprises a wide variety of subordinate clauses that, combined, account for 958 occurrences per 100,000 words. Conditional, causal, temporal and result clauses are particularly frequent in SydTV, as shown by the occurrences of subordinatiors summarised in Table 1, in line with the distribution of adverbial clauses in spontaneous conversation (Biber *et al.* 2021, p. 814).

Subordinator	Raw frequency	Normed frequency (per 100,000 words)
(Even) If	797	288
Because/Cause	532	192
When/whenever	423	153
So (that)	193	70
As	98	35
Before	92	33
Until	91	33
While	63	23
Since	60	22
Than	54	20
After	45	16
Unless	36	13
Like	32	11

Table 1
Most frequent subordinators in SydTV.[3]

These types of clauses provide the interlocutor with the necessary circumstantial information to fully understand the unfolding interaction; in TV dialogue, they are particularly important to provide the viewing audience with background knowledge about the characters and events that are represented onscreen (Formentelli *et al*. forthcoming). Relevant, yet far less frequent, are

---

[3] Only subordinators with a frequency > 10 forms per 100,000 words are included in Table 1.

nonfinite clauses (87 per 100,000 words), which mainly codify *to*-infinitive purpose clauses.

Finally, relative clauses show overall low frequencies in the TV series included in SydTV. Amongst them, finite clauses with zero relativiser and *that*-relative clauses account for most of the occurrences in this category (185 and 133 occurrences per 100,000 words, respectively), hinting at the informal and colloquial nature of the interactions represented onscreen. The marked preference for the omission of the relative pronoun is consistent with the high frequency of zero complement clauses observed above.

5.2. The functions of clausal complexity features in TV dialogue

Several types of subordinate clauses that contribute to grammatical complexity in TV dialogue were found in SydTV. The following qualitative analysis complements the quantitative findings and enriches the picture by describing the communicative functions served by different clauses at the diegetic and extradiegetic levels.

One main function of the subordinate clauses in the data is the expression of the speakers' personal stances, which is central to the interpersonal character of face-to-face conversation and the onscreen reproductions. Complement clauses are the preferred devices to encode epistemic evaluations and attitudinal meanings in English (Biber 2015), establishing a link between syntactic complexity at the clausal level and this communicative function. The high frequency of both finite and nonfinite complement clauses in SydTV is in line with the findings of previous studies of American films and television registers, which identified these types of clauses as a major source of the naturalness of onscreen dialogue, scene dramatisation and audience entertainment (Berber Sardinha, Veirano Pinto 2019; Formentelli *et al.* forthcoming; Veirano Pinto 2014; see also Bednarek 2011). The brief excerpts of TV dialogue in Examples 1 and 2 show a series of complement clauses that resonate across characters, foregrounding their points of view by means of the first person pronoun *I* and verbs that convey various degrees of epistemic certainty (*I think*, *I had no idea*), as well as strong positive and negative attitudes (*I love*, *I'd love*, *I hate*). The quick exchanges of ideas about events and states of affairs contribute to the spontaneity of the fictional interactions and simultaneously engage the viewing audience, prompting their interpersonal (dis)alignment and emotive participation.

*L*ingue e *L*inguaggi

1) Suits_std_1_10 I **think** it'd be nice to be in nature, don't you Mike? <THERESE> Well I **love** that you're a lawyer. In fact, I 'd **love** to see where you work some time.[4]

2) Gossip_Girl_std_1_17 but I **had no idea** their standards were so low. <DAN> I **hate** that I have to ask you this, but have you seen Serena?

In addition to the expression of stance, nonfinite complement clauses are used in TV series to add dynamicity to the dialogue through the concise combination of verbs in modal, aspectual and temporal lexico-syntactic constructions. In Examples 3 and 4, the deontic verbal expressions *Do you want*, *I want* and *I need you* introduce the subordinate clauses by expressing the characters' strong desires and needs, thus projecting the actions encoded in the propositional content into the future and advancing the plot.

3) How_I_Met_Your_Mother_std_1_12 <LILY> Okay, okay, sweetie, we, just calm down. **Do you want** to go somewhere and talk? <CLAUDIA> **I want** to go somewhere and drink. <LILY> Okay, meet me at MacLaren's. I I'll see you there.

4) Fringe_std_1_13 **I need you** to go to the passengers, **I need you** to collect as many sedatives or tranquilizers as you can. <MELISSA> Okay, Sir, **I need you** to take a deep breath.

Similarly, nonfinite complement clauses often combine with imperative forms in directive face-threatening acts and amplify their pragmatic force, particularly in the conflictual situations that abound in fictional TV series (Bednarek 2012). The use of the swearword *the hell* in Example 5 and the dramatic climax in Example 6, achieved through Kate's aggressive repetition of the verb *stop* followed by two complement clauses, charge the dialogue with a wave of negative emotions that impact on the fictional interlocutors and the spectators alike. This strong, concurrent orientation towards onscreen characters and viewers contributes to heightening the involvement of the latter, who are the ultimate addressees in the mediatised communicative event (Dynel 2011).

5) Lost_std_1_17 <LOCKE> Should I be writing this down? <SHANNON> Just, just **tell him** to stay the hell out of my business. <LOCKE> Do you like him ? <SHANNON> What?

6) United_States_of_Tara_std_1_08 <KATE> Okay , stop it. <GENE> Stop what, Kitty Kate? <KATE> **Stop** with the name. **Stop** singling me out at your little crab updates. **Stop** blabbing to people. Just **stop** it.

With regard to the narrative dimension of TV dialogue, both *wh*-complement clauses and adverbial clauses are used to make relations between facts and people explicit, and facilitate the audience's task of reconstructing the storyworld and understanding the plot that is unfolding onscreen. Examples 7 and 8 illustrate how *wh*-complement clauses can indicate that relevant

---

[4] Each example is introduced by a codified string with metadata on the TV series from which the extract of dialogue is taken: TV series name_standardised transcription_season number_episode number (e.g., Suits_std_1_10)

contextual information is being delivered by the characters in the on-going interaction, such as Jess' whereabouts, professional status or the details of a criminal investigation.

7) Girls_std_1_03 me change here Shoshanna. **Do you know** <u>where Jess is</u>? **Do you know** <u>when she'll be home</u>? <SHOSHANNA> Umm, probably not until late. She has a job now.

8) Human_Target_std_1_11 <WINSTON> It **explains** <u>why they wanted to keep Tony alive</u>. Plant the murder weapon in his hand. All these love letters in the ambulance make him look like a stalker.

Similarly, in Example 9, the hypothetical clauses that the two characters ironically exchange concerning the theme 'killing the partner' allow the viewers to frame Gabrielle and Carlos' romantic relationship, while the causal and the purpose clauses uttered by Charlie in Example 10 express the speaker's annoyance with Ed's behaviour and simultaneously describe an aspect of his personality.

9) Desperate_Housewives_std_1_19 <GABRIELLE> One more thing. <u>If you ever hurt me again</u>, I will kill you. <CARLOS> <u>If you ever leave me for another man</u>, I'll kill you. <GABRIELLE> Boy, with all this passion isn't it a shame that we're not having sex?

10) Anger_Management_std_1_01 as hell. <CHARLIE> You know Ed, you make jokes like that <u>because you're insecure</u> so you denigrate others <u>to feel better about yourself</u>.

As can be seen in the examples above, *wh*-clauses condense contextual information through verbal arguments, thus promoting the economy of speech that is typical of telecinematic products. Conversely, adverbial clauses are structurally less integrated in the clause complex and can occur in both initial and final positions with respect to the main clause. The position of adverbial clauses has been associated with different pragmatic functions in spoken English (Diessel 2005) and film (Formentelli *et al*. forthcoming), with initial clauses serving as the thematic basis for the remainder of the utterance and final clauses mainly adding new information and facilitating processing.

One last point to be made about the narrative function of subordinate clauses in TV dialogue concerns relative clauses. Even though relative clauses occur less frequently compared to the other types of clauses in SydTV, they are still (extra)diegetically important to identify and further define people and places.

11) Human_Target_std_1_11 Down! Gerard, we need the evidence. <GERARD> Belfast. Remember **the man** <u>we hunted together</u>? <CHANCE> Really? <GERARD> He's here in New York.

12) Dollhouse_std_1_06 those movies never end well. <PAUL> It's **the part** <u>where I run down to Tiki Port and grab us dinner and we go over my files and you give me your perspective on some stuff</u>.

Relative clauses can be employed effectively as postmodifiers to pinpoint referents introduced by general nouns, such as *man* in Example 11, or to add details about hypothetical situations that cannot be represented visually onscreen, such as the series of actions that Paul describes in Example 12 as potential scenes in a film.

Overall, the different types of subordinate clauses identified in SydTV serve a variety of register-specific communicative purposes at the diegetic and extradiegetic levels, including the realistic simulation of spontaneous interactions, narration, characterisation, and the promotion of viewers' emotional involvement and entertainment. The multifaceted realisations of clausal complexity combine with several strategic audience-oriented functions in TV series to make the fictional stories more accessible, particularly for L2 learner-viewers, who can benefit from examples of rich and sophisticated English use.

5.3. Phrasal complexity

The aspect of phrasal complexity considered in this study is premodification, as mentioned in Section 4. More specifically, the study examines premodification involving up to three attributive adjectives, premodification involving up to three nouns and premodification involving genitives. The frequencies of these phrasal complexity features in SydTV are provided in Table 2.

Premodifying nouns		
Patterns	Raw frequency	Normed frequency (per 100,000 words)
Three premodifying nouns	33	11
Two premodifying nouns	294	106
One premodifying noun	2,480	895
Genitives	651	235
Total (premodifying nouns)	3,458	1,248
Premodifying adjectives		
Patterns	**Raw frequency**	**Normed frequency (per 100,000 words)**
Three premodifying adjectives	33	11
Two premodifying adjectives	520	187
One premodifying adjective	4,663	1,684
Total (premodifying adjectives)	5,216	1,883
Grand total	8,674	3,132

Table 2
Phrasal complexity – Premodification patterns in SydTV.

The frequency counts presented in Table 2 reflect two main tendencies in SydTV. The first is that adjectival premodification is more frequent than is nominal premodification (1,883 versus 1,248 occurrences per 100,000 words). The second is a preference for light forms of premodification, while heavy premodification via multiple premodifiers is much less common, particularly when it involves more than two modifiers[5]. In other words, when a noun is premodified in SydTV, this usually involves one adjective or – albeit less frequently – one noun.

TV dialogue aligns with spontaneous conversation in terms of its reliance on succinct premodification (Biber *et al.* 2021, pp. 591-592). By contrast, it differs from expository written registers, particularly newspaper prose and academic writing, in which heavier forms of premodification are attested (Biber *et al.* 2021, pp. 591-592; Zago 2024). In both casual conversation and fictional TV dialogue, the moderate use of premodification is a recipient-friendly tendency that is in keeping with the synchronous and auditory mode of reception of these spoken registers.

5.4. The functions of phrasal complexity features in TV dialogue

Complementing the discussion of clausal complexity in 5.2, this Subsection approaches the findings obtained for phrasal complexity from a functional perspective. The aim is to illustrate what phrasal complexity features typically do for L1 and L2 viewers or, in other words, to highlight the main register-specific communicative functions of phrasal complexity features in fictional TV dialogue.

A distinct trend in SydTV is that premodifiers serve to identify or qualify characters in various ways. There are cases in which characters use premodifiers to describe and evaluate themselves (Example 13), their interlocutor or a third party (Example 14). In other cases, premodifiers clarify the characters' identities, introduce another character or specify inter-character relationships (Example 15); all these functions assist viewers to understand the narrative. This is particularly the case for L2 viewers, for whom the explicitness and specification afforded by premodifiers may be especially beneficial and advantageous.

13) How_I_Met_Your_Mother_std_1_12 am. I'm not a **commitment guy**, I'm a **single guy**. <TED> Stuart, you don't have to be one or the

14) Breaking_Bad_std_1_03 news flash. That partner of yours? He's got a **big mouth**. Walter, I don't know what you think you're

---

[5] Multiple premodification may involve combinations of the premodifiers shown in Table 2. The most common type is when attributive adjectives combine with noun premodifiers (*bad car accident*). A search for the adjective + noun + noun cluster returned 352 matches in SydTV.

15) Parks_and_Recreation_std_1_06 wife. <WENDY> Hi, I'm Lindy Haverford. <MARK> Hi you're **Tom's wife**. <WENDY> Don't hold it against me. <TOM> Look at how

When utilised as in Examples 13 to 15, premodifiers contribute to the anchorage between the characters and the viewers, as well as to character revelation and characterisation, which are essential communicative functions in TV dialogue (Bednarek 2018a; Kozloff 2000). In such cases, TV language constructs the characters for the audience at different levels.

In addition to providing the audience with information about the characters, premodifiers are used to identify or describe the general context (Example 16) or specific parts (Example 17). This facilitates another type of anchorage, namely that between the audience and the physical spaces represented on the screen (Bednarek 2018a; Kozloff 2000).

16) The_Shield_std_1_04 's with me. <DANNY> Back off, okay. This is a **private party**. <LEMONHEAD> Look back off of them. You can't be

17) Weeds_std_1_01 <SILAS> Here. <QUINN> Oh, look, Silas, look at that **cute stuffed bear** on the table. We used to have one just like it[6]

Another significant function of premodifiers in TV dialogue is that of conveying the characters' viewpoints, attitudes, emotions and feelings (Examples 18 and 19), thus acting as markers of stance and expressivity (Bednarek 2018a, p. 140-144; Quaglio 2009, p. 87-105). When they have this function, they tend to be accompanied by a variety of other expressive items, such as interjections and intensifiers (*God* and *really* in Example 18), in such a way that reproduces and even amplifies the dynamics of spontaneous colloquial conversation, in which emphasis and emotionality abound when stance is presented.

18) Girls_std_1_03 because you both already have HPV. <HANNAH> God that's a really **good point**. <JULIAN> Marnie, I think one of these paintings is up crooked

19) Desperate_Housewives_std_1_19 <REX> I don't know. I think we may be making a **huge mistake**. <BREE> We made our decision, let's just stick to it

Apart from fulfilling the broad purposes described thus far, premodifiers have genre-related uses, as is evident in TV series that are set in specialised contexts. For instance, the audiences of medical TV series often listen to discussions concerning technical matters that necessitate complex noun phrases in order to be explained (Examples 20 and 21). Both the head nouns and the premodifiers that occur in such exchanges are notably more formal than those used to perform the general communicative functions that were illustrated previously.

---

[6] While not examined quantitatively, phrasal complexity via postmodification (the prepositional phrase 'on the table' in Example 17) was taken into consideration in the qualitative stage of this study.

20) House_std_1_18 <FOREMAN> The three miscarriages make me think
it's an **underlying physiology**. <HOUSE> **Pregnancy-related
autoimmunity**. Too bad that Cameron quit, I could use an immunologist

21) Greys_Anatomy_std_1_09 to know the cause of death? <BAILEY> It's
going down as **cardiopulmonary arrest** complicated by **liver disease**.
<IZZIE> But an autopsy would...

Similar occurrences can be seen in legal TV series, in which formal and
technical premodifiers are utilised in complex noun phrases by lawyers, judges
and police officers when discussing specialised, genre-defining topics, as
shown in Examples 22 and 23.

22) Southland_std_1_02 <SALINGER> Just stay there. <RUSSELL> This
guy's got a uh, **ten-page rap sheet**. **Spousal abuse**, **drug possession**, two
eleven.

23) Suits_std_1_10 know who you are. You said that if I accepted the
severance agreement there wouldn't be a lawsuit . <JESSICA> Well,
there

The pronounced phrasal complexity observed in the medical and legal TV
series of SydTV mirrors what takes place in real medical and legal English,
two registers that are known for their extensive use of heavily modified noun
phrases to ensure referential precision (Gotti 2011). In other words, phrasal
complexity is a requirement that scriptwriters have to satisfy for the sake of
realism in these genres. While realism often entails scriptwriters simulating the
lexicogrammar of colloquial spoken English, the type addressed here involves
presenting the audience with a credible reproduction of specialised discourses,
in line with Bednarek's (2018a, p. 71) observation that "realism/authenticity
can […] be produced by successful representations of particular language
varieties (associated with professional or ethnic groups)".

When a TV series portrays specialised contexts, the accessibility of its
dialogue may decrease for many lay L1 and L2 viewers, particularly the latter,
as a result of the phrasal intricacy as well as the degree of technicality of the
language in such contexts. However, this replicates or approximates what the
same viewers are likely to experience in their daily lives when they encounter
medical or legal English, with the result being that the dialogue gains
credibility. Ultimately, in scenes such as those in Examples 20 to 23, the aim
of constructing believable specialised discourse is of such importance for
scriptwriters that it may even be pursued at the expense of a certain degree of
comprehensibility. L1 and L2 viewers may not know precisely what Dr House
means when he says *pregnancy-related autoimmunity* (Example 20);
nevertheless, they have the impression of listening to true-to-life medical
discourse. The other side of the coin is that the phrasal complexity and
technicality of medical and legal TV genres only affect those viewers who do
not belong to the legal and medical fields. For learner-users of English who are
studying law or medicine, watching TV series such as *Grey's Anatomy* or *Suits*

is an opportunity – 'disguised' as entertainment – to improve their specialised L2 competence extramurally, in addition to being a practice that has been shown to be well suited to teaching English for specific purposes (Bonsignori 2018; Dang 2020; Forchini 2018; Franceschi 2021).

5.5. Comparison with spontaneous conversation

As a final step in the analysis, the quantitative findings regarding clausal and phrasal complexity in SydTV are compared here to corpus-based data pertaining to grammatical complexity in spontaneous speech. The results of Biber's (2015) research on the interplay between grammatical complexity and the expression of stance in face-to-face conversation, based on the spoken component of the *Longman Spoken and Written English Corpus*, are particularly relevant with regard to clausal complexity. Comparative data with TV dialogue are available for the three major categories of subordinate clauses considered in the present study, namely complement, adverbial and relative clauses, with the exception of nonfinite adverbial clauses. The partial lack of data is not expected to impact on the general trend displayed by the distribution of complexity features in the two registers, as nonfinite adverbial clauses are infrequent in conversation (Biber *et al.* 2021, p. 760). The frequencies of clausal complexity features in TV dialogue and spontaneous conversation normalised per 100,000 words are plotted in Figure 3.
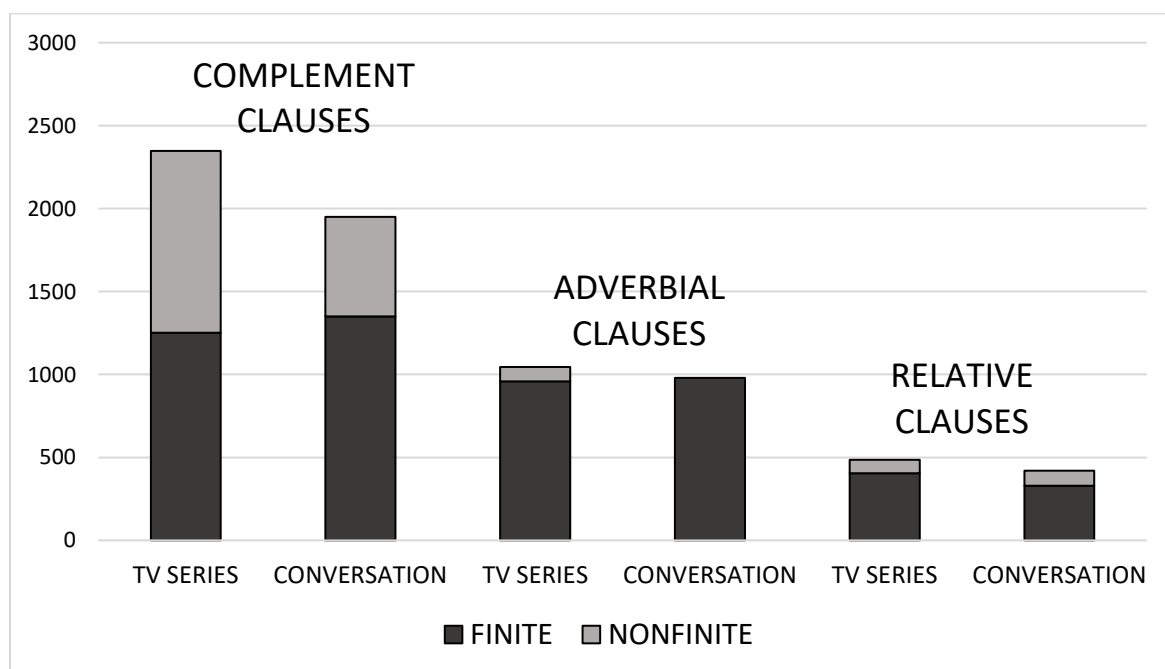

Figure 3
Frequencies of clausal complexity features in TV dialogue and spontaneous conversation
(per 100,000 words).

Complexity matters in TV dialogue as language input. A corpus-based description
of clausal and phrasal patterns

285

The distribution of clauses in the two registers reveals that TV dialogue closely approximates spontaneous conversation. Both varieties of spoken English favour the production of complement clauses over adverbial clauses and relative clauses in interactions. The clear preference for complement clauses in both TV dialogue and face-to-face conversation indicates the relevance of stance expression in spoken registers that have an interactive and interpersonal dimension, as these constructions are specialised in the codification of the speakers' attitudes, emotions and evaluations (Biber 2015, p. 12). The frequencies of each structural/syntactic subtype are also very similar in the two corpora, with the exception of nonfinite complement clauses, which are markedly more numerous onscreen. This difference might be explained by recalling a distinctive syntactic property of nonfinite clauses, namely the absence of a subject and tense marking, which allows for the expression of propositional content in a more compressed and economical way compared to finite subordinate clauses (Biber, Gray 2016), in line with the language constraints and requirements of the audiovisual medium. The pervasiveness of nonfinite complement clauses onscreen may also result in an enhanced dynamicity of exchanges amongst characters and a foregrounding of emotionality through dialogue.

Adverbial clauses occur to the same extent in TV dialogue and spontaneous conversation and are the second most frequent category of clausal complexity features. In face-to-face interactions and the onscreen representations, adverbial clauses enable the contextualisation of events and the development of common knowledge between real and fictional interactants.

The least frequent clausal subtype in the two corpora is relative clause. This finding may reflect the fact that most relative clauses serve as postmodifiers of head nouns and qualify as noun phrase constituents rather than as clausal constituents. While finite relative clauses have been found to be "relatively common in both written and spoken registers" (Biber *et al.* 2020, p. 8), it has also been found that noun postmodification is a more functional strategy in written registers than in spoken ones, as it allows for the compression of information into a few words (Biber *et al.* 2022, pp. 462-465).

Finally, TV dialogue and conversation share a preference for the explicit codification of new information through finite clause constructions, which may ease the decoding of the message for face-to-face interlocutors and comprehension for the viewing audience. Concurrently, less explicit nonfinite clausal combinations are used more frequently onscreen, where they arguably serve as time-saving strategies. In sum, the frequency and distribution of all types of clausal complexity features are comparable in the two registers with few exceptions, which is in line with recent findings concerning the complexity of Anglophone film dialogue (Formentelli *et al.* forthcoming). These results

provide additional evidence for the orientation of fictional TV dialogue towards real-life conversation, as pointed out in previous studies (Bednarek 2018a; Berber Sardinha, Veirano Pinto 2019; Quaglio 2009).

With regard to phrasal complexity in TV dialogue versus conversation, the frequencies of premodifiers in SydTV are compared here to those reported for spontaneous conversation in the *Grammar of Spoken and Written English* (GSWE) (Biber *et al.* 2021). The results of the comparison are plotted in Figure 4.

As can be seen in Figure 4, premodifiers – especially nouns – are more frequent in SydTV than in the conversation subcorpus of the GSWE. However, the difference is not marked, particularly when adjectives are considered, and therefore does not appear to indicate a substantial dissimilarity in the register profile of conversation and TV dialogue as far as premodification is concerned.
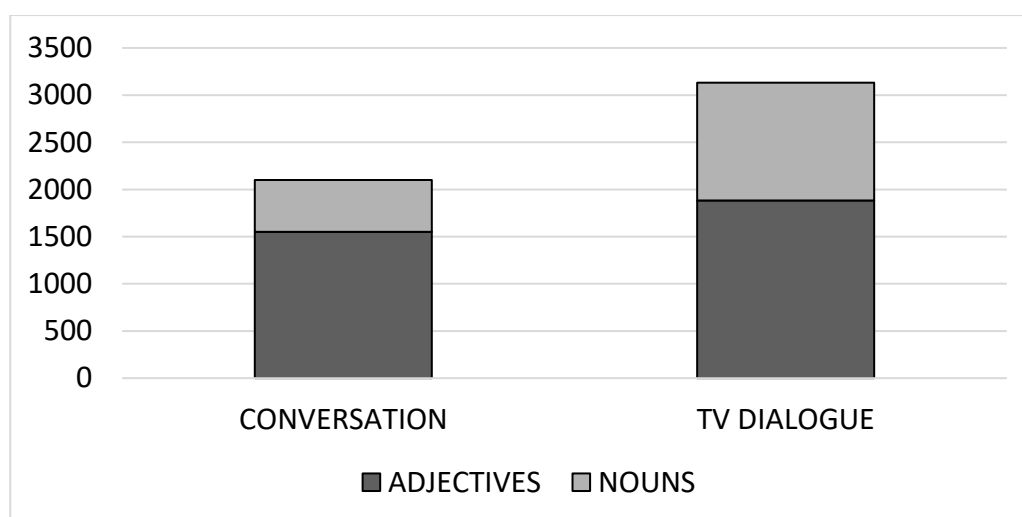
Figure 4
Phrasal complexity: premodifiers in conversation (adapted from Biber *et al.* 2021, p. 583, figure 8.7) vs. TV dialogue (frequencies normalised per 100,000 words).

As informational devices, premodifiers are ideally suited to the overarching goal of TV dialogue, namely that of narrating – a goal that is comparatively less important in conversation. This may be the main reason why premodifiers are more frequent in TV dialogue than in conversation. Given the association between phrasal modification and preplanning found in informational written registers (Biber *et al.* 2022), the slightly higher phrasal complexity observed in TV series may likewise result from the scripted nature of televisual language. In spontaneous conversation, by contrast, the left periphery of the noun phrase represents a more cognitively demanding, hence less frequently used, slot for speakers due to the online planning pressure they experience.

The moderately higher frequency of premodifiers in TV series arguably also stems from genre-related factors. For example, crime/legal and medical TV series regularly feature autopsies, police reports, investigations, legal

actions, diagnoses and surgical operations; these are activities that inevitably involve complex noun phrases and noun premodifiers by analogy with the real-world specialised varieties of English that these genres seek to imitate (Gotti 2011, p. 55-58). Another genre that deserves to be mentioned here is fantasy, as it features characters, settings and events whose extraordinary distance from reality 'calls for' description and illustration. Finally, in sitcoms, phrasal complexity also appears to be involved in the construction of humour. For instance, one of the reasons why viewers laugh when they watch *The Big Bang Theory* is that the awkward, nerdy characters in this series use a variety of English that is constantly complex and overly formal (cf. Bednarek 2023).

In summary, the comparison between the conversation subcorpus of the GSWE and SydTV suggests that TV dialogue is generally similar to casual conversation in terms of phrasal complexity. The former register relies on the referential function somewhat more extensively than does the latter, but the extent of the difference does not appear such that it 'compromises' TV dialogue as a model of spoken English for learner-users. Instead, the moderately higher presence of premodifiers in TV dialogue results in greater explicitness (in delineating a character, in contextualising, etc.), a feature that may help L1 and L2 viewers to orient themselves as they follow the storyline.

6. Concluding remarks

The quantitative findings of this study highlight some of the main types of grammatical complexity that occur in TV dialogue. With regard to the clausal dimension, complement and adverbial clauses are the preferred means of expressing complexity as opposed to relative clauses, with finite clauses being far more frequent compared to nonfinite ones. This distribution promotes narration and realism, and explicitly conveys background information about characters and events for the benefit of the audience. With regard to the phrasal dimension, attributive adjectives are more frequent than nouns as premodifiers, and single adjective/noun premodification is preferred to the use of multiple premodifiers.

The (extra)diegetic functions served by complexity features in televisual language were identified and documented through the qualitative analysis of concordances in SydTV. Many of these functions were observed across different TV series and can therefore be regarded as being typical of fictional TV dialogue in general. They include the identification and description of referents, the presentation of circumstantial information, the establishment and clarification of the relationships amongst characters and events, the marking of the characters' stance and expressivity. Other functions, such as the simulation and reproduction of the phrasal complexity of specialised discourses, are more common in TV series that depict legal activities, investigations of criminal

cases, and medical professions. This aspect emphasises the role of TV genres in generating variation within the register as a whole.

From a comparative perspective, TV dialogue was found to approximate spontaneous conversation with regard to major patterns of grammatical complexity, as is particularly evident in the occurrences of most of the types of subordinate clauses and in the rates of adjectival premodification. These findings reflect the mimetic aim of onscreen language with respect to naturally occurring spoken exchanges. At the same time, some quantitative patterns that partly distance fictional TV dialogue from face-to-face conversation were detected, such as the greater frequency of nonfinite complement clauses and of nominal premodifiers in TV series. The strong reliance on nonfinite complement clauses may reflect the medium-related need for economy of speech, whereas the difference in nominal premodification may be explained as an attempt to reproduce specialised discourse onscreen through domain-specific vocabulary in more elaborate noun phrases. Narrative exigencies, together with the scripted nature of TV dialogue, may also be proposed to account for the slightly higher phrasal complexity found in SydTV, as phrasal complexity is closely associated with the preplanned nature of informational written texts (Biber *et al.* 2022). More evidence about this aspect may be obtained by assessing the degree of phrasal postmodification through prepositional phrases as part of a future development of the present investigation.

When considered from an informal learning perspective, the results of this investigation allow us to describe TV dialogue as a rich and reliable source of input for L2 learner-users. TV dialogue qualifies as a source of realistic conversational language that learner-viewers have at their disposal and can readily use as a model of spoken English. Phrasal and clausal complexity features perform several referential and communicative functions onscreen, and the specification and contextualisation they afford can increase the accessibility of TV narratives for L2 viewers, often in tandem with visuals in a multimodal fashion. In conclusion, it should be noted that, while TV dialogue is a unidirectional, screen-to-face register, the wide availability of audiovisual products on numerous streaming platforms considerably increases the opportunities for many learner-users to access English, and probably provides more opportunities compared to bidirectional, face-to-face interactions with native speakers. This makes TV series a valuable tool for learning English outside of the classroom.

Bionotes

Maicol Formentelli, PhD, is Associate Professor of English language and Linguistics at the University of Pavia. His main research interests include sociolinguistic, pragmatic, and interpersonal aspects of language with a focus on native and non-native varieties of English in naturally occurring interactions as well as in fictional dialogue (original and dubbed films). He has carried out corpus-based studies on address strategies, interpersonal stance and swearing in English and English/Italian translation. His most recent publications include "The pragmatic functions of *tu* and *lei* in films: Converging patterns of address across translated and original Italian dialogue" (Journal of Pragmatics, 2022) and "The pragmatic dimensions of swearing in films: Searching for coherence in dubbing strategies (Journal of Pragmatics, 2023).

Raffaele Zago, PhD in English Linguistics, is Assistant Professor of English at the University of Catania, Department of Humanities. His research interests include spoken English, register variation, corpus linguistics, discourse analysis, telecinematic dialogue and informal contact with English. In these areas, he has published two monographs (*From Originals to Remakes. Colloquiality in English Film Dialogue over Time*, 2016; *Cross-linguistic Affinities in Film Dialogue*, 2018), various articles/chapters and has co-edited two volumes (*Language*, *Expressivity and Cognition*, 2023; *English Studies in Italy*: *New Directions and Perspectives*, in press). He is currently the local coordinator of the national research project *The Informalisation of English Language Learning through the Media*: *Language Input*, *Learning Outcomes and Sociolinguistic Attitudes from an Italian Perspective*.

Authors' addresses: maicol.formentelli@unipv.it; raffaele.zago@unict.it.

References

Arends J. 2001, *Simple Grammars, Complex Languages*, in "Linguistic Typology" 5 [2/3], pp. 180-182.

Arnbjörnsdóttir B. and Ingvarsdóttir H. (Eds.) 2018, *Language Development across the Life Span. The Impact of English on Education and Work in Iceland*, Springer International Publishing, New York.

Bednarek M. 2010, T*he Language of Fictional Television: Drama and Identity*, Continuum, London.

Bednarek M. 2011, *The Language of Fictional Television: A Case Study of the 'Dramedy' Gilmore Girls*, in "English Text Construction" 4 [1], pp. 54-83.

Bednarek M. 2012, *"Get Us the Hell out of Here": Key Words and Trigrams in Fictional Television Series*, in "International Journal of Corpus Linguistics" 17 [1], pp. 35-63.

Bednarek M. 2018a, *Language and Television Series. A Linguistic Approach to TV Dialogue*, Cambridge University Press, Cambridge.

Bednarek M. 2018b, *Guide to the Sydney Corpus of Television Dialogue (SydTV)*, www.syd-tv.com (1.2.2021).

Bednarek M. 2023, *Language and Characterisation in Television Series. A Corpus-informed Approach to the Construction of Social Identity in the Media*, John Benjamins, Amsterdam-Philadelphia.

Bednarek M. and Zago R. 2024, *Bibliography of Linguistic Research on Fictional (Narrative, Scripted) Television Series and Films/Movies*, version 7, https://unico.academia.edu/RaffaeleZago (30.9.2024).

Berber Sardinha T. and Veirano Pinto M. 2019, *Dimensions of Variation across American Television Registers*, in "International Journal of Corpus Linguistics" 24 [1], pp. 3-32.

Biber D. 1988, *Variation across Speech and Writing*, Cambridge University Press, Cambridge.

Biber D. 1992, *On the Complexity of Discourse Complexity: A Multidimensional Analysis*, in "Discourse Processes" 15, pp. 133-163.

Biber D. 2015, *Stance and Grammatical Complexity: An Unlikely Partnership Discovered through Corpus Analysis*, in "Corpus Linguistics Research" 1, pp. 1-19.

Biber D. and Gray B. 2016, *Grammatical Complexity in Academic English: Linguistic Change in Writing*, Cambridge University Press, Cambridge.

Biber D., Gray B., Staples S. and Egbert J. 2020, *Investigating Grammatical Complexity in L2 English Writing Research: Linguistic Description versus Predictive Measurement*, in "Journal of English for Academic Purposes" 46, Article 100869, n.p..

Biber D., Johansson S., Leech G., Conrad S. and Finegan E. 2021, *Grammar of Spoken and Written English*, John Benjamins, Amsterdam/Philadelphia.

Biber D., Gray B., Staples S. and Egbert J. 2022, *The Register-Functional Approach to Grammatical Complexity. Theoretical Foundation, Descriptive Research Findings, Application*, Routledge, New York/London.

Biber D., Larsson T. and Hancock G.R. 2023, *The Linguistic Organization of Grammatical Text Complexity: Comparing the Empirical Adequacy of Theory-Based Models*, in "Corpus Linguistics and Linguistic Theory" 20 [2], pp. 347-373.

Bonsignori V. 2018, *Using Films and TV Series for ESP Teaching: A Multimodal Perspective*, in Crawford Camiciottoli B. and Campoy Cubillo M.C. (Eds.), *Multimodal Perspectives on English Language Teaching in Higher Education*, special issue of "System" 77, pp. 58-69.

Complexity matters in TV dialogue as language input. A corpus-based description
of clausal and phrasal patterns

291

Bulté B. and Housen A. 2012, *Defining and Operationalising L2 Complexity*, in A. Housen, F. Kuiken and I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency. Complexity*, *Accuracy and Fluency in SLA*, John Benjamins, Amsterdam, pp. 21-46.

Bulté B. and Housen A. 2014, *Conceptualizing and Measuring Short-Term Changes in L2 Writing Complexity*, in "Journal of Second Language Writing" 26, pp. 42-65.

Dang T.N.Y. 2020, *The Potential for Learning Specialized Vocabulary of University Lectures and Seminars Through Watching Discipline-Related TV Programs: Insights From Medical Corpora*, in "Tesol Quarterly" 54 [2], pp. 436-459.

Diessel H. 2005, *Competing Motivations for the Ordering of Main and Adverbial Clauses*, in "Linguistics" 43 [3], pp. 449-470.

Dynel M. 2011, *'You Talking to Me?' The Viewer as a Ratified Listener to Film Discourse*, in "Journal of Pragmatics" 43, pp. 1628-1644.

Forchini P. 2012, *Movie Language Revisited. Evidence from Multi-Dimensional Analysis and Corpora*, Peter Lang, Bern.

Forchini P. 2018, *The Applicability of Movies in Legal Language Teaching*: *Evidence from Multi-Dimensional Analysis*, in "International Journal of Linguistics" 10 [6], pp. 245-262.

Formentelli M. 2014, *Exploring Lexical Variety and Simplification in Original and Dubbed Film Dialogue*, in M. Pavesi, M. Formentelli and E. Ghia (Eds.), *The Languages of Dubbing. Mainstream Audiovisual Translation in Italy*, Peter Lang, Bern, pp. 141-166.

Formentelli M., Galiano L. and Pavesi M. forthcoming, *Grammatical Complexity in Film Dialogue: A Corpus-Based Study from a Register-Functional Perspective*, in "International Journal of Corpus Linguistics".

Franceschi D. 2021, *Fostering Multimodal Literacy in ESP Teaching: the Case of Lawyer-Client Interviews*, in "ESP Across Cultures" 18, pp. 81-103.

Gotti M. 2011, *Investigating Specialised Discourse*, Peter Lang, Bern.

Halliday M.A.K. 1989, *Spoken and Written Language*, Oxford University Press, Oxford.

Jones C. 2017, *Soap Operas as Models of Authentic Conversations: Implications for Materials Design*, in Tomlinson B. and Maley A. (Eds.), *Authenticity in Materials Development for Language Learning*, Cambridge Scholars Publishing, Newcastle upon Tyne, pp.158-175.

Kozloff S. 2000, *Overhearing Film Dialogue*, University of California Press, Berkeley/Los Angeles-London.

Kusyk M. 2020, *Informal English Learning in France*, in Dressman M. and Sadler R. (Eds.), *Handbook of Informal Language Learning*, Wiley Blackwell, Hoboken, pp. 333-348.

Miestamo M., Sinnemäki K. and Karlsson F. (Eds.) 2008, *Language Complexity: Typology, Contact*, *Change*, John Benjamins, Amsterdam.

Nichols J. 2009, *Linguistic Complexity: A Comprehensive Definition and Survey*, in Sampson G., Gil D. and Trudgill P. (Eds.), *Language Complexity as an Evolving Variable*, Oxford University Press, Oxford, pp. 64-79.

Norris J.M. and Ortega L. 2009, *Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity*, in "Applied Linguistics" 30, pp. 555-578.

Pallotti G. 2015, *A Simple View of Linguistic Complexity*, in "Second Language Research" 31 [1], pp. 117-134.

Pavesi M. 2012, *The Potentials of Audiovisual Dialogue for Second Language Acquisition*, in Alderete-Díez P., Incalcaterra McLoughlin L., Ní Dhonnchadha L. and Ní Uigín D. (Eds.), *Translation, Technology and Autonomy in Language Teaching and Learning*, Peter Lang, Bern, pp. 155-174.

Pavesi M. 2015, *From the Screen to the Viewer-Learner. Audiovisual Input as a Context for Second Language Acquisition*, in S. Campagna, E. Ochse, V. Pulcini and M. Solly

(Eds.), *Languaging in and across Communities: New Voices, New Identities. Studies in Honour of Giuseppina Cortese*, Peter Lang, Bern, pp. 83-104.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case-Study of Italian Postgraduate Students*, Edizioni ETS, Pisa.

Quaglio P. 2009, *Television Dialogue*: *The Sitcom Friends vs. Natural Conversation*, John Benjamins, Amsterdam/Philadelphia.

Scheffler P., Jones C. and Dominska A. 2020, *The Peppa Pig Television Series as Input in Pre-primary EFL Instruction: A Corpus-based Study*, in "International Journal of Applied Linguistics" 31, pp. 3-17.

Sockett G. 2014, *The Online Informal Learning of English*, Palgrave MacMillan, London.

Szmrecsanyi B. 2015, *Recontextualizing Language Complexity*, in Daems J., Zenner E., Heylen K., Speelman D. and Cuyckens H. (Eds.), *Change of Paradigms – New Paradoxes: Recontextualizing Language and Linguistics,* Applications of Cognitive Linguistics 31, De Gruyter Mouton, Berlin, pp. 347-360.

Veirano Pinto M. 2014, *Dimensions of Variation in North American Movies*, in T. Berber Sardinha and M. Veirano Pinto (Eds.), *Multi-dimensional Analysis*, *25 Years on: A Tribute to Douglas Biber*, John Benjamins, Amsterdam, pp. 109-146.

Webb S. and Rodgers P.H. 2009a, *Vocabulary Demands of Television Programmes*, in "Language Learning" 59 [2], pp. 335-366.

Webb S. and Rodgers P.H. 2009b, *The Lexical Coverage of Movies*, in "Applied Linguistics" 30 [3], pp. 407-427.

Werner V. 2018, *The Language of Pop Culture*, Routledge, New York.

Zago R. 2019, *Complexity in Film Dialogue*, in "Le forme e la storia" 12 [1], pp. 123-32.

Zago R. 2024, *Complexity Markers in TV Series: A Corpus-based Analysis of Premodification*, in D'Amore M., Zago R., Bugliani P., Polatti A. and Vignozzi G. (Eds.), *English Studies in Italy: New Directions and Perspectives*, Carocci, Roma, pp. 137-150.

TV series

Abrams J. J., Kurtzman A., Orci R. (Creators) (2008), *The transformation* (Season 1 Episode 13) [TV series episode], in Abrams J.J., Burk B., Kurtzman A., Orci R., Wyman J.H., J. Pinkner, Chappelle J. (Executive Producers), *Fringe*, Bad Robot, Warner Bros. Television.

Bays C., Thomas C. (Creators), Fryman P., Greenberg R., Shea M., Harris N.P. (Directors) (2005), *The wedding* (Season 1 Episode 12) [TV series episode], in Bays C., Thomas C., Greenberg R., Fryman P., Heisler E., Heline D., Lord P., Malins G., Miller C., Harris C., Lloyd S., Kang K., Rhonheimer J., Tatham C. (Executive Producers), *How I Met your Mother*, Bays Thomas Productions, 20th Century Fox Television.

Biderman A. (Creator) (2009), *Mozambique* (Season 1 Episode 2) [TV series episode], in Biderman A., Chulack C., Wells J. (Executive Producers), *Southland*, John Wells Productions, Warner Bros. Television.

Cherry M. (Creator) (2004), *Live alone and like it* (Season 1 Episode 19) [TV series episode], in Cherry M., Spezialy T., Edelstein M. (Executive Producers), *Desperate Housewives*, Cherry Productions, Touchstone Television.

Cody D. (Creator) (2009), *Abundance* (Season 1 Episode 8) [TV series episode], in Spielberg S., Cody D., Junge A., Frank D., Falvey J. (Executive Producers), *United States of Tara*, DreamWorks Television, Showtime Networks.

Daniels G., Schur M. (Creators) (2009), *Rock show* (Season 1 Episode 6) [TV series episode], in Daniels G., Schur M., Klein H., Miner D., Sackett M., Holland D., Goor D. (Executive

Complexity matters in TV dialogue as language input. A corpus-based description
of clausal and phrasal patterns

293

Producers), *Parks and Recreation*, Open 4 Business Productions, Deedle-Dee Productions, Fremulon, 3 Arts Entertainment, Universal Television.

Dunham L. (Creator) (2012), *All adventurous women do* (Season 1 Episode 3) [TV series episode], in Dunham L., Apatow J., Konner J., Landress I.S., Kaplan B.E. (Executive Producers), *Girls*, Apatow Productions, I Am Jenni Konner Productions, HBO Entertainment.

Gilligan V. (Creator) (2008), *And the bag's in the river* (Season 1 Episode 3) [TV series episode], in Gilligan V., Johnson M., MacLaren M. (Executive Producers), *Breaking Bad*, High Bridge Entertainment, Gran Via Productions, Sony Pictures Television.

Helford B. (Creator) (2012), *Charlie goes back to therapy* (Season 1, Episode 1) [TV series episode], in Helford B., Estevez R., Sheen C., Caplan D., Burg M., Totino V., Roth J., Kushell B., Maron R. (Executive Producers), *Anger Management*, Mohawk Productions, Revolution Studios, Estevez/Sheen Productions, Twisted Television, Debmar-Mercury, Lionsgate Television.

Kohan J. (Creator) (2005), *You can't miss the bear* (Season 1 Episode 1) [TV series episode], in Kohan J. (Executive Producers), *Weeds*, Tilted Productions, Lionsgate Television, Showtime Networks.

Korsh A. (Creator) (2011), *The shelf life* (Season 1 Episode 11) [TV series episode], in Korsh A., Liman D., Bartis D. (Executive Producers), *Suits*, Untitled Korsh Company, Hypnotic Films & Television, Universal Cable Productions, Open 4 Business Productions.

Lieber J., Abrams J.J., Lindelof D. (Creators) (2004), *... in translation* (Season 1 Episode 17) [TV series episode], in Abrams J.J., Lindelof D., Burk B., Cuse C., Bender J., Pinkner J., Williams S., Kitsis E., Horowitz A., Higgins J., Sarnoff E. (Executive Producers), *Lost*, Grass Skirt Productions, Bad Robot, Touchstone Television.

Rhimes S. (Creator) (2005), *Who's zoomin' who?* (Season 1 Episode 9) [TV series episode], in Rhimes S., Pompeo E., Heinberg A., Reaser A., Beers B., Allen D., Parriott J.D., Renshaw J., Rafner J., Rater J., Hodder K., Vernoff K., Gordon M., Wilding M., Noxon M., Marinis M., Horton P., Corn R., McKee S., Mulholland S., Phelan T., Harper W., Clack Z., Dumont O. (Executive Producers), *Gray's Anatomy*. Shondaland, Lionsgate Television, 20th Television.

Ryan S. (Creator) (2002), *Dawg days* (Season 1 Episode 4) [TV series episode], in Ryan S., Brazil S., Mazzara G., Eglee C.H., Sutter K., Rosenbaum S., Fierro A. (Executive Producers), *The Shield*, The Barn Productions, MiddKid Productions, Columbia TriStar Domestic Television, Fox Television Studios.

Schwartz J., Savage S. (Creators) (2007), *Woman on the verge* (Season 1 Episode 17) [TV series episode], in Schwartz J., Savage S., Levy B., Morgenstein L., Stephens J. (Executive Producers), *Gossip Girl*, Warner Bros. Television, Alloy Entertainment, College Hill Pictures, Inc.

Shore D. (Creator) (2004), *Babies and birthwater* (Season 1 Episode 18) [TV series episode], in Attanasio P., Jacobs K., Shore D., Singer B., Moran T.L., Friend R., Lerner G., Yaitanes G., Laurie H. (Executive Producers), *House*, Heel and Toe Films, Shore Z Productions, Bad Hat Harry Productions, NBC Universal Television Studio.

Steinberg J.E. (Creator) (2010), *Victoria* (Season 1 Episode 11) [TV series episode], in Steinberg J.E., Kern B., Hooks K., McG, Johnson P., Miller M. (Executive Producers), *Human Target*, Wonderland Sound and Vision, DC Entertainment, Bonanza Productions, Warner Bros. Television.

Whedon J. (Creator) (2009), *Man on the street* (Season 1 Episode 6) [TV series episode], in Whedon J., Solomon D., Minear T. (Executive Producers), *Dollhouse*, Mutant Enemy Productions, 20th Century Fox Television.

NOUN PHRASE COMPLEXITY IN FILM DIALOGUE AS INPUT FOR SECOND-LANGUAGE LEARNING
A corpus-based study
from a register-functional perspective

LIVIANA GALIANO
UNIVERSITÀ SAN RAFFAELE ROMA

Abstract – Since audiovisual dialogue represents a common means of informal contact with English (Pavesi, Ghia 2020), it is important to understand its characteristics. Following a study on clausal complexity in film dialogue (Formentelli *et al.* forthcoming), the present work investigates phrasal complexity, thus contributing to the provision of a broader overview of grammatical complexity in film dialogue. The results provide insights into which phrasal structures second-language (L2) learners accessing English through films are likely to encounter notice and internalise after repeated exposure (Kerswill, Williams 2002). The study investigates nominal pre- and post-modification in a corpus of 34 anglophone film dialogues by examining adjectival and nominal pre-modifiers and prepositional phrases as post-modifiers. Further analysis of the types of the most frequently modified nominal heads is also conducted. The results of the analysis are interpreted by adopting a register-functional approach to complexity (Biber *et al.* 2022); they suggest a degree of phrasal complexity in film dialogue that approximates trends in natural conversation while also performing register-specific functions.

Keywords: film dialogue; phrasal complexity; register-functional; English.

1. Introduction

Register-functional studies of grammatical complexity (see Biber *et al.* 2022) have shown that phrasal complexity is pivotal in the production of specialised written academic texts. Despite the fact that phrasal complexity finds its greatest expression in written registers, a study thereof can provide useful insights when applied to texts simulating spoken interaction, such as are found in films. The interest in studying the phrasal complexity of film dialogue is two-fold: first, it helps in establishing whether the written nature of film dialogue is mirrored in its grammatical complexity; second, since films have been found to be a frequently accessed source of English among second-language (L2) learners outside of the classroom (Sundqvist 2009; Pavesi, Ghia 2020), knowing more about the phrasal complexity of film dialogue contributes to a description of this crucial type of input for L2 acquisition.

Following previous investigations into the clausal complexity of film dialogue (Formentelli *et al.* forthcoming), the present study aims to contribute to building a more comprehensive overview of the grammatical complexity of this register by examining its phrasal dimension. A corpus of anglophone film dialogues is analysed to find patterns of phrasal complexity (pre- and post-modification patterns). The trends highlighted for film dialogue are subsequently compared to the data on phrasal complexity in spontaneous conversation that Biber *et al.* (2021) provided.

The paper is structured as follows: Section 2 briefly introduces the notion of phrasal complexity and the register-functional approach; Section 3 discusses the main characteristics of film dialogue as a register and as input for informal L2 learning; Section 4 outlines the research questions and describes the corpus and methodology used in the present study; Section 5 presents the results of the analysis of phrasal complexity features in film dialogue; Section 6 compares the phrasal complexity of film dialogue and natural conversation; Section 7 presents the results and discusses the adoption of the register-functional approach.

2. Phrasal complexity and the register-functional approach

The study of phrasal complexity is rather a recent matter. As Staples *et al.* (2016) point out, only a few studies on the development of academic writing have included a number of features of phrasal complexity (Crossley *et al.* 2011; Haswell 2000; Lu 2011), and even fewer have investigated the relation of such features with genre (cf. Lu 2011; Beers, Nagy 2009). This tendency seems to be related to the fact that for a long time, the study of grammatical complexity has actually meant study of subordinate clauses, which have long been considered the highest expression of syntactic complexity in English (cf. Bulté, Housen 2012; Biber *et al.* 2022). In the latest years, however, in particular in the wake of register-functional studies of grammatical complexity (Biber *et al.* 2022), it has been evidenced how subordinate clauses only represent one type of grammatical complexity. In fact, different registers display a preference for either clausal or phrasal complexity.

Corpus-based inductive analyses (Biber *et al.* 2022) have shown that grammatical complexity is better regarded as a multidimensional construct that is expressed in structurally different ways along the speaking-to-writing continuum. Specifically, the key difference between spoken and written registers is the reliance of the former on clausal complexity and the latter on phrasal complexity. While clausal complexity indicates long and elaborate dependent clauses, phrasal complexity refers to dense and compact pre- and

post-modified noun phrases (Biber, Gray 2016; Biber *et al.* 2022). These register-specific tendencies have been linked primarily to the circumstances of production and secondarily to the communicative purposes of a text (Biber *et al.* 2022, p. 463; cf. also Biber 1992). Since spoken language is produced on-line, with few or no opportunities for planning or revising, it makes extensive use of dependent clauses, which are added incrementally following the speaker's stream of thought (cf. Ferreira, Slevc 2007). Moreover, among dependent clauses complement clauses are a preferred strategy to express stance (e.g. *I think that…*, *I love how…*), thus a suitable linguistic strategy for the involved character of spontaneous, informal conversation (see example 1) (see Biber *et al.* 2021). Contrastingly, since written language can exploit the availability of planning as well as revision time and in its most prototypical forms it is used for informational purposes, it exploits phrasal complexity in order to compact as much information as possible (see example 2).

> [1] Excerpt from a conversation (Biber, Gray 2016, p. 89)
>
> Oh, I have to do three different kinds of reviews in my senior seminar class and so, it's a good excuse <u>to go out for dinner</u>. <laughing.> I – you know – my teacher was suggesting <u>that we review a restaurant</u> so I was trying <u>to think of a good restaurant</u>.

> [2] Excerpt from an engineering textbook (Biber, Gray 2016, p. 71)
>
> <u>For applications with parallel shafts</u>, <u>straight spur, stepped, helical, double helical, or herringbone gears</u> are usually used. <u>In the case of intersecting shafts</u>, <u>straight bevel, spiral bevel, or face gears</u> are employed.

As a component of grammatical complexity, phrasal complexity is defined as the elaboration of phrases (nominal, adjectival, adverbial, prepositional), with a focus on the noun phrase and its pre- and post-modification (Biber *et al.* 2022). In her study on noun phrase complexity in English, Berlage (2014) points out that what is considered more or less complex in noun phrases remains an unresolved issue. She defines phrasal complexity as a combination of at least three factors: length, the number of embedded phrases and whether or not verb phrases are embedded in the noun phrase. She also claims that quantitative parameters alone cannot be determinants of the degree of noun phrase complexity, as the qualitative dimension also needs to be considered. According to Berlage (2014, p. 2), this dimension corresponds to noun phrases being or not being sentential, i.e. featuring verb phrases in their embeddings. Maintaining an integrated approach combining quantitative and qualitative dimensions of grammatical complexity, Biber *et al.* (2022) propose examining not only the types of syntactic patterns but also their function and distribution in a particular register. Therefore, the degree of phrasal complexity is dependent on how much sophistication can be found in the noun phrase in terms of how many and which types of elements (e.g.

Lingue e Linguaggi

adjectives, nouns, etc.) are added to its basic structure, as represented by a single noun, and where (pre-modification vs. post-modification).

Considering the findings of previous studies on register variation and complexity, the expectation for film dialogue is to find a low degree of phrasal elaboration, following the phrasal complexity tendencies found in natural conversation, which is the register film dialogue aims to simulate. In spoken English, noun phrases are not very complex since they feature few elements; they usually contain a head noun alone or with one pre-modifying element. Furthermore, post-modification is even less frequent than pre-modification; rarely are nouns both pre- and post-modified simultaneously (Biber *et al.* 2021, p. 572).

3. Film dialogue as L2 input in informal access to English

Audiovisual dialogue has been observed to be a preferred means to access English extramurally, outside of the educational context (Sundqvist 2009; Pavesi, Ghia 2020). In all models of L2 acquisition, both the quantity and quality of the input are considered essential to the learning process (cf., among others, Krashen 1985; Ellis, Wulff 2015; VanPatten 2015; Montag *et al.* 2018; Anderson *et al.* 2021). According to Caruana (2009), the incidental acquisition of an L2 is particularly supported in films and television because verbal, paraverbal and non-verbal elements coincide frequently therein. Additionally, L2 audiovisual input creating an immersive experience for the audience is considered an advantage for L2 learners, whereby they can become emotionally receptive if the L2 does not represent a barrier (Caruana 2009; Pavesi 2015; Pavesi, Ghia 2020). Furthermore, since the focus is on entertainment rather than language learning, incidental learning is more likely to occur (Reinders, Benson 2017), as evidenced in media psychology (d'Ydewalle, Pavakanun 1996, 1997; Green *et al.* 2004; Wissmath *et al.* 2009).

The widely evidenced similarity between film and television language to naturally occurring conversation (Quaglio 2009; Bednarek 2010, 2018; Forchini 2012, 2021; Zago 2015; Werner 2021) suggests that the former is ideal for the acquisition of informal English (Pavesi 2015). Moreover, audiovisual dialogue is mostly free from dysfluency phenomena, which may be obstacles to information processing among L2 learners. A further advantage to using audiovisual language as input for an L2 is found in its staging situations characterised by informal language use, which are difficult to reproduce in a standard learning environment. Obtaining access to informal language also means coming in contact with the interpersonal, involved and emotional uses of language due to the freer syntactic organisation of discourse (Finegan, Biber 2001; Warren 2006) and the variety

of pragmatic strategies used, including emphasis, information structure management and politeness.

However, audiovisual input is not free from the structural complexity that characterises English and spoken language, the latter bearing specificities related to on-line production and the constant management of the speaker-hearer relationship. Furthermore, film bears register-specific characteristics related to its narrative structure, time constraints and salient artistic–symbolic intent (Alvarez-Pereyre 2011). Concurrently, film has a sophisticated architecture due to the interplay between the diegetic and extradiegetic levels, the level of the story portrayed on screen and the level on which the dialogue is indirectly addressing the audience (Lorenzo-Dus 2009; Dynel 2011; Guillot, Pavesi 2019). This may all impact the linguistic make-up and grammatical complexity of film.

Studying the grammatical complexity of film dialogue is relevant to L2 acquisition in two main ways: quantitatively since frequently encountered linguistic expressions and patterns are more likely to be noticed and internalised (Bley-Vroman 2002; Bybee 2008); qualitatively since it is important to know which structure types are found in the input and what functions they perform in order to formulate expectations about learning outcomes.

4. Research questions, data and methodology

The present study aims to describe the main features of phrasal complexity in film dialogue by examining different types of pre- and post-modification strategies and their frequencies of occurrence. Drawing on the assumption that film dialogue is meant to imitate natural conversation, the expectation is to find a low degree of phrasal complexity, short phrases with little or no modification (Biber *et al.* 2023). However, since film dialogue is first produced in written form and is then performed by actors and interpreted by the audience as spoken language, it may bear traces of its production circumstances. Specifically, time for pre-planning and medium-related time limitations may result in more complex and compressed phrasal expressions. To qualify the phrasal complexity of film dialogue, the following research questions (RQs) were formulated:

1. What length and type of pre-modification characterise film dialogue?
2. What type of post-modification characterises film dialogue?
3. What nouns are most often pre- and post-modified in film dialogue?
4. Are the patterns of pre- and post-modification in film dialogue similar to those characterising natural conversation?

*L*ingue e
*L*inguaggi

The analysis relies on data from 34 orthographic transcriptions of British and American film dialogues, which constitute the anglophone component of the Pavia Corpus of Film Dialogue[1] (PCFD), for a total of 380,219 words. The films included in the PCFD present stories about everyday life that are likely to trigger the representation of spontaneous conversation onscreen (Pavesi 2014). The PCFD is entirely POS-tagged using the CLAWS7 tagset (Galiano, Semeraro 2023) to enable the semi-automatic retrieval of syntactic patterns.

To answer RQs 1 and 2, the patterns selected from Biber *et al.*'s (2022, pp. 14-15) taxonomy of grammatical complexity features in English (detailed in Table 1) were analysed using the AntConc concordancer[2] (version 4.2.0).

Structural type	Syntactic function within structural type	Specific structural/syntactic features	Examples
Dependent phrase	Noun phrase modifier	Attributive adjective as premodifier Noun as premodifier Prepositional phrase as postmodifier	You know I bought Chris an ancient Greek fertility charm? (*Match Point*, Allen, 2005) Well, I had to pick a place for my party, and I read that article about you. (*Autumn in New York*, Chen, 2000)
	Adverbial phrase	Prepositional phrase modifying a clause	Has your mother talked to you guys about this stuff? (*Boyhood*, Linklater, 2014)

Table 1
Phrasal complexity features in English
(adapted from Biber *et al.* 2022, p. 14-15).

The noun phrase was analysed in its pre-modification by examining both adjectives and nouns as well as post-modification by prepositional phrases (PPs). In addition, PPs used as adverbials modifying clauses (both verb arguments and additional information) were included in the analysis for functional comparison (N-attachment vs. V-attachment). The analysis of pre- and post-modifying patterns relied on the semi-automatic retrieval of patterns of POS-tags. That is, all possible combinations of pre- and post-modifiers were converted into combinations of tags, as reported in Table 2. The use of wildcards (e.g. *) allowed for more comprehensive searches (see also Figure 1).

---

[1] https://studiumanistici.unipv.it/?pagina=p&titolo=pcfd
[2] https://www.laurenceanthony.net/software/antconc/

Noun phrase complexity in film dialogue as input for second-language learning.
A corpus-based study from a register-functional perspective

301

Pattern	Description	Tag sequence
(Adj) + (Adj) + (Adj) + Adj + N	Any noun preceded by up to four adjectives	(*_JJ*) (*_JJ*) *_JJ* *_NN*
(N) + (N) + (N) + N + N	Any noun preceded by up to four nouns	(*_NN*) (*_NN*) *_NN* *_NN*
N's + (Adj) + N	Any noun (pre-modified or not) preceded by Saxon Genitive	*_GE* (*_JJ*) *_NN*
(Adj) + Adj + N + N	Any noun-noun combination preceded by up to two adjectives	(*_JJ*) *_JJ* *_NN* *_NN*
Adj + N + N + N	Any noun-noun combination preceded by adjective+noun pre-modifying combination	*_JJ* *_NN* *_NN* *_NN*
Adv + Adj + N	Any pre-modifying sequence containing an adverb modifying the adjective	*_RR* *_JJ* *_NN*
N + P	Any noun followed by any preposition	*_NN* *_I*

Table 2
Pre- and post-modification queries.

The concordances retrieved through the search for POS-tag combinations in AntConc were manually checked to prevent false positives, and in some cases, concordances were sorted into categories following functional criteria. For example, Figure 1 displays the concordances obtained when searching for nouns followed by prepositions. Since these contained both post-modifying and adverbial PPs, the occurrences needed to be sorted into two categories. Once the data were cleaned and sorted, frequencies of occurrence were normalised per 100,000 words.
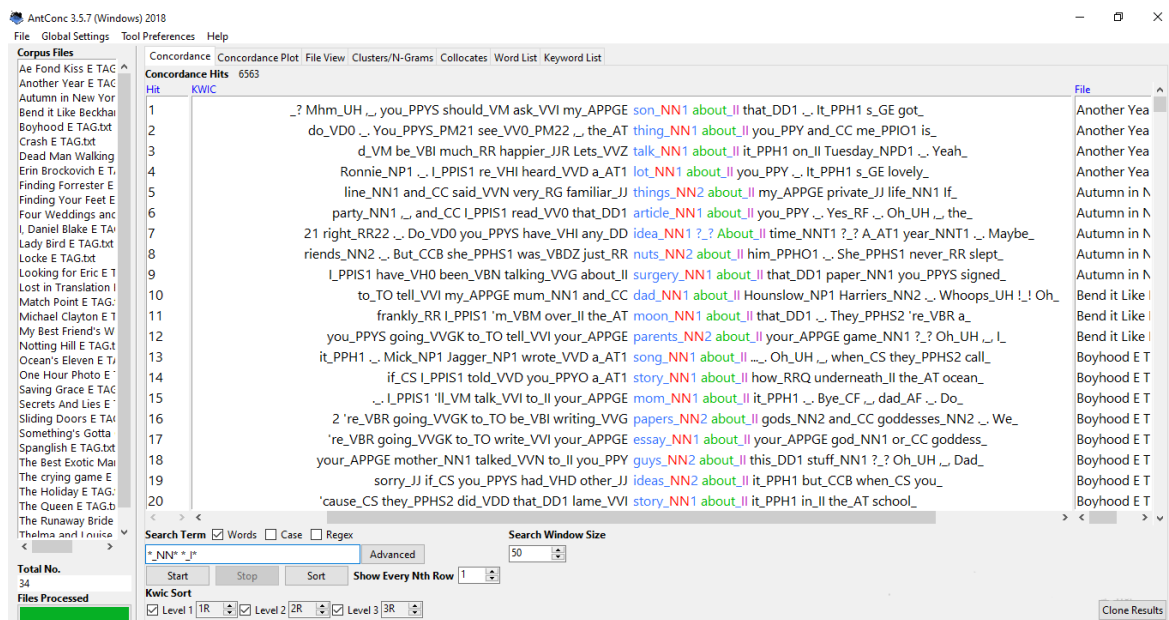

Figure 1
Sample of query results using the POS-tag sequence 'noun + preposition'.

Compounds were considered single units only when hyphenated (see examples 6 and 7), whereas complex routinised units (e.g. *prime minister*) were considered as being composed of a head noun (e.g. *Minister*) and a modifier (e.g. *prime*).

To answer RQ3, the noun heads most often (Mutual Information [MI] scores > 3; see Hunston 2002, p. 71) pre-modified by at least one adjective or noun and post-modified by a PP were analysed. The aim here was to determine what types of nouns are most frequently pre- and post-modified in film dialogue, as well as which nominal heads frequently occur with the various types of pre- and post-modifiers (e.g. nouns modified by adjectives vs. nouns modified by nouns). This part of the analysis used both collocate and cluster analyses of POS-tag combinations. To find the nominal heads most frequently pre-modified by adjectives, the POS-tag for adjectives ('JJ') was used as the search term, and a collocate search was performed for the element immediately following the adjective. Therefore, the span was reduced to two words and restricted to the right of the adjective (0L 2R). The minimum collocate frequency (MCF) was set to 20. Cluster analysis was used to obtain the nominal heads of the most frequent noun-noun combinations by setting the cluster size to a maximum of four elements and the minimum frequency to five. Finally, the nominal heads of post-modifying PPs were retrieved by conducting collocate analysis. The tag for 'any preposition in the corpus' ('I') was used as the search term, and the software was asked to find any element preceding the preposition, thus restricting the span to the left (2L 0R) with an MCF of 20 occurrences. Since a preposition can be preceded by elements belonging to different word classes, the 20 most frequent nouns were selected from the list of the most frequent collocates of prepositions. Nominal heads were subsequently analysed according to their semantic traits (e.g. general vs. specific, human vs. non-human) so that trends of combinations between the nominal head type and the pre- or post-modifying pattern could be discovered.

RQ4 was answered by comparing the corpus data with information on noun phrase complexity in natural conversation from Biber *et al.*'s (2021) study. Differences and similarities between the two registers were interpreted by intersecting the quantitative and qualitative aspects of the phrasal complexity of each register in a functional interpretation that considered production circumstances, the audience and medium-related specificities (Biber *et al.* 2022).

5. Noun phrase complexity in the PCFD

By examining the frequencies of the selected phrasal complexity patterns, film dialogue appears to have a register in which pre-modification plays a preponderant role compared to post-modification. That is, pre-modification is

Noun phrase complexity in film dialogue as input for second-language learning.
A corpus-based study from a register-functional perspective

303

three times as frequent as post-modification, with a strong preference for adjectival pre-modification over nominal pre-modification (see Table 3). As will be discussed in more detail in the following sub-sections, these tendencies may be related to the narrative requirements of film, which would also explain how different pre- and post-modification patterns trigger different categories of nominal heads.

Pattern	Tokens	Frequency (per 100,000 words)
Adjectival premodification	7857	2066
Nominal premodification	4261	1120
Post modification with prepositional phrase	4095	1077
Prepositional phrases as adverbials	10044	2641

Table 3
Phrasal complexity in the PCFD: Frequencies of phrasal patterns.

5.1 Pre-modification

The analysis of noun phrase pre-modification in the PCFD showed that adjectival pre-modification is twice as frequent as nominal pre-modification (2066 vs. 1120 per 100,000 words, respectively). The most frequent pre-modification pattern includes only one pre-modifier and, occasionally, a determiner (see example 3; see Table 4). Longer sequences of (three or more) pre-modifiers only comprise 0.6% of the pre-modification in film dialogue. Pre-modified nominal heads occur 2847.3 times every 100,000 words in film dialogue, suggesting that approximately one in four nouns in the corpus is pre-modified by at least one word.

As can be observed in Table 4, both adjectival and nominal pre-modification patterns containing only one pre-modifier (3) are more frequent than those containing two pre-modifiers (4), which, in turn, are more frequent than those containing three pre-modifiers (5).

> [3]
> You see, this is an <u>unfair advantage</u>. (*Thelma and Louise*, Scott, 1991)
> I'm calling about the, um, <u>job ad</u>. (*Erin Brockovich*, Soderbergh, 2000)
> [4]
> She's with her <u>new best friend</u>. (*Lady Bird*, Gerwig, 2017)
> We got back from the <u>pardon board hearing</u>. (*Dead Man Walking*, Robbins, 1995)
> Do you realise that a <u>severe anxiety attack</u> can masquerade as a heart attack? (*Something's Gotta Give*, Meyers, 2003)

Great, listen, Harry has a <u>really bad headache</u>. (*Something's Gotta Give*, Meyers, 2003)
[5]
Go, let her go, she's got that <u>big old bad joker</u> there […]. (*Dead Man Walking*, Robbins, 1995)
You act like they're going to play a <u>World Series championship game</u> or something. (*Finding Forrester*, Van Sant, 2000)

The only exception to the trend is the Saxon genitive pre-modification, which is less frequent than N + N + N combinations. Overall, the tendency is a preference for shorter pre-modifying sequences. This was expected since film dialogue is meant to imitate natural conversation, which tends to rely on clausal rather than phrasal features, as mentioned in Section 2 (see Biber *et al.* 2022).

Pattern	Tokens	Normalised (100,000 words)
(det) JJ + NN	6394	1681
(det) JJ + JJ + NN	441	116
(det) JJ + JJ + JJ + NN	3	0.8
(det) NN + NN	2847	749
(det) NN + NN + NN	351	92
(det) NN + NN + NN + NN	12	3.1
(det) NN's + NN	301	79
(det) JJ + NN + NN	269	70.7
(det) JJ + JJ + NN + NN	33	8.7
(det) JJ + NN + NN + NN	16	4.2
(det) NN's + JJ + NN	15	4
(det) ADV + JJ + NN	144	37.9
Total number of pre-mod nominal heads	10826	2847.3
Premod ADJ	7857	2066
Premod N	4261	1120
Premod (tot)	12118	3187

Table 4
Frequencies of pre-modifying patterns in the PCFD.

However, film dialogue also displays complex pre-modifying patterns, although these occur much less frequently than short pre-modifying patterns. The longest pre-modification pattern in the corpus includes five pre-modifiers – or six if *special-edition* is considered to comprise two words – realising a complex pre-modifying sequence (see 6). Other examples of long sequences of pre-modifiers in film dialogue feature three and four pre-modifiers (see 7).

[6]
[special-edition plastic Burger King tray cups] (*Crash*, Haggis, 2004)
[7]
The [Evans County Threat Management Unit] (*One Hour Photo*, Romanek, 2002)
[cold-decking teen beat cover boys] (*Ocean's Eleven*, Soderbergh, 2001)
[T]he [strongest possible criminal attorney] (*Michael Clayton*, Gilroy, 2007)

Functionally, pre-modification allows for the identification and characterisation of the referent of the head noun by adding information about a variety of parameters such as physical appearance (e.g. *large, thin, heavy, pale*), evaluation (e.g. *bad, nice, beautiful*), classification (e.g. *German, chemical, legal*), etc. (cf. Feist 2012; Biber *et al.* 2022). Relying on pre-modification also allows for conveying information concisely, by using as few words as possible, to save space and time. This is an efficient way to provide substantial information about characters and events, especially in the context of film dialogue, in which a story needs to be told in a limited time. However, compacting information also leads to a lack of explicitness in the relationship between the head noun and its pre-modifiers (Biber *et al.* 2022). For example, in (8), the relationship between the three modifiers *little, horse* and *shit* and the head *attitude* is not explicitly coded.

[8]
Why don't you say goodbye to that <u>little horse shit attitude</u>? (*Boyhood*, Linklater, 2014)

By drawing on background knowledge, the expression can be interpreted as [little [horse shit [attitude]]]. That is, *attitude* is first modified by the already complex N + N expression *horse shit,* and the complex expression *horse shit attitude* is further modified by the adjective *little*.

The possibility of misunderstandings due to the ambiguity caused by the reduced explicitness of pre-modifiers is generally solved by drawing on contextual information: the speaker is expected to be aware of the information that the interlocutors already know, who then use less or non-explicit constructions (Feist 2012). For film dialogue, aside from relying on the dialogue itself, film can make referents and concepts accessible with images and sounds, which may allow for the broader use of inexplicit references through pre-modification. Such a relationship between dialogue and video, however, could not be investigated in the present study due to the nature of the corpus.

Although adjectives are believed to be the prototypical pre-modifying word class, evidence has been given for the increasing use of nouns as pre-modifiers over time (cf. Biber, Clark 2002; Biber *et al.* 2022). The N + N sequence represents a convenient strategy to compact information, as it allows for using nouns with an attributive function without needing to add morphological material, as is required with derived adjectives (e.g. *danger > danger-ous, care > care-ful, love > lovely*). Similar to attributive adjectives, N + N combinations express a variety of meaning relations between the head and the modifier, such as composition (9a), source (9b) and location (9c)

(Biber *et al.* 2021, p. 584), which are not immediately identifiable, thus adding to the vagueness of language onscreen (Quaglio 2009).

[9]

a. [...] we flew a <u>paper airplane</u> off. (*Boyhood*, Linklater, 2014)

b. [...] smelled like <u>chicken fat</u>. (*Erin Brockovich*, Soderbergh, 2000)

c. […] the duty officer at the <u>Belmount Police Station.</u> (*Locke*, Knight, 2013)

The strong prevalence of pre-modification with adjectives in film dialogue indicates a preference that may be related to the function of characterisation. Specifically, adjectives tend to add characteristics that are directly related to the referent of the head they modify. These thus represent a convenient strategy to provide information about the personality, appearance and cultural background of film characters. N + N structures, on the other hand, encode relations between the N1 and N2 making up the expression and likely refer to objects rather than people (see Example 9 and Section 5.3) and so are less explicit. Hence, their meaning must be inferred from the context or background knowledge.

5.2 Post-modification

In film dialogue, nominal post-modification with a PP is not particularly frequent compared to both pre-modification and PPs with an adverbial function. Approximately, only one in 10 nouns is post-modified by a PP in the corpus compared to one in four nouns that are pre-modified. Moreover, adverbial PPs are more than twice as frequent as post-modifying PPs (2604 vs. 1063 per 100,000 words, respectively; see Table 5). This tendency highlights the preference to add information at the clausal level in film dialogue, even when phrases are used, which adds to the findings on clausal complexity in film dialogue (Formentelli *et al.*, forthcoming).

Prepositional phrases	Token	Norm.
NP + PP (post-mod)	4095	1077
PP (adverbial)	10044	2641

Table 5
Post-modifying compared to adverbial PPs in the PCFD.

Functionally, PPs add information more explicitly than pre-modifiers since the preposition clarifies how the content of the PP relates to the item it modifies (see Example 10).

[10]

Like, the Freshmen satisfaction rate <u>for new roommates</u> used to be like 60%, and now it's 100%. (*Boyhood*, Linklater, 2014)

Noun phrase complexity in film dialogue as input for second-language learning.
A corpus-based study from a register-functional perspective

307

A balance between inexplicitness and explicitness is essential in film dialogue since it needs to be understood by a wide audience who does not share much background information about the characters and story portrayed onscreen until such information is given. Simultaneously, the higher frequency of adverbial PPs, with their elements of clausal complexity, may be ascribed to the imitation of the spoken register and thus the reproduction of tendencies that are typical of conversation (see Section 2) and that are aimed at achieving naturalness.

5.3 Pre- and post-modified nominal heads

As mentioned in Section 4, the most frequent pre- and post-modified nominal heads in the PCFD were obtained via collocate and cluster analyses of the tags for adjectives, nouns and any element preceding a preposition. Due to spatial limitations, the 20 most frequently modified head nouns for each category of pre- and post-modifier (adjective, noun, PP) were selected and analysed.

Beginning with the nominal heads that were found to be most frequently pre-modified by adjectives, the collocate analysis returned the list of nouns reported in Table 6.

Collocate	MI score
Minister	18,52
News	17,96
Ones	17,79
Birthday	17,48
Person	17,48
Girl	17,35
Deal	17,29
Thing	17,24
Hell	17,23
Side	17,01
Lady	17,01
School	16,97
Idea	16,97
Boy	16,94
Woman	16,91
Part	16,66
Night	16,65
Men	16,65
Friend	16,64
Bit	16,58

Table 6
Nouns collocating with the tag 'JJ' (adjectives).

Notably, nouns that are frequently pre-modified by adjectives tend to be generic and underspecified. Therefore, they need to be further characterised in order to be identified more clearly, which is the function of pre-modifying

adjectives. This is applicable to underspecified words such as *person*, *ones*, *girl*, *thing*, *lady*, *part* and *men* (see 11).

> [11]
> a. I am this <u>horrible person</u>. (*Boyhood*, Linklater, 2014)
> b. Ah! Yeah, I like the <u>old ones</u>. (*Match Point*, Allen, 2005)
> c. I've never seen an <u>Indian girl</u> into football. (*Bend it like Beckham*, Chadha, 2002)
> d. Can you do <u>one other big thing</u> for me? (*Something's Gotta Give*, Meyers, 2003)
> e. Don't mind the <u>funny lady</u>. (*Another Year*, Leigh, 2010)
> f. This is just the <u>best part</u> of my day. (*Ocean's Eleven*, Soderbergh, 2001)
> g. Would there be any <u>wealthy single men</u> in this evening? (*The Best Exotic Marigold Hotel*, Madden, 2012)

Furthermore, the referents indicated by nominal heads pre-modified by adjectives most often identify people and everyday life events: eight out of the 20 nouns refer to people, namely *minister, person*, *girl*, *lady*, *boy*, *woman*, *men* and *friend* (see Examples 11a, c, e and g); six out of the 20 nouns refer to everyday life events and situations, namely *news*, *birthday*, *school*, *idea*, *night* and *deal* (see Example 12).

> [12]
> a. Well, <u>excellent news</u>. (*Four Weddings and a Funeral*, Newell, 1994)
> b. <u>Happy birthday</u>, dear Roxanne. (*Secrets and Lies*, Leigh, 1996)
> c. Erm, he was from this <u>private school</u>. (*Finding Forrester*, Van Sant, 2000)
> d. Arthur downtown was not a <u>good idea</u>. (*Michael Clayton*, Gilroy, 2007)
> e. See you soon. <u>Good night</u>. (*I, Daniel Blake*, Loach, 2016)
> f. Oh, what's the <u>big deal</u>? (*Erin Brockovich*, Soderbergh, 2000)

As mentioned, pre-modification in film is preferred for providing more information about the characters whom the audience 'meets' for the first time. Therefore, if *girl*, *lady*, *boy*, *woman* and *men* provide information about gender and age, adjectives provide information about personal inclinations and features of the characters. Nouns referring to people were found to display greater variation in collocating adjectives compared to other nouns (see Appendix A). Adjectives also significantly pre-modify nouns referring to everyday life events and situations, often occurring in formulaic chunks expressing speech acts such as leave-taking (12e), good wishes (12b), comments (12a, f), etc.

For the cluster analysis of N + N combinations, the 20 most frequent nouns pre-modified by at least one noun are reported in Table 7. As shown, the referents of N + N combinations are all inanimate, except for *decision maker* and *football team*. The remainder of clusters refer to material things that are part of everyday life (e.g. *phone number*, *eye contact*,

thousand/million dollars, *ice cream*), except for *death row* and *pardon board,* which are overrepresented due to their high frequency of occurrence in the film *Dead Man Walking*.

Cluster	Frequency (tokens)
Heart attack	24
Million dollars	22
Parking lot	17
Death row	13
Phone number	13
Thousand dollars	13
Marigold Hotel	12
Decision maker	10
Eye contact	9
Phone call	9
Ice cream	8
Room service	8
Pardon board	7
Blood pressure	6
Bus stop	6
Cell phone	6
Dance floor	6
Detector test	6
Football team	6
Hundred years	6

Table 7
Cluster analysis of N + N combinations in the PCFD.

Finally, the collocation analysis of items followed by PPs resulted in the following list of the most frequent nominal collocates (Table 8).

Collocate of the PP	MI score
Rid	9.10
Lots	9.04
Cup	8.96
Couple	8.70
None	8.60
Piece	8.37
Middle	8.31
Kind	8.29
Part	8.17
Lot	8.14
Front	8.10
Hands	7.91
Top	7.83
Deal	7.79
End	7.74
Bit	7.60
Sort	7.45
Side	7.37
Friend	6.88
One	6.79

Table 8
Nouns collocating with the tag '*_I*' (prepositions).

This group of nominal heads is of another type, namely identifiers of the quantities (e.g. *part, lot, bit, deal*), positions and parts (e.g. *front, top, side*) of the entities that are referred to in the dependent PP. In the most frequent N + PP combinations, therefore, the semantic core is found in the PP rather than in the nominal head, as shown in Example (13).

> [13]
> Your case looked good, <u>lots of **motives**</u>. (*Match Point*, Allen, 2005)
> I only had a <u>couple of **wines**</u>! (*Bend it like Beckham*, Chadha, 2002)
> Not in the <u>middle of **winter**</u>. (*Finding your feet*, Loncraine, 2017)

Exceptions to this trend are the nouns *hands* and *friend*. The former is frequently found in a command shouted by the police, as in (14), while the latter is followed by PPs describing a friendship (see 15).

> [14]
> a. **Hands** <u>in plain sight</u>! (*Crash*, Haggis, 2004)
> b. **Hands** <u>in plain view</u>! (*Thelma and Louise*, Scott, 1991)
> [13]
> Jamal Wallace is a **friend** <u>of yours</u>? (*Finding Forrester*, Van Sant, 2000)

The nominal heads that significantly collocate with different types of pre- and post-modifiers were found to occur in a seemingly complementary distribution, different categories of nouns occur with different modification strategies. Adjectives tend to modify nouns referring to people and everyday life events and situations, whereas nouns tend to modify nouns referring to material things and common objects. Finally, PPs tend to modify nouns describing quantities, parts of an object and object locations. What is common to all, although it is less true of N + N combinations, is the under-specificity of the semantics of nominal heads that are most frequently pre- and post-modified, thus calling for further characterisation through modification.

5.4 Comparison with natural conversation

By comparing the data on the pre- and post-modification of the noun phrase with those provided by Biber *et al.* (2021), both similarities and differences are found between the two registers of film dialogue and natural conversation. Generally, the PCFD and natural conversation display similar tendencies of pre- and post-modification (see Figure 2). In both registers, adjectival pre-modification is more frequent than nominal pre-modification, while post-modifying PPs are much less frequent than adverbial PPs. Simultaneously, film dialogue displays a higher frequency of both pre-modifying adjectives and nouns, as well as PPs, compared to natural conversation (see Figure 2).

Lingue e Linguaggi

Noun phrase complexity in film dialogue as input for second-language learning.
A corpus-based study from a register-functional perspective

311

If, on the one hand, similarities in the trends of both pre- and post-modification point to a successful imitation of the register, on the other hand, film dialogue appears to rely more on the use of phrasal complexity features than natural conversation. Such differences, however, do not seem to affect the overall tendencies observed for phrasal structures, thus indicating that register-specific needs in film dialogue (compacting information and characterisation) do not eliminate the need to sound natural.
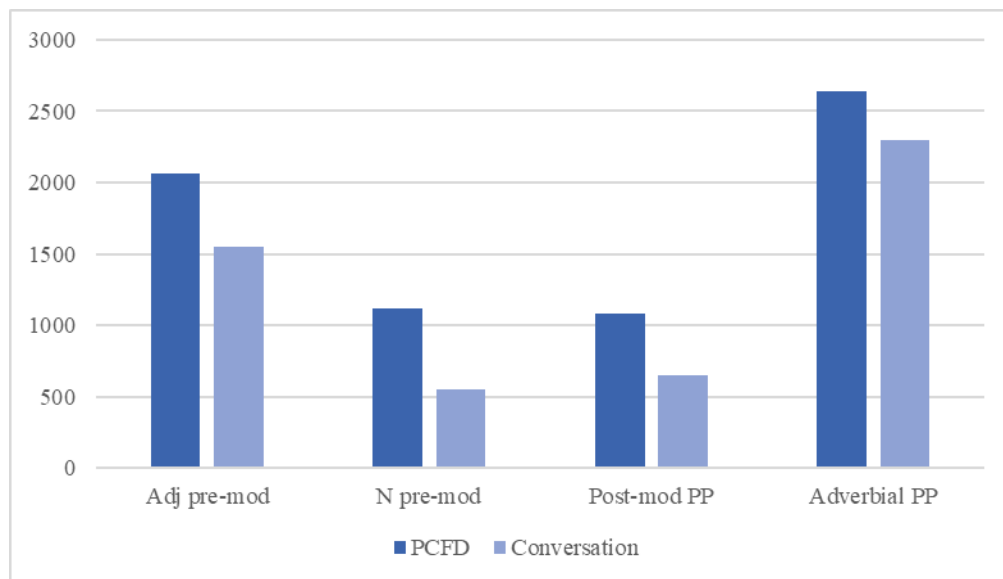

Figure 2
Adjective compared to noun pre-modification and PP modification
in film dialogue and conversation.

A final difference between pre-modification in film dialogue and conversation relates to the length of the pre-modifying sequence. Biber *et al.* (2021, p. 591) found that in conversation, a noun is not pre-modified by more than two words, whereas film dialogue displays occurrences of four-word pre-modifying sequences (excluding determiners, which would represent a fourth pre-modifying element; see Example 7).

A heavier reliance on phrasal features such as pre- and post-modification, combined with the observed increased length of the pre-modifying sequences compared to natural conversation, corresponds to a higher degree of phrasal complexity in film dialogue. This may be seen as suggestive of the production circumstances of film dialogue: since it starts as written language, there is time for pre-planning and editing. As Fox (2007, p. 314) posited, writers have no time constraints placed on their production, which presumably allows for the use of more complex syntactic structures. Simultaneously, since phrasal features allow for maximising the compacting of information, they can be used in film dialogue since much information

about the characters and the story portrayed onscreen needs to be provided in the brief time given by the film format.

6. Conclusions

The present paper investigated the phrasal complexity of film dialogue by examining nominal pre- and post-modification and the typology of nominal heads that are most frequently modified. It was observed that film dialogue most often relies on nominal pre-modification with an attributive adjective, which is followed in frequency by pre-modification with a noun and post-modification with a PP. Since phrasal elaboration allows for conveying the most information using minimal coding (Levi 1978; Biber *et al.* 2022), these represent convenient choices for film dialogue, where much needs to be said in a limited time. Both pre- and post-modified noun phrases appear to be more frequent in film dialogue compared to natural conversation, which may reflect the register-specific requirements and production circumstances in the text's complexity. Using dense nominal expressions, however, also implies a loss of explicitness in textual information, especially in the relationship between the head and its modifiers. This requires more shared and contextual knowledge for disambiguation (Varantola 1993), which can be provided through the use of multiple modes (images and sounds) in film.

The analysis of the typology of nominal heads in film dialogue revealed a complementary distribution of nouns according to the modification strategy with which they occur: pre-modification with adjectives frequently occurs with nouns denoting people and everyday situations or events; pre-modification with nouns frequently occurs with concrete referents such as objects; finally, nouns that are most frequently post-modified by a PP mainly indicate quantities or parts of the referent found in the PP.

Aside from showing an increased degree of phrasal complexity in film dialogue, considered a consequence of register-functional linguistic strategies, the comparison between film dialogue and natural conversation yielded a second observation: the tendencies displayed for both pre- and post-modification in film dialogue are similar to those displayed by natural conversation, confirming that film dialogue successfully imitates natural conversation in many ways, including phrasal complexity. This may have important implications for L2 learning since learners encounter input that is both rich and faithful to natural spoken language. Furthermore, because of the frequency of exposure, it can be hypothesised that frequent structures in input may lead to noticing and internalisation (Bley-Vroman 2002; Bybee 2008). Future studies addressing the impact of accessing media in English outside of the classroom should consider formulating expectations about the learning

*L*ingue e
*L*inguaggi

paths of phrasal complexity features following the typology and distribution of these features in film dialogue and other frequently accessed audiovisual products (see Formentelli, Zago this volume). Such learning expectations can then be tested against actual learners' productions, ideally in longitudinal studies, to gain deeper insights into the relationship between informal exposure to audiovisual English content and the learning outcomes of several grammatical complexity features.

Bionote: Liviana Galiano is tenure-track researcher at San Raffaele University of Rome. In 2020, she earned her AHRC-funded Ph.D. in English Linguistics from the University of Lancaster (UK). As former research fellow at the Universities of Catania (2023-2024) and Pavia (2020-2023), she has conducted longitudinal case studies for the 2020 PRIN project "The informalisation of English language learning through the media" and worked on the expansion and tagging of the Pavia Corpus of Film Dialogue. She has also conducted research on pragmatic aspects of film dialogue (e.g. pragmatic markers, speech acts and attitude) and the grammatical complexity of film dialogue as input for L2 learners. Liviana Galiano's main research interests include English corpus linguistics, languages of traditional and new media, pragmatics, grammatical complexity, English Lingua Franca (ELF).

Author's address: liviana.galiano@uniroma5.it

References

Alvarez-Pereyre M. 2011, *Using film as linguistic specimen: Theoretical and practical issues*, in Piazza R., Bednarek M. and Rossi F. (Eds.), *Telecinematic Discourse: Approaches to the Language of Films and Television Series*, John Benjamins, Amsterdam, pp. 47-67.

Anderson N.J., Graham S.A., Prime H., Jenkins J.M. and Madigan S. 2021, *Linking Quality and Quantity of Parental Linguistic Input to Child Language Skills: A Meta-Analysis*, in "Child Development" 92 [2], pp. 484-501.

Bednarek M. 2010, *The Language of Fictional Television: Drama and Identity*, London, Continuum.

Bednarek M. 2018, *Language and Television Series. A Linguistic Approach to TV Dialogue*, Cambridge University Press, Cambridge.

Beers S.F. and Nagy W.E. 2009, *Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre?*, in "Reading and Writing" 22, pp. 185-200.

Berlage E. 2014, *Noun Phrase Complexity in English*, Cambridge University Press, Cambridge.

Biber D. 1992, *On the complexity of discourse complexity: A multidimensional analysis*, in "Discourse Processes" 15, pp. 133-163.

Biber D. and Clark V. 2002, *Historical shifts in modification patterns with complex noun phrase structures*, in Fanego T., Pérez-Guerra J. and López-Couso M.J. (Eds.), *English Historical Syntax and Morphology*, John Benjamins, Amsterdam, pp. 43-66.

Biber, D. and Gray B. 2016, *Grammatical Complexity in academic English: Linguistic change in writing,* Cambridge University Press, Cambridge.

Biber D., Johansson S., Leech G., Conrad S. and Finegan E. 2021, *Grammar of Spoken and Written English*, John Benjamins, Amsterdam/Philadelphia.

Biber D., Gray B., Staples S. and Egbert J. 2022, *The Register-functional Approach to Grammatical Complexity: Theoretical Foundation, Descriptive Research Findings, Application*, Routledge, London.

Biber D., Larsson T. and Hancock G.R. 2023, *The linguistic organization of grammatical text complexity: comparing the empirical adequacy of theory-based models*, in "Corpus Linguistics and Linguistic Theory" 20 [2], pp. 347-373.

Bulté B. and Housen A. 2012, *Defining and operationalising L2 complexity*, in Housen A., Kuiken F. and Vedder I. (Eds.), *Dimensions of L2 Performance and Proficiency Investigating Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia, pp. 21-46.

Bley-Vroman R. 2002, *Frequency in production, comprehension and acquisition*, in "Studies in Second Language Acquisition" 24 [2], pp. 209-213.

Bybee J. 2008, *Usage-based grammar and second language acquisition*, in Robinson P. and Ellis N.C. (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, Routledge, New York, pp. 216-236.

Caruana S. 2009, *'The Italian Job': The impact of input from television on language learning*, in Borg J., Hillman A. and Lauri M.A. (Eds.), *Exploring the Maltese Media Landscape*, Allied Newspapers Ltd, Malta, pp.173-185.

Crossley S., Weston J., McLain Sullivan S. and McNamara D. 2011, *The development of writing proficiency as a function of grade level: A linguistic analysis*, in "Written Communication" 28, pp. 282-311.

d'Ydewalle G. and Pavakanun U. 1996, *Le sous-titrage à la télévision facilite-t-il l'apprentissage des langues?*, in Gambier Y. (Ed.), *Les transferts linguistiques dans*

les médias audiovisuels, Presses Universitaires du Septentrion, Villeneuve d'Ascq, pp. 217-223.

d'Ydewalle G. and Pavakanun U. 1997, *Could enjoying a movie lead to language acquisition?*, in Winterhoff-Spurk P. and van der Voort T.H.A. (Eds.), *New Horizons in Media Psychology*, Westdeutcher Verlag, Opladen, pp. 145-155.

Dynel M. 2011, *'You talking to me?' The viewer as a ratified listener to film discourse*, in "Journal of Pragmatics" 43, pp. 1628-1644.

Ellis N. and Wulff S. 2015, *Usage-based approaches in second language acquisition*, in VanPatten B. and Williams J. (Eds.), *Theories in SLA. An Introduction*, Routledge, London, pp. 75-93.

Feist J. 2012, *Premodifiers in English: Their Structure and Significance*, Cambridge University Press, Cambridge.

Ferreira V.S. and Slevc L.R. 2007, *Grammatical encoding*, in Gaskell M.G. (Ed.), *The Oxford Handbook of Psycholinguistics*, Oxford University Press, Oxford, pp. 453-469.

Finegan E. and Biber D. 2001, *Register variation and social dialect variation: The register axiom*, in Eckert P. and Rickford J.R. (Eds.), *Style and Sociolinguistic Variation*, Cambridge University Press, Cambridge, pp. 235-267.

Forchini P. 2012, *Movie Language Revisited. Evidence from Multi-Dimensional Analysis and Corpora*, Bern, Peter Lang.

Forchini P. 2021, *Corpora and movie discourse: Marvel and DC Studios compared*, in Friginal E. and Hardy J.A. (Eds.), *The Routledge Handbook of Corpus Approaches to Discourse Analysis*, Routledge, London, pp. 183-201.

Formentelli M., Galiano L., Pavesi M., forthcoming, *Grammatical complexity in film dialogue: a corpus-based study from a register-functional perspective*, in "International Journal of Corpus Linguistics".

Formentelli M. and Zago R. this volume, *Complexity matters in TV dialogue as language input. A corpus-based description of clausal and phrasal patterns*.

Fox B. 2007, *Principles shaping grammatical practices*, in "Discourse Studies" 9, pp. 299-318.

Galiano L. and Semeraro A. 2023, *Part-of-Speech and pragmatic tagging of a corpus of film dialogue: a pilot study*, in "Corpus Pragmatics" 7, pp. 17-39.

Green M.C., Brock T.C. and Kaufman G.F. 2004, *Understanding media enjoyment*, in "Communication Theory" 14 [4], pp. 11-27.

Guillot M-N. and Pavesi, M. 2019, *AVT as intercultural mediation*, in "Multilingua" 38 [5], pp. 495-504.

Haswell R.H. 2000, *Documenting improvement in college writing: A longitudinal approach*, in "Written Communication" 17, pp. 307-352.

Hunston, S. 2002, *Corpora in Applied Linguistics*, Cambridge University Press, Cambridge.

Kerswill P. and Williams A. 2002, *'Salience' as an explanatory factor in language change: evidence from dialect levelling in urban England*, in "Contributions to the Sociology of Language" 86, pp. 81-110.

Krashen S. 1985, *The Input Hypothesis: Issues and Implications*, Longman, London.

Levi J.N. 1978, *The Syntax and Semantics of Complex Nominals*, Academic Press, New York.

Lorenzo-Dus N. 2009, *'You're barking mad, I'm out': Impoliteness and broadcast talk*, in "Journal of Politeness Research" 5 [2], pp. 159-187.

Lu X. 2011, *A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writer's language development*, in "TESOL Quarterly" 45, pp. 36-61.

Montag J.L., Jones M.N. and Smith L.B. 2018, *Quantity and Diversity: Stimulating Early Word Learning Environments*, in "Cognitive Science" 42 [2], pp. 375-412.

Pavesi M. 2014, *The Pavia Corpus of Film Dialogue: a means to several ends*, in Pavesi M., Formentelli M. and Ghia E. (Eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*, Peter Lang, Lusanne, pp. 29-55.

Pavesi M. 2015, *From the screen to the viewer-learner. Audiovisual input as a context for second language acquisition*, in Campagna S., Ochse E., Pulcini V. and Solly M. (eds.), *Languaging in and across Communities: New Voices, New Identities. Studies in Honour of Giuseppina Cortese*, Peter Lang, Lusanne, pp. 83-104.

Pavesi M. and Ghia E. 2020, *Informal Contact with English: A Case-Study of Italian Postgraduate Students*, ETS, Pisa.

Quaglio P. 2009, *Television Dialogue: The Sitcom Friends vs. Natural Conversation*, John Benjamins, Amsterdam/Philadelphia.

Reinders H. and Benson P. 2017, *Research agenda: Language learning beyond the classroom*, in "Language teaching" 50 [4], pp. 561-578.

Staples S., Egbert J., Biber D. and Gray B. 2016, *Academic writing development at the university level: phrasal and clausal complexity across level of study, discipline and genre*, in "Written Communication" 33 [2], pp. 149-183.

Sundqvist P. 2009, *Extramural English Matters: Out-of-School English and Its Impact on Swedish Ninth Graders' Oral Proficiency and Vocabulary*, unpublished doctoral thesis, Karlstad University.

VanPatten B. 2015, *Input processing in adult SLA*, in VanPatten B. and Williams J. (Eds.), *Theories in SLA. An Introduction*, Routledge, London, pp. 113-134.

Varantola K. 1993, *Modification of nouns by nouns. Bad by definition?*, in Jucker A.H. (Ed.), *The Noun Phrase in English. Its Structure and Variability. Anglistik and Englischunterricht 49*, Winter, Heidelberg, pp. 69-83.

Warren M. 2006, *Features of Naturalness in Conversation*, Amsterdam, John Benjamins.

Werner V. 2021, *A diachronic perspective on telecinematic language*, in "International Journal of Corpus Linguistics" 26 [1], pp. 38-70.

Wissmath B., Weibel D. and Groner R. 2009, *Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment*, in "Journal of Media Psychology" 21 [3], pp. 114-125.

Zago R. 2015, *'That's none of your business, Sy': The pragmatics of vocatives in film dialogue*, in Dynel M. and Chovanec J. (Eds), *Participation in Public and Social Media Interactions*, John Benjamins, Amsterdam/Philadelphia, pp. 183-207.

Films

Allen W. (Director) 2005, *Match Point* [Film], BBC Films, Thema Production, Jada Productions.

Chadha G. (Director) 2002, *Bend it like Beckham* [Film], Kintop Pictures, Bend It Films, Roc Media, Road Movies, Filmproduktion.

Chen J. (Director) 2000, *Autumn in New York* [Film], Lakeshore Entertainment, Metro-Goldwyn-Mayer.

Gerwig G. (Director) 2017, *Lady Bird* [Film], IAC Films, Scott Rudin Productions, Management 360.

Gilroy T. (Director) 2007, *Michael Clayton* [Film], Samuels Media, Castle Rock Entertainment, Mirage Enterprises, Section Eight Productions.

Haggis P. (Director) 2004, *Crash* [Film], Lions Gate Films, Bob Yari Productions.

Noun phrase complexity in film dialogue as input for second-language learning.
A corpus-based study from a register-functional perspective

317

Leigh M. (Director) 1996, *Secrets and Lies* [Film], Thin Man Films, Ciby 2000, Channel Four Films.

Leigh M. (Director) 2010, *Another Year* [Film], Thin Man Films, Film4, Focus Features.

Linklater R. (Director) 2014, *Boyhood* [Film], IFC Productions, Detour Film.

Loach K. (Director) 2016, *I, Daniel Blake* [Film], Wild Bunch, The Criterion Collection.

Loncraine R. (Director) 2017, *Finding your Feet* [Film], Eclipse Films, Powder Keg Pictures, Catalyst Global Media, Fred Films, Bob & Co., Twickenham Studios, Ultimate Pictures, Protagonist Pictures.

Madden J. (Director) 2011, *The Best Exotic Marigold Hotel* [Film], Dune Entertainment, Ingenious, Participant Media Imagenation Abu Dhabi FZ, Blueprint Pictures.

Meyers N. (Director) 2006, *Something's Gotta Give* [Film], Columbia Pictures, Warner Bros. Pictures, Waverly Films.

Meyers N. (Director) 2006, *The Holiday* [Film], Columbia Pictures, Universal Pictures, Relativity Media, Waverly Films.

Newell M. (Director) 1994, *Four Weddings and a Funeral* [Film], PolyGram Filmed Entertainment, Channel Four Films, Working Title Films.

Robbins T. (Director) 1995, *Dead Man Walking* [Film], Working Title Films, Havoc.

Romanek M. (Director) 2002, *One Hour Photo* [Film], Catch 23 Entertainment, Killer Films, John Wells Productions.

Scott R. (Director) 1991, *Thelma and Louise* [Film], Pathé Entertainment, Percy Main Productions, Star Partners III Ltd..

Soderbergh S. (Director) 2000, *Erin Brockovich* [Film], Universal Pictures, Columbia Pictures, Jersey Films.

Soderbergh S. (Director) 2001, *Ocean's Eleven* [Film], Village Roadshow Pictures, NPV Entertainment, Jerry Weintraub Productions, Section Eight Productions.

Van Sant G. (Director) 2000, *Finding Forrester* [Film], Columbia Pictures.

Annexes

Adjectival pre-modifiers collocating with nominal heads

Collocate	Premodifying adjectives
Minister	Prime
News	Good, excellent, bad, unwelcome, great, nightly, dreadful, local, worse
Ones	Small, little, sissy, whole, long, old, new, darling, loved, bad, white, new, unloved, larger
Birthday	Happy, big
Person	Bad, young, whole, lovely, cooperative, horrible, real, serious, white, black, missing, bad, nice, special, elderly, right, favourite, other, nervous, good, fragile, quiet, odd, new, complicated, smart, criminal, wrong, brave, stupid, influential, terrible, suitable, wonderful, stupidest, sweetest
Girl	Good, Irish, white, gorgeous, underage, sad, old-fashioned, Indian, blonde, young, good, lucky, pretty, busy, little, silly, sweet, big, American, adorable, friendly, beautiful, poor, nice, wrong, interesting, bad, depressing, British, lovely, funniest, greatest
Deal	Great, whole, big, good, bigger
Thing	Right, whole, wonderful, other, good, thoughtless, beautiful, Indian, big, small, little, terrible, poor, polite, simple, important, silly, crucial, strange, extraordinary, unforgivable, big, exciting, different, honourable, wrong, bad, favourite, stupid, clinical, wonderful, horrible, awesome, nice, communal, ridiculous, mad, new, fun, calm, funny, toxic, dreadful, weird, terrific, best, dumbest, sexiest
Hell	Bloody, fucking, living
Side	Other, bright, wrong, American
Lady	Young, lovely, funny, old, pregnant, special, leading, nice, little
School	Catholic, ordinary, non-denominational, medical, old, whole, high, middle, prep, private, new, agricultural, public, good
Idea	Good, great, whole, fun, bad, general, daft, stupid, bright, brilliant, marvellous, wrong, superb, better
Boy	Good-looking, good, naughty, pretty, clean-shaven, white, Indian, English, small, black, little, poor, clever, handsome, nice, lucky, wonderful
Woman	Strange, married, young, mature, poor, other, perfect, proper, English, old, white, free, beautiful, middle-aged, decent, divorced, non-threatening, amazing, Mexican, promiscuous, wild, happier, younger, best
Part	Whole, big, chemical, famous, hard, great, sane, other, important, better-lit, best, easiest, deepest, creepiest
Night	Late, wild, silent, holy, good, tough, cold, other, lucky, lovely, great, rough, insane, clear, busy, early, important, wonderful, extraordinary, remarkable, interesting, special, frantic, wrong, best
Men	Strong, gay, straight, favourite, black, white, accused, great, single, cruel, divorced, modern, needy, older
Friend	Old, Indian, feathered, good, dear, furry, Scottish, special, best
Bit	Wee, little, nice, tiny, easy, best, worst