# PROFILING INFORMAL EXPERIENCE WITH ENGLISH An instrument for data collection 

MARIA PAVESI, ELISA GHIA, TIZIANA CIABATTONI ${ }^{1}$<br>Università di Pavia


#### Abstract

Within the thriving area of research on informal access to English, the present paper focuses on the methodology and design of valid, reliable and effective instruments for data collection. Studies on informal language learning have addressed different target populations, with a preference for teenagers and young adults, and have mainly used questionnaires and users' reports to gather large amounts of information. Methodological reflections on questionnaire design, however, are not always in the foreground - while sharing methodology choices is crucial for the comparability and replicability of findings. Moving from these premises, the current study describes a new instrument to investigate the experiencing and learning of English outside the classroom through multiple media. It includes a newly-developed questionnaire (the Informal English Contact and Learning questionnaire, IECoL; see Appendix) and a receptive vocabulary test adapted from Nation (1990). The tool was designed as part of an ongoing large-scale national (PRIN) project on the informalisation of English language learning among university students in Italy. Both components are described in detail, focusing on the rationale, the innovative aspects, the development and piloting phases, and the final structure and layout. Cronbach's Alpha coefficient was calculated on the 4,000 valid questionnaires, showing the reliability and internal consistency of the instrument as a whole. We conclude by outlining the innovativeness of the instrument and by underlining the paramount importance of data collection methods and tools in a field where users' reports and recall are the main sources of information about respondents' behaviour and private worlds.


Keywords: informal English learning; questionnaires; data collection; test validation; media input.

## 1. Introduction

Recent years have seen a novel interest towards informal - i.e. spontaneous, untutored and mainly incidental - language learning (Caruana 2021; Dressman, Sadler 2020; Pavesi, Ghia 2020). This interest goes hand-in-hand with L2 users' growing access to English outside the classroom in today's language and media-saturated world (e.g. Dressman, Sadler 2020; Reinhardt 2022). Information about actual L2 users' behaviour, however, is difficult to access directly and researchers have had to rely on learners' reports, as mainly obtained through questionnaires and less frequently interviews and language diaries. The rationale and design of the data collection instruments are thus crucial and need to be described and reflected upon to guarantee the transparency, comparability and replicability of results within the research community - even more so as the development of the research tools and the interpretation of participant responses often go unaddressed in the literature in applied linguistics (Dörnyei, Dewaele 2023).

[^0]Based on these considerations, this article proposes a new instrument for investigating the experiencing and learning of English outside the classroom and through a variety of media. The tool combines the newly constructed Informal English Contact and Learning questionnaire - IECoL - and a receptive vocabulary test. Both components of the tool were developed as part of an ongoing large-scale national (PRIN) project on the informalisation of English language learning through the media involving four large- and middle-sized universities in Italy. By using a mixed-method approach of data collection, the project probes Italian university students' private worlds and the undetected processes that are shaping English informal learning today. As for the general framework, it moves from the assumption that input is essential to second language acquisition (SLA) in terms of frequency, salience and interactional modifications, in line with Krashen's Comprehensible Input Hypothesis (Krashen 1985), usage-based approaches (Ellis 2003) and the Interaction Hypothesis (Gass 1997; Long 1996).

The present article is organised as follows. In the next section, we provide an overview of research on informal learning focusing on survey data collection methods. Subsequently, IECoL and the receptive vocabulary test are presented and discussed, along with the phases involved in their design - rationale, validation, piloting, item analysis. We then describe the upload of the questionnaires onto a tailor-made online platform and the procedures adopted to guarantee participants' anonymity and privacy. A general discussion coupled with an acknowledgement of the research limitations and concluding remarks ends the article.

## 2. Background to the study

Out-of-the-classroom or beyond the classroom L2 learning has become crucial in a world where globalisation and mobility have expanded language affordances and multiplied the settings in which language acquisition can take place. Yet, such remarkable changes in learning spaces and in the dynamics of L2 users' personal and social behaviours escape easy observation. Moreover, it is difficult to assess the learning outcomes that issue from such language practices, although many studies are now available which empirically support the positive impact of informal exposure to English on the learning of the language (Azzolini et al. 2022; De Wilde et al. 2020; Muñoz, Cadierno 2021; Puimège, Peters 2019). Investigations mainly divide into theoretical analyses of the constructs and contexts of informal language learning, and empirical investigations of learner-users' habits, media-induced learning trajectories and proficiency levels. The former research thread has come up with an abundance of terms to refer to informal language learning, including LBC (Language Learning Beyond the Classroom, Benson 2011), OILE (Online Informal Learning of English, Sockett 2014) and IDLE (Informal Digital Learning of English, Lee, Dressman 2018). Whereas research in this realm is inextricably linked with sociological analyses of the new contexts of L2 learning and the changing nature of media multiliteracies (Lütge 2022), empirical investigations involve large-scale studies of different populations or case surveys to explore the nature and extent of informal contact with English (e.g. Arnbjörnsdóttir, Ingvarsdóttir 2018; Azzolini et al. 2022; Kusyk 2020; Muñoz 2020; Pavesi, Ghia 2020, among others). In Europe, systematic research on informal language contact and learning has been conducted in several geographical areas over the last 20 years, at first mostly in Northern European countries (Arnbjörnsdóttir, Ingvarsdóttir 2018; Leppänen et al. 2011; Puimège, Peters 2019; Sundqvist, Sylvén 2016; Verspoor et al. 2011), while more recently including France, Germany (Kusyk 2020;

Sockett 2014), Spain and Italy (Muñoz 2020; Pavesi, Ghia 2020). Overall, the digital age has recently brought radical changes across the continent, mainly through the advent of the social Web and the spread of subscription video-on-demand (SVOD) platforms, which may be generally narrowing the gap between foreign and second language learning in younger generations (Ghia, Pavesi 2021; Sockett 2014).

Although studies on informal language learning keep flourishing, some issues still call for researchers' in-depth attention. First, there is a need for a fine-grained description of which language media, registers and genres learner-users of English access, since comprehensible input - unidirectional, interactional, spoken, written, oralised, multimodal etc. - is a sine qua non of SLA (VanPatten, Williams 2020). The correlation between exposure types and the development of different L2 skills also requires extensive research to unveil which learning outcomes are more likely with given types of exposure in different settings and language combinations. Attitudes have also been found to correlate with learning success (Aiello 2018) and have been seen to change through exposure to the language in ways that depend on different factors including amount, length and type of extramural exposure to English, perceived difficulty or accessibility of the language, and experience with language instruction (Berns et al. 2007; Arnbjörnsdóttir, Ingvarsdóttir 2018; De Wilde et al. 2020). More information is hence required on L2 users' perceptions, assessments and beliefs developed when experiencing English in the wild. As a result, more research is called for to design increasingly refined instruments to collect data on changing linguistic and learning landscapes and measure different constructs - informal contact, engagement, motivation, learning outcomes (see Arndt 2023).

### 2.1. Surveys on informal experience with English: An overview

Questionnaires have become widespread means to investigate several research issues in linguistics and applied linguistics (Dörnyei, Dewaele 2023). They allow researchers to collect vast amounts of data within a reduced timespan and guarantee high comparability among participants. Thanks to recent technological advances, they are also relatively easy to administer and process for statistical analysis (Iwaniec 2019). Despite their advantages and broadening use in applied linguistics research, several criticisms have been moved against questionnaires as means to collect data on language behaviour and language learning (Dörnyei, Dewaele 2023, pp. 8-9). Questionnaires have been often criticised for yielding data that are not reliable or valid and for not bringing out historical ontology. The use of surveys can also be challenged as these may provide superficial data on the issues they investigate, a possible consequence of the little time and effort participants are typically willing to invest in the task. Questionnaires, moreover, are viewed as limited because they lead respondents through set paths and options, hence constraining their agency. Many of these criticisms, however, can be addressed if researchers are aware of the limitations of the tool and complement survey results with deeper qualitative data through mixed method approaches. When applied, etically- and emically-oriented research methodology allows for triangulation, a crucial resource for data validity. Moreover, provided care is taken in designing, validating and piloting the tools, questionnaires can yield valid and reliable results (Iwaniec 2019). At the same time, agency can be exercised thanks to the social-interactive nature of questionnaire-taking as a communicative practice that relies on the meaning-making strategies of everyday conversation when responses are freely supplied (Schrauf 2016 in Dörnyei, Dewaele 2023, p. 8).

In the first, large-scale study of informal contact with English, Berns et al. (2007) investigated the habits of groups of school pupils in Belgium, the Netherlands, France and

Germany (for a total of 2,248 respondents). Data were collected through paper-and-pencil questionnaires which comprised biographical and socio-economic questions, items about English language proficiency and use - individually and within the family -, queries on everyday contact with English out of school and attitudes towards the language. Behavioural questions addressed both frequency and length of exposure to English through different media and live-abroad experiences. Attitudinal questions and selfassessment statements were also included in the form of Likert scales: Respondents had to rate the importance of English and their appreciation of the language as well as the ease or difficulty encountered in performing given activities in the L2. The detailed questionnaire was coupled with an EFL Vocabulary test (Meara 1992), whose results allowed for statistical analyses on which variables were "causally related" (Berns et al. 2007, p. 71).

Many survey-based investigations on informal language learning in Europe followed, with countries whose citizens had long experienced informal contact with the L2 through subtitled English-language audiovisual programmes carrying out major surveys first. Leppänen et al. (2011) conducted a nation-wide survey on the status of English in Finland by administering an extensive questionnaire to 1,495 Finnish citizens aged 15-79. The questionnaire investigated respondents' experience with English, attitudes towards it and opinions about English in Finland. Given the long-standing use of English in the media, Belgium and the Netherlands have also been live laboratories for research on informal access to English since the turn of the century (see also Kuppens 2010; Puimège, Peters 2019; Verspoor et al. 2011, among many). More recently, De Wilde et al. (2020) explored the habits and the level of media-induced competence in English of about 800 primary school pupils in Flanders prior to the start of classroom instruction in English which begins at 12-13 years old in the region. Being administered to very young participants, the questionnaire was not overly detailed, but allowed the researchers to isolate out-of-the-classroom exposure as a variable affecting knowledge of the foreign language. With a focus on respondents’ life span, Arnbjörnsdóttir and Ingvarsdóttir (2018) conducted extensive research in Iceland based on telephone surveys administered across time to people of different ages. It should be underlined that thanks to a solid survey methodology the study has relayed the longitudinal picture of an entire country, where contact with English in everyday life has been pervasive for many decades. In another context of high exposure to English, i.e. Sweden, Sylvén and Sundqvist (e.g. Sundqvist, Sylvén 2014; Sylvén, Sundqvist 2012) used questionnaires to investigate the habits of young Swedish learners of English engaging in a variety of extramural English activities and their motivations for informal exposure. The questionnaires specifically targeted school learners, who completed the survey during class time, allowing for the exploration of both formal and informal learning.

Research on informal contact with English has also been carried out in typically dubbing countries, where access to English-language media is a relatively recent phenomenon. A series of surveys conducted by Geoffrey Sockett and his colleagues are an example (e.g. Kusyk, Sockett 2012; Toffoli, Sockett 2010). The studies drew on questionnaires to explore informal access to different English media by French university students. Both frequency and length of exposure were investigated, thus enabling the researchers to quantify total hours of contact with English (Sockett 2014). Pavesi and Ghia (2020) in turn presented the first study on university students' informal contact with English in Italy and looked at participants' habits and motivations for accessing different input sources in L2 English. Among the different input types potentially available to L2 users, the questionnaire specifically zoomed in on audiovisual input from TV series and films, which was predicted - and confirmed - to be the major source of contact with

English as a result of the recent availability of subtitled products. Concurrently, Kusyk (2020) explored the habits of French and German L2 users of English (specifically, 953 university students), while also tapping into participants' motivations for engaging in different types of media-based activities. Innovatively, Kusyk's fine-grained questionnaire was coupled with longitudinal follow-up case studies to access L2 acquisitional paths through Complexity-Accuracy-Fluency measures as a result of media exposure. The same methodological framework was adopted by De Riso (2023), who combined a questionnaire on type and extent of informal experience with longitudinal case studies on individual L2 development in high-exposure learner-users of English attending an Italian university. A set of studies were also conducted in Spain by Muñoz and her colleagues. Muñoz (2020) explored the language habits of secondary school and university students from Catalonia. Data collection was carried out through a previously validated and tested questionnaire including a biographical section and sections on English language learning and frequency of out-of-school L2 activities.

Contrary to the general tendency to focus on single countries (but see Berns et al. 2007 and Lindgren, Muñoz 2013), a wide-scale study was carried out exploring the language competences of European teenagers from 14 European countries with different backgrounds in terms of foreign language instruction and opportunities for informal contact with English (Azzolini et al. 2022). This major study assembled EU data using the SurveyLang questionnaire on individual characteristics, social/educational background, out-of-the-classroom exposure habits and formal instruction in foreign languages. Focusing on English, the authors gathered over 20,000 questionnaires and importantly correlated the above variables, alongside language distance, with the outcomes of standardised proficiency tests.

Within research on informal language learning, a crucial issue is whether and how survey participants' proficiency in the L2 is also measured. Most survey-based studies include self-assessment questions, which require respondents to report on their perceived level of proficiency in English (Kusyk 2020; Pavesi, Ghia 2020). This, however, is a rather unsatisfactory procedure on its own since it is strongly subjected to respondents' non-objective evaluation. Some studies instead rely on pre-assessed/default samples, i.e. specific student cohorts assigned to a given proficiency level based on school grade or level of education (Muñoz, Cadierno 2021; Sundqvist, Sylvén 2014), a procedure which however allows for limited inter-subject and inter-group comparability. Only in few cases have language assessment tests accompanied questionnaire administration, mainly in the form of receptive vocabulary tests (De Wilde et al. 2020). Performance in these tests allows systematic correlations with respondents' behaviour. As an example, Berns et al. (2007)'s questionnaire was followed by a paper-based version of Meara's 1992 vocabulary test, in which learners were presented with 120 words and non-words and were asked to discriminate between them. Kusyk and Sockett (2012), in turn, tested vocabulary knowledge through a vocabulary knowledge scale made up of frequent 4 -grams extracted from a corpus of TV-series dialogue. González Fernández and Schmitt (2015) accompanied a questionnaire on formal and informal access to the L2 with a productive collocation test. In very few studies, more L2 skills were tested through integrated tasks. De Wilde et al. (2020) administered questionnaires to $10-12$-year-old pupils while concurrently testing their listening comprehension, reading, writing and speaking skills in L2 English. Azzolini et al. (2022) gathered data on participants' English proficiency from reading, writing and listening tests based on the Common European Framework of Reference (CEFR).

In all these investigations, the results of the questionnaires were set in relation to participants' test scores and generally showed a positive correlation with out-of-school access to English input, but often with other variables as well, such as formal instruction or parents' socio-economic status. Moreover, as most questionnaires plus test tools were administered to secondary school students, little insight is available into the relationship between informal experience with English and L2 competence in higher education.

## 3. The study

### 3.1. The IECoL questionnaire

### 3.1.1. Background, rationale and aims

IECoL draws on major research on language contact profiles (Collentine, Freed 2004; Freed et al. 2004) and informal language learning (Berns et al. 2007; Sockett 2014; see Section 2). It builds on previous work conducted at the University of Pavia, a historical, multidisciplinary middle-sized university in Italy, where several investigations were carried out on informal language learning with an attention to data collection tools (e.g. Casiraghi 2016; Cravidi 2016; Ghia, Pavesi 2021; Pavesi, Ghia 2020). Pavesi and Ghia (2020) designed an all-inclusive questionnaire that relied on the outcomes of questionnaires, focus groups and semi-structured interviews previously administered at Pavia (Casiraghi 2016; Cravidi 2016) and integrated emic aspects that had emerged from those initial data collections. The questionnaire specifically explored various forms of informal contact while focusing on audiovisual input in the varied forms it can take in today's world (dubbed or subtitled, with English or Italian subtitles, in other languages) among postgraduate students who were having little or no instruction in English at the time. The participants, who did not include language majors, showed varied contact with different types of out-of-the-classroom input. They tended to engage in receptive activities such as watching films, videos and TV-series, while they scarcely interacted with other speakers online or offline. Learners' reported reasons for accessing English input informally varied from language-oriented to hedonic-affective and social ones.

As IECoL is an updated version of the questionnaire in Pavesi and Ghia (2020), it targets university students, given their centrality in the process of language change (Ferguson 2015, p. 15). At the same time, it is meant as a broader and more flexible resource to investigate different university contexts in Italy, so as to paint a bigger picture of the ways English is accessed out of formal educational settings by students in higher education in Italy. Its present make-up is meant to allow for adaptations and applications to different age, schooling and social groups. Like previous profiling questionnaires, it investigates frequency and length of contact with English in different settings but it additionally focuses on learner-users' reasons for accessing English in a variety of unguided environments. This is in line with recent research on informal second language practices, which explore learners' engagement as a key, multidimensional construct involving behavioural, cognitive and language-specific components (Arndt 2023).

### 3.1.2. The structure of the questionnaire

The questionnaire is in Italian, it is anonymous and is divided into three macrosections. The first macrosection contains factual questions and gathers general information about

the participants' language background, including: Number of known languages, L1, age of onset of English instruction at school, enrollment in extramural English language courses. The second macrosection of the questionnaire includes behavioural questions and focuses on exposure to various input types; it is in turn divided into subsections: Films, TV-series and TV-programmes; YouTube videos; video games; songs and song lyrics; the internet further divided into reading on social networks; posting contents on social network; reading blogs and forums; posting contents on blogs and forums; reading web pages; listening to podcasts; listening to internet radio; using apps; using online shopping websites; books, magazines and newspapers; additional input sources (email writing; online video calls; face-to-face interaction; chats; using English when travelling). For each input type, the questionnaire explores overall access, frequency of access, length of access per session, access modality (e.g. with or without subtitles when watching audiovisuals), subgenres (e.g. film genres; video game categories), supports and motivations. The following example shows two questionnaire items about web pages with reference to the same set of activities (English translation in brackets):
(1.a) 5.1. Quanto spesso svolgi le seguenti attività in lingua inglese? [How often do you engage in the following activities in English?]

Activities: 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]

Frequency ranges: 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]
(1.b) 5.2. Per quanto tempo usi internet in lingua inglese per le seguenti attività complessivamente il giorno in cui lo fai? [How long do you surf the internet to engage in the following activities in English as a whole on the day you do so?]

Activities: See the activities listed in 1.a
Time ranges: Più 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]

The third macrosection includes more behavioural and attitudinal questions, as it gathers participants' demographic data while broadening the picture concerning participants' linguistic profiles through additional details about language background and language activities outside the traditional language classroom. These include age of onset of informal access to English (if this occurred at all); attendance of EMI university programmes; attitudes towards English; informal access to other L2s; study abroad experience; perceived proficiency level in English. Additional social and cognitive variables are investigated for further data correlations, i.e. parents' education and occupation - following the Italian National Institute of Statistics (ISTAT)'s Classification of Occupations, and possible participants' learning disorders.

Most items are closed questions requiring multiple-choice and drop-down answers. Open questions were kept to a minimum to avoid excessive variability in learners' answers
and to facilitate data grouping and analysis. However, an open-ended "other" option is included in most items, so that participants could indicate and describe alternative answers to the ones listed, at the same time allowing to gather an emic perspective on a few relevant issues.

### 3.1.3. Innovative aspects of IECoL

With reference to previous questionnaires, IECoL presents several innovative features that were introduced with the aim of capturing respondents' experience with English at a high level of granularity. Hence, together with the usual questions on duration and type of formal instruction received, items target the whole range of activities L2 users may be involved in in the wild by delving in detail into the text-types and multimodal genres students may access - the underlying assumption being that different types of language input as provided by different genres and registers will impact on the L2 learner-users will develop. The items in 2 and 3 show examples of the subgenres considered for audiovisual input and YouTube:
(2) 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]
(3) 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]
IECoL also introduces more refined questions about exposure to music, which previous studies have repeatedly shown to be a major, often the most widespread source of contact with English (Toffoli, Sockett 2014; Ludke 2020). By drawing on interviews with students and group discussions among team members, we focused on modality of access to song lyrics. The following questionnaire item (Example 4) taps students' degree of attention to lyrics when listening to English-language songs and access to song lyrics through audio
streaming platforms - two activities that are more likely to promote language learning than mere listening to the audio-track (Ludke 2020; McCarthy 2009; Toffoli, Sockett 2014):
(4) 4.1 Quando ascolti canzoni in lingua inglese, ti capita di concentrarti sul testo delle canzoni? [When you listen to songs in English, do you ever focus on lyrics?]
sì [Yes]
no [ No ]
4.4. Accedi al testo delle canzoni in lingua inglese tramite programmi di streaming audio o altre risorse sul web? [Do you search for English song lyrics through audio streaming apps or other web sources?] sì [Yes]
no $[N o]$
We believe that asking about whether respondents are involved with the lyrics rather than simply listen to English language songs (presumably mostly without attending to the verbal text) will help clarify the contradiction often reported in research on informal language learning between massive exposure to music and limited or null acquisition gains resulting from such exposure (e.g. De Wilde et al. 2020, p. 178; González Fernández, Schmitt 2015; Kuppens 2010).

Along similar lines, in IECoL systematic attention has been paid to the distinction among receptive, productive and interactional involvement by asking questions about reading, watching and listening to media texts vis-à-vis engaging in various types of writing activities and interacting with other English language users. Items on the internet, whenever relevant, differentiate between reception and production, as in Example 1.a above (e.g. "I read posts and contents on social networks" vs. "I write contents on social networks"). A distinction between the two is relevant acquisitionally, as predicted by the comprehensible output hypothesis (Swain 1995; 2005). Interaction in turn is posited to be crucial to comprehension and L2 development, as it combines positive and negative input, internal learner capacity and language output (Gass, Mackey 2020; Long 1996). New media provide key interactive and negotiation opportunities for naturalistic SLA that have yet to be fully explored. Questions on video games specifically tackle receptive versus interactive practices by asking whether respondents play video games on their own or with other players online (see also Muñoz 2020). The following questionnaire item (Example 5) presents a list of game types, which instantiate more or less interactive playing modalities:

```
(5) 3.3. Quali categorie di giochi usi in genere? [Which video game types do you usually play?]
__ Avventure grafiche [Graphic adventures]
__ Visual novel e/o interactive fiction [Visual novel and /or interactive fiction]
| Action-Adventure (per es., Stealth, Survival horror, ecc.)
__ Azione / Picchiaduro / Sparatutto [Action/fighting/beat 'em up/shooter games]
_ _ Immersive sim (per es. city-building, artificial life, sandbox, ecc.)
___ Giochi di ruolo (per es. RPG, MMORPG, Open world) [Role-playing]
_ Simulatori (per es. di volo; di guida; sportivi, ecc.) [Simulator (e.g. flight simulators; racing simulators;
sports simulators,etc.]
__| Strategia (real-time, turn-based, ecc.) [Strategy]
__| Puzzle games
__| Giochi di tipo educational [Educational games]
```

Overall, the questionnaire is meant to extract data on engagement with new web 2.0 media affordances in younger generations, for whom we expect a transition to greater interaction online (Sockett 2014). Interestingly, De Wilde et al. (2020)'s recent investigation on preinstruction informal learning of English among Dutch-speaking children and young
adolescents has revealed that interactive activities including using social media, speaking in English and gaming are most beneficial for the development of overall L2 proficiency.

Importantly, due the great variability of the input accessed informally and the wide range of Englishes available through the media, attention was also given to the native/nonnative dimension. In particular, a survey item (Example 6) asks how often participants interact online in English with both native and non-native English speakers - an aspect that can lead to meaningful insights into the spread of English as a lingua franca on the web and into users' awareness of and attitudes towards different Englishes (cf. Aiello 2018). The nativeness vs. non-nativeness of the input learner-viewers receive is also relevant to the acquisition of English. We might expect different proficiency levels depending on whether native or non-native English input is accessed - if native norms are assumed to be the target - and whether interactions occur with an orientation towards English as native language or English as a foreign language or lingua franca.
(6) 5.7 Su internet interagisci in inglese: [On the web you interact in English:]

Interlocutors: Con parlanti nativi di inglese [With native English speakers]; Con parlanti non nativi di inglese [With nonnative English speakers]

Frequency ranges: Molto spesso [Very often]; Spesso [Often]; Qualche volta [Sometimes]; Raramente [Rarely]; Mai [Never]

More generally, current research has stressed the multilingual context in which language learning increasingly develops today (Ortega 2019). The questionnaire hence explores respondents' access to other foreign languages through the media, asking about which languages and which media are involved (Example 7).
(7) 12. Nel tuo tempo libero, accedi ad altre lingue straniere? [In your free time, do you access any other foreign languages?]
sì [Yes]
no $[\mathrm{No}]$
12.1. A quali altre lingue straniere accedi nel tuo tempo libero? [Which other languages do you access in your free time?]

In IECoL, we have extended the questions about the reasons for access to the various activities and we have differentiated them according to the specificities of each one. Reasons range from a desire to learn the language and an appreciation of the language itself to availability and sociality. Exploring learner-users' reasons for accessing different input sources in English allows for a more in-depth description of behavioural patterns in the sample. It also lends itself to correlations with participants' self-reported or assessed proficiency in the foreign language. Additionally, the main driving factors to L2 exposure may be linked to the affective and attitudinal dimensions, which are strong predictors of acquisition (De Wilde et al. 2020).

A final specificity of IECoL are the queries on the supports used to access media genres, which include TV, computers, tablets and smartphones. This information is essential to investigate L2 users' spatial collocation and mobility in their contact with English. In line with the expansion of places in which SLA at present can take place and in which English is used as a lingua franca, we expect an increase in access to English in the wild while users are on the go. In particular, using mobile devices such as smartphones
implies freedom of access to many English media in any place and at any time (KukulskaHulme 2009). Details on which supports are used for each activity will inform us about the spatial dimension of informal language learning (Benson 2021) and tell us how much the English language permeates and melts in L2 users' daily lives. They will also unveil important aspects on the social dimension of media access, which may have changed through the years and with the diffusion of individual practices - through personal computers and online platforms - as opposed to more traditional and intrinsically more communal and sociable TV watching (Caruana 2021).

The questionnaire was developed in stages, which also involved several piloting sessions and revisions. In its final format, IECoL includes a total of 72 items. To avoid the risk of fatigue and irritation, the structure of the questionnaire is highly varied in terms of item format and question distribution. Furthermore, its migration onto an online platform (see Section 3.3) allows for an interactive navigation through items based on respondents' answers and the extraction of exact data relating to completion time.

The full questionnaire is available in the Appendix. It is not in its original format but has been compacted due to space limitations.

### 3.1.4. Piloting and item analysis

As mentioned above, the initial item pool drew on previous questionnaires on language exposure (Berns et al. 2007; Collentine, Freed 2004; Freed et al. 2004; Pavesi, Ghia 2020) and exploratory data (Casiraghi 2016; Cravidi 2016). IECoL was progressively validated and refined through discussion within the PRIN research team, with its piloting being concurrently carried out.

Initial pilot versions of the questionnaire were administered on paper to small groups of adults, including both specialists and non-specialists in the field of data collection, and to small classes of university students. This phase involved a general examination of questionnaire support (paper vs. digital), overall structure and constituting items. Through item analysis, we were able to identify problematic areas in questionnaire structuring and phrasing. For example, since respondents found it hard to precisely quantify exposure to different input sources, we introduced time range options to be selected from checklists (e.g. more than one hour, between one and two hours, less than 30 minutes). In parallel, small changes were made to item phrasing while the questionnaire was migrated onto an online version on Google forms. This made filling out more manageable and less time consuming as respondents are directly led to the sections pertaining to the activities they actually engage in. Moreover, the online format allows researchers to analyse and process questionnaire responses more immediately and effectively. By replacing most open-ended questions with multiple-choice and drop-downoption answers, a greater level of accuracy was ensured in the selection of response categories. The digital version of the questionnaire was tested in a following piloting phase on larger university classes, i.e. a similar population to the target sample. Respondents appreciated the variety of activities included in the questionnaire, which they said closely reflected what they actually experienced in their lives and made them aware of the potential of informal input for language learning.

Following the piloting phases, IECoL was administered in university classrooms to avoid the self-selection bias (see Dörnyei, Dewaele 2023) - since candidates who volunteer to take part in a study are usually the most motivated ones. Members of the research team visited the target classrooms and provided instructions for accessing the questionnaire through QR -codes and individual login credentials. On average, it took
respondents 24 minutes to complete IECoL, a reasonable time for filling out a questionnaire before fatigue or irritation were experienced by participants, which may jeopardise the validity of the instrument ${ }^{2}$.

### 3.2. The vocabulary test: Design and development

In line with much research on informal access to English, we decided to integrate the questionnaire with a vocabulary test, with a view to supporting and integrating participants' self-assessment of their English proficiency and correlating vocabulary test scores with exposure data. A vocabulary test appeared particularly suitable "since lexical knowledge has been shown to relate to success in reading, writing, and general language proficiency" (Laufer et al. 2004, p. 202; Schmitt et al. 2020). It thus allowed us to avoid a long testing session.

Numerous attempts were made to find the most suitable test type, including Meara's EFL Vocabulary Test (1992), Webb et al. (2017)'s updated Vocabulary Level Tests and Martinez and Schmitt (2012)'s test on lexical expressions. Several piloting sessions of each test and adaptations of the same tests were conducted on different classes and individual groups. These tests were not considered ideal for our target population for different reasons: They were not challenging enough for the participants' expected proficiency levels, they contained many Latinate words or they tested recognition rather than comprehension.

Our final choice is an adaptation of Nation's Vocabulary Levels Test (VLT, available online at https://www.lextutor.ca/tests/vlt/?mode=test). VLT was selected since it includes items from different frequency ranges based on corpus data (most frequent $2,000,3,000,5,000$ and 10,000 word families) and is suitable for assessing vocabulary size and knowledge at different levels of EFL proficiency. VLT measures receptive word knowledge and form-meaning mapping, as it asks test takers to match target words with their definition. By relying upon word families, the test also assesses deeper word knowledge, involving morphological aspects and word-class assignment (Nation 2021). An advantage of Nation's VLT is the presence of the 10,000 word family frequency range, which allows to test higher levels of lexical knowledge while limiting the risk of inflating learners' vocabulary size (Nation 2021, p. 3).

The test is structured into checkboxes containing definitions and matching and non-matching items (i.e. distractors); test takers are required to pair each definition with the corresponding item. Piloting sessions were conducted on VLT to perform item analysis, which led us to introduce a few changes to the original test format. Editing mainly ensued from the characteristics of the target sample in terms of expected L2 proficiency and participants' L1, as well as from contextual factors such as time available for test administration and test fatigue. Cognate words of Latinate origin (e.g. celebration, agriculture, meditate) had a high recognition rate (between $83 \%$ and $100 \%$ ) in spite of their low frequency. For this reason, they were removed from the test and replaced with words of Germanic origin. The new target items were chosen from the distractors for the same test query, so that they belonged to the same frequency range (e.g. ridge, rope, toss), and definitions for those new items were phrased at the corresponding CEFR level

[^1](matching the level of the original items included in the test). Since students in the pilot group took up to 25 minutes to fill out the test, the initial version was abridged by deleting some lower-level items. The final test includes a total of 63 items ( 12 items in the 2,000 frequency range; 15 in the 3,000 range; 18 in the 5,000 range; 18 in the 10,000 range). Items were grouped into 21 checkboxes containing three target words each. Being uploaded onto a Google Form, all test items were randomised at each access, so that learners would not perceive the growing difficulty in levels and would be discouraged from cheating. Figure 1 shows a sample checkbox from the 5,000 frequency range:


Figure 1
Checkbox from the 5,000 frequency range.
In the final editing phase we decided to add a 10 -minute timer to the form, as timed tests favour access to implicit rather than explicit knowledge - and the type of learning we intended to tap may most likely be linked to implicit knowledge (Muñoz, Cadierno 2021, p. 194). This addition also allowed us to homogenise the time required to fill out the test, thus limiting inter-subject variability in completion time. The test did not automatically close after 10 minutes, but test takers were encouraged to send out their answers once the time limit had been reached, with time of test submission being automatically recorded.

### 3.3. Privacy, online IT platform and automation of the administration procedure

After the initial piloting phases, both IECoL and the vocabulary test were migrated onto an ad hoc online IT platform. The reasons for using an IT platform were several. First and foremost, it caters to the need to grant participants' anonymity and protect their privacy by complying with the Privacy Protection Rules on sensitive data, Art. 13 EU Regulation 2016/679-RGDP. Uploading the two Google Forms - IECoL and vocabulary test - onto the platform granted full anonymity and data encryption and entailed no automatic saving of the respondents' email addresses. At the same time, a website was created to collect sensitive data from those candidates who were willing to take part in the follow-up sections of the study. Secondly, the IT platform allowed participants to access both forms in a single session thus generating form pairs (i.e. each questionnaire matching a lexical test by the same candidate).

In the initial piloting phases, candidates accessed the two forms separately through individually generated credentials. This led to a high margin of error, since respondents sometimes misspelt or changed the identification code in the two forms or two or more respondents generated the same code. To obviate the problem, we implemented a procedure for generating unique alphanumeric codes to assign to every participant. Each
student received an anonymous identification code and a QR Code for logging in to the platform. Students who after completing the questionnaire and the lexical tests were interested in participating in the subsequent phases of the research project were asked to provide their contact details. These data were later archived on the database created specifically to protect the privacy of sensitive information and complying with the University of Pavia's regulations for handling research data. The whole procedure was approved by the University's Data Protection and Privacy Officer.

IECoL and the vocabulary test tested very high for reliability on the Cronbach's alpha test, the Kaiser-Mayer-Olkin test and the Bartlett test of sphericity with zero significance (Table 1).

| Cronbach's Alpha <br> coefficient $(\alpha)$ | Kaiser-Mayer-Olkin <br> test (KMO) | $\|\mathrm{R}\|$ | Approx. Chi- <br> Square | df | Significance of <br> Bartlett's test |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 0.933 | 0.965 | $3.86 \mathrm{E}-85$ | 797784.14 | 16471 | 0 |

Table 1
Reliability test results.
At the end of data collection, about 4,000 valid questionnaire-test pairs were gathered. In the sample, the appreciation of the instrument is further suggested by the high response rate to both IECoL and the receptive vocabulary test: Only $6 \%$ of the total sample ( 280 respondents) quit the platform after filling out the questionnaire and without taking the vocabulary test.

## 4. Concluding remarks

Informal learning of English is booming, as evidenced by the growing interest within the field of applied linguistics. In this scenario, methodology issues arise in the search for valid and reliable data collection tools capable of gathering large amounts of information and providing replicable models for further research. Starting from these premises, in the present contribution we have reported on the development and validation of an instrument including a language experience questionnaire (IECOL) and a receptive vocabulary test. The centrality of input for SLA - mainly in terms of frequent, comprehensible and interactionally modified input - is assumed at the core of this research.

We have aimed to create a fine-grained tool which includes a broad variety of media and media genres, distinguishing between receptive and productive informal practices in L2 English and paying attention to the multilingual dimension of present-day L2 use and learning of English as a foreign/second language and as a lingua franca. As location is a paramount dimension of language learning (Benson 2021), we believe that exploring the supports used for accessing input gives access to crucial aspects of the learning process such as the sociability drive as opposed to the private experience often pursued in experiencing the media. Given the personal, self-directed nature of informal learning, attention to the motivational perspective is also fundamental, and was addressed in IECoL through questions on participants' reasons for differentiated media access.

Importantly, IECoL was paired with a receptive vocabulary test, and the choice of a web format allowed for wide-scale administration of both the questionnaire and the vocabulary test. The web-based format of the questionnaire also enabled respondents to complete the task successfully within a reasonable time. This preserved the validity of the
instrument, as it also emerged from the piloting sessions preceding actual administration. Following tests of internal consistency additionally showed the high reliability of the tool.

Some limitations of the data collection tool emerged and will need to be fully addressed in future research. One major drawback of questionnaires in SLA research is their elicitation of reported behaviour - an aspect that differentiates questionnaires from actual observation and that requires participants to rely on their memory of past events (Arndt 2023). Although the issue cannot be fully resolved, we have phrased questions in such a way to reduce the recall effort by restricting and specifying the time span in which the activities are performed. Furthermore, the web format redirects respondents to specific questionnaire items based on their answers, hence not presenting participants with the whole set of questions in case they declared not to access a given input type. This means that, in case of item misunderstanding, participants will miss individually relevant followup questions. Great care must thus be taken in carefully phrasing and piloting questionnaire items. Finally, in addition to gathering reported behavioural data, our tool assesses competence in L2 English through self-evaluation and the vocabulary test. While self-assessment obviously yields a subjective perspective on one's competence in L2 English, the vocabulary test is restricted to receptive lexical knowledge. These limitations in assessing language competence, along with the etic and reported nature of questionnaire data, call for additional investigations from a longitudinal and emic perspective. Triangulation can additionally help increase the validity and reliability of results, as rich data directly coming from participants who voluntarily share their experience and express their point of view can make findings more solid overall.

Acknowledgements: This work has been supported by the PRIN 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 (MUR) - Bando 2020 grant 2020NNJTW3_001.
Our deepest gratitude goes to all the PRIN colleagues, students, friends and family who participated in the different validation and piloting phases of the project and offered their support. We are also thankful to Flavio Ceravolo for his insightful suggestions about questionnaire design. We would like to thank Marco Zappatore, who provided precious technological support in the initial stages of the project. Last but not least, we wish to warmly thank all the colleagues who welcomed us into their classrooms and all the students who participated in the study. The project would not have been possible without them.

## Bionotes:

Maria Pavesi is Full Professor of English language, Linguistics and Translation at the Department of Humanities, University of Pavia. Her research has addressed various topics in English applied linguistics focussing on English second language acquisition, film language and audiovisual translation. She has published very widely and recently co-authored Informal Contact with English. A Case Study of Italian Postgraduate Students (Edizioni ETS, 2021) and authored "Corpora and the language of films: Exploring dialogue in English and Italian" (Routledge 2022). Maria Pavesi has taken part in several governmentfunded, EU and international research projects. Currently, she is the PI of the PRIN project "The informalisation of English language learning through the media".
Elisa Ghia, PhD in Linguistics, is an Assistant Professor of English at the University of Pavia. She has been visiting fellow at the University of Turku, Katholieke Universiteit Leuven, the University of Melbourne and Michigan State University and she has taught English Linguistics and ESP at the University for Foreigners of Siena, the University of Eastern Piedmont and the University of Milan. Her research interests include foreign language acquisition, audiovisual translation, corpus linguistics and the study of spoken English. With Maria Pavesi, she is co-author of the book Informal Contact with English. A case study of Italian postgraduate students (2020).
Tiziana Ciabattoni works in Business Intelligence as Data Architect, Data Analyst and Project Manager at the University of Pavia. She graduated in Physics at the University of Pavia with a theoretical-experimental
thesis in Condensed Matter Structure, with which she won the "R. Sala" Study Prize. At the same University Tiziana Ciabattoni obtained a Master's in Materials Science and Technology and a PhD in Physics. She collaborates with numerous research groups in physics and linguistics automating measurement procedures, test administration and data analysis and creating software for various laboratory instruments and online IT platforms for data collection and analysis.

Author's addresses: mariagabriella.pavesi@unipv.it; elisa.ghia@unipv.it; tiziana.ciabattoni@unipv.it

## References

Aiello J. 2018, Negotiating Englishes and English-speaking identities, Routledge, London.
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. 2023, Construction and validation of a questionnaire to study engagement in informal second language learning, in "Studies in Second Language Acquisition", pp. 1-25.
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. 2011, Language learning and teaching beyond the classroom: An introduction to the field, in Benson P. and Reinders H. (eds.), Beyond the Language Classroom, Palgrave Macmillan, Basingstoke, pp. 7-16.
Benson P. 2021, Language Learning Environments, Multilingual Matters, Bristol.
Berns M., de Bot K. and Hasebrink U. 2007, In the Presence of English: Media and European Youth, Springer, New York.
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, Università degli Studi di Pavia.
Collentine J. and Freed B. F. 2004, Learning context and its effects on Second Language Acquisition, in "Studies in Second Language Acquisition" 26 [2], pp. 153-171.
Cravidi G. 2016, Accessing English Input Outside University: A Qualitative and Quantitative Investigation, Unpublished postgraduate thesis, Università degli Studi di Pavia.
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.
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, WileyBlackwell, Hoboken.
Ellis N. C. 2003, Constructions, Chunking and Connectionism, in Doughty C. and Long M. (eds.), The Handbook of Second Language Acquisition, Blackwell Publishing, Malden, pp. 63-103.
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.
Freed B. F., Dewey D. P., Segalowitz N. and Halter R. 2004, The language contact profile, in "Studies in Second Language Acquisition" 26 [2], pp. 349-356.
Gass S. M. 1997, Input, Interaction, and the Second Language Learner, Lawrence Erlbaum Associates, Mahwah.
Gass S. M. and Mackey A. 2020, Input, Interaction, and Output in L2 Acquisition, in VanPatten B. and Williams J. (eds.), Theories in SLA. An Introduction, Routledge, London, pp. 192-222.
Ghia E. and Pavesi M. 2021, Choosing between dubbing and subtitling in a changing landscape, in "Lingue e Linguaggi" 46, pp. 161-177.
González Fernandez B. and Schmitt N. 2015, How much collocation knowledge do L2 learners have?: the effects of frequency and amount of exposure, in "ITL - International Journal of Applied Linguistics" 166 [1], pp. 94-126.
Iwaniec J. 2019, Questionnaires: Implications for effective implementation, in McKinley J., Rose H. (eds.), The Routledge handbook of research methods in applied linguistics, Routledge, London, pp. 324-335.
Krashen S. 1985, The Input Hypothesis: Issues and Implications, Longman, London.
Kukulska-Hulme A. 2009, Will mobile learning change language learning?, in "ReCALL" 21 [2], pp. 157165.

Kuppens A. 2010, Incidental foreign language acquisition from media exposure, in "Learning, Media and Technology" 35 [1], pp. 65-85.
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. and Sockett G. 2012, From informal resource usage to incidental language acquisition: Language
uptake from online television viewing in English, in "Asp. La revue du GERAS" 62, pp. 45-65.
Laufer B., Elder C., Hill K. and Congdon P. 2004, Size and Strength: Do We Need Both to Measure Vocabulary Knowledge?, in "Language Testing" 21, pp. 202-226.
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.
Leppänen S., Pitkänen-Huhta A., Nikula T., Kytölä S., Törmäkangas T., Nissinen K., Kääntä L., Räisänen T., Laitinen M., Pahta P., Koskela H., Lähdesmäki S. and Jousmäki H. 2011, National survey on the English language in Finland: Uses, meanings and attitudes, in "Studies in Variation, Contacts and Change in English" 5.
Lindgren E. and Muñoz C. 2013, The influence of exposure, parents, and lin- guistic distance on young European learners' foreign language comprehension, in "International Journal of Multilingualism" 10 [1], pp. 105-129.
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.
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.
Lütge C. (ed.) 2022, Foreign Language Learning in the Digital Age, Routledge, Oxon.
Martinez R. and Schmitt N. 2012, A Phrasal Expressions List in "Applied Linguistics" 33, pp. 299-320.
McCarthy T.M. 2009, Music and song beyond the classroom: Strategies to aid the language learning process, in "Cam TESOL Conference on English Language Teaching: Selected Papers" 5, pp. 252261.

Meara P. 1992, EFL Vocabulary Tests, Swansea University, Swansea.
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.
Nation I.S.P. 1990, Teaching and Learning Vocabulary, Newbury House, New York.
Nation I. S. P. 2021, Critical Commentary. Thoughts on word families, in "Studies in Second Language Acquisition" 43, pp. 1-4.
Ortega L. 2019, SLA and the Study of Equitable Multilingualism, in "The Modern Language Journal" 103 [1], pp. 23-38.
Pavesi M. and Ghia E. 2020, Informal Contact with English. A case study of Italian postgraduate students, ETS, Pisa.
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.
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.
Schmitt N., Nation P. and Kremmel B. 2020, Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation, in "Language Teaching" 53 [1], pp. 109-120.
Sockett G. 2014, The Online Informal Learning of English, Palgrave Macmillan, London.
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.
Sundqvist P. and Sylvén L.K. 2016, Extramural English in Teaching and Learning: from Theory and Research to Practice, Palgrave Macmillan, London.
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.
Swain M. 2005, The output hypothesis: theory and research, in Hinkel E. (ed.), Handbook of Research in Second Language Teaching and Learning, Lawrence Erlbaum Associates, Mahwah, pp. 471-483.
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.
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. and Sockett G. 2014, English language music: Does it help with learning?, in "Researching and Teaching Languages for Specific Purposes" 33 [2], pp. 192-209.
VanPatten B. and Williams J. (eds.) 2020, Theories in SLA. An Introduction, Routledge, London.

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.
Webb S., Sasao Y. and Ballance O. 2017, The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT, in "ITL - International Journal of Applied Linguistics" 168 [1], pp. 34-70.

## Appendix

The Informal English Contact and Learning questionnaire - IECoL
Abridged and compacted version
Questionario sul contatto con la lingua inglese tramite i media

* Required

1. Codice identificativo *

Questo è il tuo codice identificativo. Accertati che sia presente e clicca su Avanti per avviare il Questionario.
Background linguistico
2. Numero di lingue conosciute oltre alla lingua madre (almeno a livello * elementare, incluso il dialetto)
[only one option possible]
$1 / 2 / 3 / 4$ / più di 4
3. Quale lingua si
parla nel tuo
paese (la tua
lingua madre)? *
[one or more
answers are
possible]
Italiano /
Other: $\qquad$
4. Quanti anni avevi quando hai iniziato a studiare la lingua inglese a scuola? *
[only one option possible]
5/6/7/8/9/10/11/12/13/14/Dopoi14 anni
5. Hai mai seguito o segui corsi extrascolastici di lingua inglese? *
[only one option possible]
Sì / No
Modalità con cui accedi alla lingua inglese al di fuori dei corsi di lingua inglese
6. Guardi film, serie TV o altri programmi in lingua inglese? * [only one option possible]
Sì [respondent redirected to question 7]
No [respondent redirected to question 15]
7. 1.1. Quanto spesso guardi in lingua inglese...?
[only one option per row possible]

|  | Molto <br> spesso <br> (tutti i <br> giorni <br> 0 <br> quasi) | Spesso (2 <br> 03 volte <br> la <br> settimana) | Qualche <br> volta (1 <br> volta alla <br> settimana) | Raramente <br> $(102$ <br> volte al <br> mese) | Mai |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | $\square$ |  |  |  |  |
| Film |  |  |  |  |  |
| Serie TV |  |  |  |  |  |
| Altri <br> programmi <br> (documentari, <br> talk show, <br> news, ecc.) |  |  |  |  |  |

8. 1.2. Per quanto tempo guardi film, serie TV e/o altri programmi in lingua inglese complessivamente il giorno in cui lo fai?
[only one option per row possible]

|  | Per più̀ <br> di due <br> ore | Da una <br> a due <br> ore | Per <br> circa <br> un'ora | Da 30 <br> minuti <br> a <br> un'ora | Per <br> meno <br> di 30 <br> minuti |
| :--- | :---: | :---: | :---: | :---: | :---: | | Mai |
| :--- |
| Guardo film |
| Guardo serie |
| TV |

9. 1.3. Se guardi film e serie TV in lingua inglese, li preferisci: [only one option possible]

In lingua originale senza sottotitoli / Con sottotitoli in inglese / Con sottotitoli in italiano
10. 1.4. Se guardi altri programmi (documentari, talk show, news, ecc.) in lingua inglese, li preferisci: [only one option possible]
In lingua originale senza sottotitoli / Con sottotitoli in inglese / Con sottotitoli in italiano
11. 1.5. A parte film e serie TV, quali altri generi audiovisivi/ televisivi ti piace guardare in lingua inglese? [One or more answers are possible]

| News | Sport | Talk show | Reality show | Concerti |
| :--- | :--- | :--- | :--- | :--- |
| Varie | Documentari | Programmi di cucina | Stand-up comedy | Other: |

12. 1.6. Quale supporto utilizzi in genere? È possibile indicare più di una opzione.

|  | Televisore | Computer | Tablet | Smartphone |
| :--- | :--- | :--- | :--- | :--- |
| Per film e <br> serie TV | $\square$ | $\square$ | $\square$ | $\square$ |
| Per altri <br> programmi <br> (documentari, <br> talk show, <br> news, ecc.) | $\square$ | $\square$ | $\square$ | $\square$ |

13. 1.7. Per quali ragioni guardi programmi in lingua inglese originale? È possibile indicare più di una opzione.

14. 1.7.1. Se hai scelto Altro, puoi specificare:
15. 2. Guardi YouTube in lingua inglese? *
[only one option possible]
Sì [respondent redirected to question 16]
No [respondent redirected to question 22]
1. 2.1. Quanto spesso guardi YouTube in lingua inglese?
[only one option possible]
Molto spesso (tutti i giorni o quasi) / Spesso (2 o 3 volte la settimana) / Qualche volta (1 volta alla settimana) / Raramente (1 o 2 volte al mese)
2. 2.2. Per quanto tempo guardi YouTube in lingua inglese complessivamente il giorno in cui lo fai?
[only one option possible]
Più di due ore / Da una a due ore / Circa un'ora / Da 30 minuti a un'ora / Meno di 30 minuti
3. 2.3. Se guardi video su YouTube in lingua inglese, li preferisci:
[only one option possible]
In lingua originale senza sottotitoli / Con sottotitoli in inglese / Con sottotitoli in italiano
4. 2.4. Quali tipi di video guardi su YouTube in inglese? È possibile indicare più di una opzione.

Video musicali Tutorial Recensioni Scene di film e serie TV Sport
$\begin{array}{lll}\text { Video comici } & \text { Documentari } & \text { Ricette di cucina Talk show }\end{array}$ Gameplay
News
YouTuber
Other: $\qquad$
20. 2.5. Se guardi video su YouTube in inglese, quale supporto utilizzi?

É possibile indicare più di una opzione.
Televisore / Computer / Tablet / Smartphone
21. 2.6. Per quali ragioni guardi video su YouTube in lingua inglese?

É possibile indicare più di una opzione.
Per svago/intrattenimento
Per studio
Per migliorare l'inglese
Per parlarne con altre persone
Per accedere a informazioni
Per imparare nuove parole ed espressioni di uso corrente e slang
Perché mi piace la lingua inglese
Other: $\qquad$
22. 3. Giochi a videogames in lingua inglese? *
[only one option possible]
Sì [respondent redirected to question 23]
No [respondent redirected to question 29]
23 3.1. Quanto spesso giochi a videogames in inglese...?
[only one option possible]


24 3.2. Per quanto tempo giochi a videogames in inglese da solo/a e/o in multiplayer complessivamente il giorno in cui lo fai?
[only one option possible]

25. 3.3. Quali categorie di giochi usi in genere in lingua inglese? È possibile indicare più di una opzione.

Avventure grafiche
Visual novel e/o interactive fiction
Action-Adventure (per es., Stealth, Survival horror, ecc.)
Azione / Picchiaduro / Sparatutto
Immersive sim (per es. city-building, artificial life, sandbox, ecc.)
Giochi di ruolo (per es. RPG, MMORPG, Open world)
Simulatori (per es. di volo; di guida; sportivi, ecc.)
Strategia (real-time, turn-based, ecc.)
Puzzle games
Giochi di tipo educational
Other: $\qquad$
$\mathcal{L o}_{\text {inguaggi }}$
26. 3.4. Se giochi a videogames in lingua inglese, li preferisci:
[only one option possible]
In lingua originale senza sottotitoli / Con sottotitoli in inglese / Con sottotitoli in italiano
27. 3.5. Se giochi a videogames in inglese, quale supporto utilizzi in genere?

È possibile indicare più di una opzione.

3.6. Per quali ragioni giochi a videogames in lingua inglese?

È possibile indicare più di una opzione.
Per svago/intrattenimento
Per migliorare l'inglese
Per stare insieme ad altre persone
Perché mi piace la competizione
Per apprendere contenuti
Perché mi piace la lingua inglese
Other:
29. 4. Ascolti canzoni in lingua inglese? *
[only one option possible]
Sì [respondent redirected to question 30]
No [respondent redirected to question 34]
30. 4.1. Quando ascolti canzoni in lingua inglese, ti capita di concentrarti sul testo * delle canzoni? [only one option possible]
Sì [respondent redirected to question 31]
No [respondent redirected to question 34]
Modalità con cui accedi alla lingua inglese al di fuori dei corsi di lingua inglese (parte 4.c)
31. 4.2. Quanto spesso ti concentri sul testo delle canzoni che ascolti in lingua inglese?
[only one option possible]
Molto spesso / Spesso / Qualche volta / Raramente
32. 4.3. Per quanto tempo ti concentri sul testo inglese delle canzoni complessivamente il giorno in cui lo fai?
[only one option possible]
Più di due ore / Da una a due ore / Circa un'ora / Da 30 minuti a un'ora / Meno di 30 minuti
33. 4.4. Accedi al testo delle canzoni in lingua inglese tramite programmi di streaming audio o
altre risorse sul web?
[only one option possible]
Sì / No
34. 5. Usi Internet in lingua inglese? *
[only one option possible]
Sì [respondent redirected to question 35]
No [respondent redirected to question 44]
35. 5.1. Quanto spesso svolgi le seguenti attività in lingua inglese?
[only one option per row possible]

36. 5.2. Per quanto tempo usi internet in lingua inglese per le seguenti attività complessivamente il giorno in cui lo fai?
[only one option per row possible]

37. 5.3. Se e quando accedi ai social network, qual è la percentuale approssimativa di contenuti in inglese?
[only one option possible] $100 \%$ / $75 \%$ / $50 \% / 25 \% / 0 \%$
38. 5.4. Se accedi a social network in inglese, quali usi?

É possibile indicare più di una opzione.

$$
\mathcal{T}_{9 \text { ingue e }}
$$

| Facebook Instagram TikTok | Twitter | Pinterest |  |  |
| :--- | :--- | :--- | :--- | :--- |
| Tumblr | Other: |  |  |  |

39. 5.5. Se accedi a pagine web in inglese, a quali accedi?

È possibile indicare più di una opzione.

| Wikipedia | Altri wiki | Dizionari di inglese | Siti di notizie e attualità |
| :--- | :--- | :--- | :--- |
| Hobby e cucina | Other: |  |  |

40. 5.6. Se accedi a blog e forum in inglese, a quali accedi?

E possibile indicare più di una opzione.

| Gaming | Musica | Viaggi | Estetica e moda | Tecnologia |
| :--- | :--- | :--- | :--- | :--- |
| Cucina | Libri | Grammatica e uso dell'inglese | Cinema |  |
| Auto/moto | Sport | Other: |  |  |

41. 5.7. Su internet interagisci in inglese:
[only one option per row possible]
Con parlanti nativi di inglese / Con parlanti non nativi di inglese
42. 5.8. Per quali ragioni svolgi le seguenti attività su internet in lingua inglese?

È possibile indicare più di una opzione.

43. 5.8.1. Se hai scelto Altro, puoi specificare:
44. 6. Leggi libri, riviste o quotidiani in lingua inglese? *
[only one option possible]
Sì [respondent redirected to question 45]
No [respondent redirected to question 49]
45. 6.1. Quanto spesso leggi libri, riviste o quotidiani in lingua inglese?
[only one option per row possible]

46. 6.2. Per quanto tempo leggi libri, riviste o quotidiani in lingua inglese complessivamente il giorno in cui lo fai?
[only one option per row possible]

|  | Più di <br> due <br> ore | Da una <br> a due <br> ore | Circa <br> un'ora | Da 30 <br> minuti <br> a <br> un'ora | Meno <br> di30 <br> minuti |
| :--- | :---: | :---: | :---: | :---: | :---: | | Mai |
| :--- |
| Leggo <br> libri |
| Leggo <br> riviste o <br> quotidiani |

47. 6.3. Per quali ragioni leggi libri, riviste o giornali in lingua inglese?

È possibile indicare più di una opzione.

48. 6.3.1. Se hai scelto Altro, puoi specificare:
49. 7. Svolgi le seguenti attività usando la lingua inglese?
[only one option per row possible]

50. 7.1. Per quanto tempo svolgi le seguenti attività in lingua inglese complessivamente il giorno in cui lo fai?
[only one option per row possible]

$$
\mathcal{T}_{0 \text { ingue e }}
$$


51. 7.2. Usi l'inglese viaggiando?
[only one option possible]
Molto spesso / Qualche volta / Raramente o mai
52. 8. Segui corsi nelle tue discipline in lingua inglese (ad esclusione dei corsi di * lingua inglese)?
[only one option possible]
Sì / No
53. 9. Quanti anni avevi quando hai iniziato ad accedere alla lingua inglese nel tuo tempo libero (per esempio, ascoltando i testi di canzoni o guardando video in inglese)?
[only one option possible]
Meno di $3 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 15 / 16 / 17 / 18 / 19 / 20 /$ Più di 20
54. 10. Quanto importante per te conoscere l'inglese su una scala da 1 a 10 ? *
[only one option possible]
55. 11. Quanto ti piace la lingua inglese da 1 a 10 ? *
[only one option possible]
Accesso ad altre lingue straniere al di fuori dei corsi di lingua
56. 12. Nel tuo tempo libero, accedi ad altre lingue straniere? *
[only one option possible]
Sì [respondent redirected to question 57]
No [respondent redirected to question 59]
57. 12.1. A quali altre lingue straniere accedi nel tuo tempo libero?
58. 12.2. In quali attività accedi alle altre lingue straniere?
59. 13. Hai vissuto per un periodo all'estero in un paese in cui dovevi usare l'inglese per comunicare (con un programma di mobilità o per altre ragioni)? *
[only one option possible]
Sì [respondent redirected to question 60]
No [respondent redirected to question 61]
60. 13.1. Se hai vissuto per un periodo all'estero in un paese in cui dovevi usare l'inglese per comunicare, in quali paesi e per quanto tempo?
61. 14. Ritieni di conoscere l'inglese:
[only one option possible]
A un livello elementare - A2 (corrispondente a certificazione Key English Test)
A un livello intermedio - B1 (corrispondente a certificazione PET)
A un livello intermedio-superiore - B2 (corrispondente a certificazione FCE - First Certificate in English)
A un livello avanzato - C1 (corrispondente a certificazione CAE - Certificate in Advanced English)
A un livello nativo o quasi nativo - C2 (corrispondente a certificazione CPE - Certificate of
Proficiency in English)

Informazioni anagrafiche
62. Età *

1. Indicare l'età scegliendo dal menu a tendina:
[only one option possible]
$19 / 20 / 21 / 22 / 23 / 24 / 25 / 26 / 27 / 28 / 29 / 30 /$ più di 30
2. 2. Corso di Laurea: *
[only one option possible]
Triennale / Magistrale / A ciclo unico
1. 3. Nome del corso di laurea *
1. 4. Anno di corso *
[only one option possible]
Primo / Secondo / Terzo / Altro
1. 5. Provincia di residenza *
[options from a drop-down menu; only one option possible]
1. 6. Genere *
[only one option possible]
F / M / Preferisco non rispondere
1. 7. Scuola superiore frequentata *
[only one option possible]
Istituto Tecnico / Istituto Professionale / Liceo classico / Liceo scientifico / Liceo linguistico / Altro liceo
1. 8. Titolo di studio dei genitori *
[only one option possible]

|  | Licenza <br> media | Diploma <br> di scuola <br> superiore | Laurea | Dottorato/Specializz. <br> post-laurea |
| :--- | :---: | :---: | :---: | :---: |
| Madre |  |  |  |  |
| Padre |  |  |  |  |

70. 9. Occupazione della madre
[only one option possible]
Imprenditrice e alta dirigenza
Professioni intellettuali, scientifiche e di elevata specializzazione (per esempio docente, medico, farmacista, ingegnere, avvocato, giornalista, specialista)
Professioni tecniche (per esempio tecnica informatica, fisioterapista, pilota, infermiera)
Professioni impiegatizie (per esempio impiegata, contabile, centralinista)
Attività commerciali e nei servizi (per esempio commessa, esercente commerciale, ristoratrice, assistente di volo, cura della persona, assistente alla persona, vigile urbano, Polizia di Stato)
Professione di operai specializzati, agricoltori e artigiani
Attività di conduttori di impianti, operai di macchinari e conducenti di veicoli (per esempio autista, conduttrice e addetta a macchinari, marinaio)
Occupazione non specializzata (per esempio addetta alle consegne di merci, addetta alla pulizia, addetta alla custodia di locali)
Casalinga
1. 10. Occupazione del padre
[only one option possible]
Imprenditore e alta dirigenza
Professioni intellettuali, scientifiche e di elevata specializzazione (per esempio docente, medico, farmacista, ingegnere, avvocato, giornalista, specialista)
Professioni tecniche (per esempio tecnico informatico, fisioterapista, pilota, infermiere)
Professioni impiegatizie (per esempio impiegato, contabile, centralinista)
Attività commerciali e nei servizi (per esempio commesso, esercente commerciale, ristoratore, assistente di volo, cura della persona, assistente alla persona, vigile urbano, Polizia di Stato)
Professione di operai specializzati, agricoltori e artigiani
Attività di conduttori di impianti, operai di macchinari e conducenti di veicoli (per esempio autista, conduttore e addetto a macchinari, marinaio)
Occupazione non specializzata (per esempio addetto alle consegne di merci, addetto alla pulizia, addetto alla custodia di locali)

1. 11. Hai mai avuto una certificazione di BES, DSA o di altre condizioni che possono influire sull'apprendimento linguistico?
[only one option possible]
Sì / No

[^0]:    ${ }^{1}$ The article is the result of joint work by the authors. Maria Pavesi and Elisa Ghia designed IECoL and adapted the vocabulary test, while Tiziana Ciabattoni developed the online IT platform. Maria Pavesi wrote sections 1., 2., 3.1.1., 3.1.2 and 3.1.3; Elisa Ghia wrote sections 2.1., 3.1.4 and 3.2.; they jointly wrote section 4.; Tiziana Ciabattoni wrote section 3.3.

[^1]:    ${ }^{2}$ When questionnaires are too long and repetitive, respondents may skip items or provide superficial or imprecise answers.

