International Journal of English Linguistics; Vol. 16, No. 3; 2026
ISSN 1923-869X E-ISSN 1923-8703
Published by Canadian Center of Science and Education

Vagueness in Learner English: Exploring General Extenders and Informal Learning among Italian EFL Students

Filippo Saettoni[1]

[1] Department of Philology, Literature and Linguistics, University of Pisa, Pisa, Italy

Correspondence: Filippo Saettoni, Department of Philology, Literature and Linguistics, University of Pisa, Via Santa Maria 67, 56126 Pisa, Italy. E-mail: filippo.saettoni@phd.unipi.it; filsae@gmail.com

Received: February 2, 2026 Accepted: April 5, 2026 Online Published: April 20, 2026

doi:10.5539/ijel.v16n3p88 URL: https://doi.org/10.5539/ijel.v16n3p88

Abstract

This study investigates the phenomenon of vague language within the framework of informal English language learning among Italian EFL learners. While recent societal and technological developments have been affecting the boundaries between formal and informal learning, little attention has been paid to how Italian learners, through informal exposure, first acquire and subsequently use pragmatic features (e.g., vague expressions). The research adopts a mixed-methods approach and combines corpus-based analysis from a learner corpus (the Trinity Lancaster Corpus) with data from a student questionnaire to investigate how learners employ general extenders (e.g., *and so on*, *or something like that*) and to what extent their usage reflects informal acquisition rather than explicit classroom instruction. Findings indicate that Italian EFL learners are generally aware of vague language but tend to use it in limited ways, overlooking its broader interpersonal and relational functions. Such limitations seem to be connected to the absence of explicit instruction on vague language in institutional educational settings, despite learners' frequent exposure to such expressions through English-language media and social platforms.

Keywords: informal language learning, Italian EFL learners, digital media, vague language, general extenders

1. Introduction

In their special issue on the informalisation of English language learning, Pavesi and Bianchi (2024) highlight the limited number of studies on this topic within the Italian academic context – particularly when compared to other European and Asian countries that have investigated this phenomenon on a larger scale. This shift is particularly notable if one considers the recent societal changes that have been affecting the learning experiences of many EFL students, both internationally and nationally.

Specifically, factors such as multilingualism, multiculturalism, technological advancements (e.g., the Internet, satellite communication, and wireless technologies), and "an unprecedented mobility of people and linguistic artefacts" (Pavesi & Bianchi, 2024, p. 7) are contributing to a growing process of informalisation in language learning. This means that, while traditionally it has been associated with the classroom environment (Pavesi, 2025), language learning is now increasingly evolving into an informal experience (Wagner, 2015), allowing students to engage with a wide range of linguistic input beyond formal educational settings – especially through English-language media. As Pavesi (2024, p. 24) illustrates, "[u]sing 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".

Among the scholars who have examined how learners engage with English beyond formal educational settings, both international and national researchers can be identified. International research consistently highlights the central role of digital environments, together with individual differences – such as autonomy, engagement, and motivation – in shaping informal English learning experiences across contexts (Dressman & Sadler, 2020; Toffoli, Sockett, & Kusyk, 2023). In line with these findings, Cole and Vanderplank (2016) show that fully autonomous, Internet-mediated learners are likely to outperform classroom-trained students, suggesting that informal learning and learner motivation are crucial factors in the acquisition of high-level English.

Among Italian EFL learners, informal exposure to English has been extensively investigated in the PRIN project (Pavesi & Bianchi, 2024; Pavesi et al., 2025), which highlights patterns of media use, learner preferences, and engagement. University students frequently access TV series and YouTube content (Ghia, 2024), with TV series

being the most popular genre and YouTube used especially for tutorials, music videos, interviews, and film/TV scenes (Pavesi et al., 2025). Learners typically prefer intralingual subtitles (English audio with English captions; Ghia, 2024) and engage mainly receptively, spending more time reading and watching than producing language (Ghiselli, 2024). Music is another widely accessed input source, though few students engage extensively with lyrics, while videogames are less popular. In addition, social media platforms such as Instagram, TikTok, and Pinterest have become key spaces for English exposure, offering micro-lessons, vocabulary content, and grammar tips within learners' everyday feeds (Cappelli & Simi, 2024; Manca, 2024). Learners now favour screen-mediated interactions over face-to-face communication, preferring brief exchanges such as instant messaging rather than extended conversations (Pavesi et al., 2025). Overall, these findings indicate that Italian learners' informal English exposure is largely media-driven, predominantly receptive, and increasingly shaped by social and digital platforms.

In light of the above – and given the scarcity of research on informal language learning within the Italian educational context – this research seeks to bridge this gap by examining a pragmatically relevant phenomenon that remains largely overlooked in language education: vague language (a brief theoretical background on this topic will be provided in the following Section). This linguistic feature plays a crucial role in the everyday speech of native speakers and helps them engage in successful conversations, yet it is seldom considered in Italian EFL formal instruction, with the consequence that learners are not fully aware of its communicative potential and thus risk to produce language which may not be as effective as that of native speakers (Samigoullina, 2020). Given the limited attention to vague language, this study adopts a mixed-methods approach (Section 2) and investigates how learners use vague expressions – in particular general extenders (Sections 3.1 and 3.2) – and how these have been acquired through informal exposure rather than formal classroom instruction (Section 3.3). Finally, Section 4 discusses the main findings and pedagogical implications.

1.1 Vague Language and Native Speakers

Vague language is a linguistic phenomenon that has long attracted the attention of scholars. When exploring the literature on the topic, two distinct research paths emerge: one philosophical and the other linguistic. The first path has been pursued by philosophers, thinkers, and logicians who, each in their own way, have reflected on the vague nature of words and expressions. Put simply, language is not set in stone and the meanings of the words at our disposal are not always easily definable, as we lack specific boundaries and criteria to express matters with absolute certainty (Machetti, 2006). Consider the bald man paradox for instance, where *bald* is a vague predicate – no one can state with absolute certainty when an individual qualifies as bald, since baldness lacks a clear-cut threshold separating the bald from the non-bald (Machetti, 2006). This represents a perfect example of *systemic vagueness* (Voghera, 2012; Voghera & Collu, 2017).

Conversely, following Peirce's work (1902), scholars – and linguists in particular – have focused on *intentional vagueness*, analysing it from a specifically linguistic perspective (Machetti, 2006). As its name suggests, this particular type of vagueness is related to the speakers' intentions rather than to the inherently undefined nature of words and expressions they can draw on. In this regard, during conversation, speakers may choose to use certain linguistic devices – which can be regarded as vague – for various reasons; these different uses have been categorised into three main types of intentional vagueness, namely *informational*, *discourse-related*, and *relational* vagueness (Voghera & Collu, 2017).

In the first case, speakers may lack specific information about the topic under discussion or may not know the precise words needed in a given context; consequently, they will likely tend to use vague expressions to manage these situations. In the second case, speakers may experience difficulty during spontaneous conversation due to the real-time pressure of formulating sentences; under this perspective, they may resort to vague expressions to maintain the conversational flow. Finally, in the third case, speakers might possess all the necessary information but choose not to be too direct in their assertions. Hence, to soften their claims, they may deliberately employ vague language as a strategic communicative choice.

Among the many expressions speakers can employ to achieve these communicative goals, one can find (i) placeholder nouns (e.g., *stuff*, *things*, *thingie*); (ii) approximators (*about*, *around*, *or so*); (iii) round numbers; (iv) non-numerical quantifiers (e.g., *bags of*, *a bit of*, *several*, *often*); (v) vague adverbs of likelihood (e.g., *probably*, *maybe*, *perhaps*); (vi) hedges (e.g., *sort of*, *like*, *kind of*); (vii) discourse markers (e.g., *I mean*, *you know*); (viii) modal verbs; (ix) copular verbs; and (x) other markers of vagueness (e.g., *if* clauses, intonation, and indefinite pronouns; Cutting, 2007).

In addition to these vague devices, native speakers also tend to use general extenders (Channell, 1994). These are fixed or semi-fixed expressions whose main pragmatic function is to evoke a broader set of items related to the topic being discussed. Morphologically, general extenders are easily recognizable because they typically follow a

prototypical structural pattern, which may vary depending on the speakers' lexical choices (Tagliamonte & Denis, 2010). In most cases, they are introduced by a conjunction – those beginning with *and* are known as *adjunctive* general extenders, whereas those beginning with *or* are referred to as *disjunctive* forms. The conjunction may be followed by a quantifier, a necessary generic noun (which may also take the form of an indefinite pronoun), and possibly a comparative element (e.g., *like this*) – whose presence determines whether the general extender is classified as a short or long form. In addition to the most prototypical forms, other fixed expressions (e.g., *and so on*, *et cetera*) can be considered general extenders as well.

The various types and devices described above show that vague items are an essential and valuable feature of native speakers' speech; even so, vague language has long been regarded as unworthy of scholarly attention (Cutting, 2007; Samigoullina, 2020). For many years, scholars have considered it a reflection of poor linguistic or cognitive abilities, and thus as a marker of unsuccessful communication. Only in recent years, have linguists begun to recognise its value in conversation from a variety of perspectives. In this sense, an appropriate use of vague expressions can contribute to successful conversations by helping speakers (i) navigate lexical or informational gaps, (ii) cope with the demands and pressures of real-time speech production, and (iii) build closer and stronger connections with their interlocutors through shared knowledge.

Building on these perspectives, scholars have also begun to highlight the pragmatic strengths of vague language in the speech of EFL learners, who may not be fully aware of its communicative potential. In this regard, many have argued that greater attention should be devoted to vague language in order to (i) enrich learners' linguistic repertoire (Cutting, 2007), (ii) help them compensate for inadequacies in their communicative competence (Sabet & Zhang, 2015), and (iii) enable them to participate in more successful, native-like interactions.

2. Research Questions and Methodology

The aim of the present research is two-fold. First, to investigate whether Italian students know – and how they pragmatically use – vague expressions, specifically general extenders; second, to ascertain how they acquired such linguistic items during their educational journey. In order to address these goals, the study was conducted along two lines of investigation: a corpus-based inquiry and the analysis of a questionnaire.

Regarding the first goal, the Trinity Lancaster Corpus – henceforth TLC – was chosen as a primary source of data (Note 1). Compiled through a collaboration between the Centre for Corpus Approaches to Social Science (CASS) at Lancaster University and Trinity College London (i.e., a major international testing board), the TLC is the largest corpus of L2 spoken English (Gablasova, Brezina, & McEnery, 2019). Specifically, it contains 4.2 million words of interaction between L1 and L2 speakers of English; however, the most pivotal element is the inclusion of L2 speakers from diverse linguistic and cultural backgrounds, such as Italy, Spain, Mexico, Argentina, Brazil, China, India, Sri Lanka, and Russia. The data comes from recordings of Trinity's Graded Examinations in Spoken English (GESE) across grades from B1–C2 on the CEFR scale, thus offering unique insights into how EFL learners manage interaction and build meaning based on their own identity.

The analysis was made possible thanks to the Trinity Lancaster Corpus Hub – or TLC Hub – (Gilquin & Granger, 2022), a user-friendly interface that allows researchers to access a comparable subset of the TLC, i.e., the Trinity Lancaster Corpus – Conversation and Discussion (TLC-C&D), which consists of about 2,000 learners of English. To investigate the corpus, users enter a keyword or phrase into the search bar to obtain concordance lines; these results can then be sorted according to specific sociolinguistic variables, including gender, age, proficiency level, and country of origin. The focus of the present research, however, is on the Italian subcomponent, which is made of 343 learners of English, divided according to the CEFR scale into 145 at B1 (42%), 130 at B2 (38%), 58 at C1 (17%), and 10 at C2 (3%) (Lancaster University et al., n.d.).

The present analysis was thus carried out by using the following procedure. First, a reference list of all possible general extenders was compiled, drawing on the main literature on the topic (Channell, 1994; Overstreet, 1999; Cutting, 2007). Then, the linguistic items were inserted into the search bar, concordance lines were downloaded and manually filtered to make sure all retrieved instances were actual uses of vague language. For instance, when exploring the vague expressions containing *something*, exchanges like (1) were suitable for the analysis, while others – such as (2) – were discarded. In this sense, while in the former example, *something* functions as a general extender and signals an open-ended category of possibilities, in the latter, the word is used as an indefinite pronoun without serving an extending function and was therefore not considered relevant to the analysis.

(1)

S: it becomes a nightmare because er er I us= I used to dream to have nightmares er

E: mm

S: during my childhood

E: mm

S: being er without erm without dreams actually being erm alone without

E: mm

S: er a job *or something like that* which is quite strange for a child

E: so are you saying that anxiety causes us to have nightmares because they're a reflection of our worries in real life?

S: yeah

(2)

S: life but they can also talk about er thing *something* that never happened their their only way is to represent

Once the filtering phase was completed, all retrieved sentences were manually analysed through close reading, in order to capture (i) the frequency of the preferred general extenders (as will be discussed in Section 3.1) and (ii) the type of intentional vagueness (i.e., informational, discourse, or relational) that each vague expression was thought to convey (Section 3.2). To this end, all general extenders were first inspected manually. Each instance was then classified according to two criteria. First, the vague expressions were distinguished based on whether they were used to refer to facts in an objective manner or to express opinions in a subjective way. In the former case, the instance was further categorised as either informational, discourse-related, or both, depending on the pragmatic function it appeared to serve; in the latter case, the instance was classified as relational. Finally, instances that performed more than one pragmatic function – and could be categorised in more than one type of vagueness – were noted as multifunctional.

As for the second research question, data was obtained thanks to a self-designed questionnaire administered to a group of 209 Italian students. The questionnaire, which was part of a related project on the analysis of vague language, was designed to assess the students' knowledge – and awareness – of specific vague expressions (i.e., placeholder nouns and general extenders). With this in mind, the questionnaire was based on one main assumption: vague expressions (just like other pragmatic devices) serve specific pragmatic functions thus presupposing a relatively high level of English proficiency (Neary-Sundquist, 2014); for this very reason, the survey targeted high school and university students, of legal age, whose pragmatic competence was expected to be fairly advanced after several years of English study. In order to ensure representativeness – and to include a comprehensive range of proficiency levels –, the sample included a variety of different high school and university courses and, accordingly, different branches of study; these comprised the linguistic, scientific, and technological sectors for the upper secondary schools and the economic, juridical, and linguistic sectors for the higher education students.

The questionnaire was structured as follows. The first section was designed to collect the participants' background information (both personal and educational), such as their age, gender, level of education (at the time of the questionnaire administration), and native language; also, participants were asked to provide details on the number of years they had spent studying English and whether they had an English language certification indicating their level of proficiency in accordance with the CEFR scale.

In the second section, participants were presented with three discourse completion tasks, where they had to formulate a vague response according to the informal context provided – and described – in the introductory part of the task itself. For the sake of illustration, consider the following note introducing Item 8: *You are cooking with your mother, and you need her to pass you a kitchen utensil that is next to her, but you don't remember its name, so you ask her: Could you ... ?* The third section of the questionnaire, instead, was devoted to five multiple choice tasks, each introduced by a brief phrase (e.g., *during an oral exam*) situating the task in a specific conversational context – either formal or informal. The final part of the questionnaire addressed the students' personal experience with vague language, exploring whether they had been explicitly taught the linguistic devices in question and how they felt about using them in more or less formal contexts. Such answers will be the focus of Section 3.3.

3. Results

3.1 Quantitative Results

The first categorisation of general extenders in the Italian subcomponent of the TLC focused on the presence and nature of the conjunction introducing the vague expression. Based on this criterion, the retrieved general extenders were divided into three categories: (i) those introduced by *and*, (ii) those introduced by *or*, and (iii) those which were not introduced by any English conjunction; this last category only included instances of the fixed expression

et cetera since other fixed expressions could not be retrieved in the corpus. The frequencies of such categorisation are summarised in Table 1 below.

Table 1. Distribution of general extenders in the TLC (Italian Subcomponent) according to their form

Form	Frequency	%
adjunctive	48	65.8%
disjunctive	19	26.0%
no conjunction	6	8.2%
Total	73	100%

As highlighted in Table 1, the majority of Italian candidates (65.8%) resorted to adjunctive general extenders, while relatively few (26%) employed disjunctive forms or the category without any English conjunction (8.2%). Similarly, a few studies have addressed the preferences of EFL speakers when it comes to choosing general extenders. In this regard, in her study on the ELFA corpus (i.e., a database of one million words consisting of recorded spoken academic ELF interactions collected from four Finnish universities), Metsä-Ketelä (2012) found that, when being vague, the sampled language users mainly resorted to general extenders and, more specifically, to adjunctive forms. In a similar vein, Parvaresh et al. (2012) observed that Persian EFL speakers preferred adjunctive phrases to disjunctive ones; by contrast, the Dutch EFL speakers analysed by Buysse (2014), demonstrated a more frequent use of disjunctive forms.

However, by examining the expressions in greater depth, additional patterns can be identified, along with the candidates' preferred expressions; these are summarised in Table 2 below.

Table 2. Lexical variants and frequency of general extenders in the TLC

Category	Variant	Frequency	%
other fixed expressions	and so on (28) et cetera (6) and all (3) and so forth (1) or this (1)	39	53.4%
forms with *thing(s)* in indefinite pronouns	or something like that (5) and something like that (3) and everything (3) or something (2) or something like this (1) or something else (1) and something like this (1)	16	21.9%
forms with *things*	and other things like that (1) and many other things (1) and things like that (1) and things like this (1) and other things (1) or things (1) or things like that (1) or any things like that (1) or other thing (1) or other things (1)	10	13.7%
forms with *whatever*	or whatever (3) or whatever else (1) and whatever (1)	5	6.8%
forms with *stuff*	and stuff (2) and stuff like that (1)	3	4.1%
Total		73	99.9%

First, one can notice that, when deciding how to express vagueness, more than half of the Italian students (53.4%) tend to prefer the category containing fixed expressions, followed by the forms containing the vague noun *thing* in indefinite pronouns (21.9%), and by the expressions containing *thing* (13.7%) – either in its singular or plural form. As for the two remaining categories – i.e., those characterised by the pronoun *whatever* and the vague noun *stuff* –, they have very low usage rates, especially if compared to the other three categories.

Second, by examining the frequency values in the Table, it is possible to identify which expressions are preferred – and are more typically used – by the Italian students during their examinations. Among the most used general

extenders, one can observe *and so on* (28 hits), *et cetera* (6 hits), *or something like that* (5 hits), followed by *and all*, *and something like that*, *and everything*, and *or whatever* – each occurring three times. Interestingly, the Italian students in the TLC tend to rely heavily on the fixed form *and so on*, thus showing a preference for already-made, pre-fabricated phrases that can be used when vagueness is required.

Third, if one considers similar studies on the preferences of EFL speakers, comparable patterns can be observed across learners of different nationalities. For instance, both the French group investigated by De Cock (2004) and the Finnish group studied by Metsä-Ketelä (2012) primarily used three general extenders: *et cetera*, *or something (like that)*, and *and so on* – with the latter being the most frequently used in both studies. Similarly, Parvaresh et al. (2012) observed that their Persian speakers mainly resorted to *and blah blah blah*, *or something*, and *and so on*. Finally, the Dutch speakers in Buysse's (2014) study largely relied on *or something (like that)*, *and stuff (like that)*, and – once again – *and so on*. What is particularly noteworthy here is that specific expressions (i.e., *and so on*, *et cetera*, and *or something* – as well as its longer variant) tend to recur among speakers from different L1 backgrounds. A possible reason behind this pattern may be that, regardless of their mother tongue, EFL students tend to memorise specific already-made and pre-fabricated expressions and stick to them throughout their conversations when vagueness is required. Such learning mechanism has also been observed in other related studies, in the field of second language acquisition (e.g., Pawley & Syder, 1983; Wood, 2002; Wray, 2002; Schmitt & Carter, 2004), addressing the concept that learners use stored expressions to cope with communicative demands as well as to sound more natural in their speech.

Along these lines, EFL speakers may use these expressions more often, as they (i) commonly appear in classroom textbooks and other pedagogical materials (with several authors noting, for instance, that *and so on* is typical of written rather than spoken registers; see e.g., Biber et al., 1999, p. 117 and Overstreet, 1999, p. 7), and (ii) are likely perceived as fixed or formulaic in nature; consequently, learners may retrieve them more readily from their repertoire of learned expressions rather than relying on more creative ways of expanding other, more colloquial vague expressions (e.g., *and all that kind of stuff*, *or things of that kind*) they might know.

3.2 Qualitative Results

The first aspect that stands out when analysing concordance lines is that no general extender is used to convey attitudes or express a personal judgement about the topic being discussed; on the contrary, all vague expressions are more content-oriented and are thus used to refer to objective or informational content – as exemplified in (3) to (5) below.

(3)

E: do you think it's possible to be rich and famous without losing your privacy?

S: oh it's an hard question erm I think there isn't some person some famous person or rich person er that don't have lose er his ide-ideas erm but er er I I don't know if I can do it erm because er I er when you have money a lot of money a lot of money a lot of er fans *et cetera*, er I think some things change in you.

(4)

E: did you find it easier in Madrid than in Barcelona?

S: yeah Madrid is is is more er open to

E: more international

S: yeah it's more international is more open to to to to people *and so on* erm maybe even the erm er the occasion the I erm Word Youth Day er is is more s erm is more erm er easy er to to

E: yeah

(5)

S: and I see wow it's really it's really beautiful I like er study the stars the planets *and* er *something like that* because I think it's really interesting

In these examples, in fact, all three general extenders are clearly used to expand lists or specific sets of referents rather than to express the candidate's stance regarding the topic being discussed. Notably, the only instance of a more stance-oriented use found in the Italian subcomponent is *and whatever* (example 6).

(6)

E: so erm let's talk about er e-equal opportunities s-some people assert that women are taking over the world what would you say to them?

S: erm well I can not say that I'm totally convinced about this

E: mm mm mm

S: but at the same times I share a lot of the the majority of the because now a nowsaday you can see a lot of reverse situation in the past the man the work

E: mm

S: in n-not only in work but in also in in house er and managing *and whatever* but nowsaday the situation's a little bit different so the woman they have a lot of control a lot of the power the decision power so yes I can not say that's one hundred per cent true true but I think that it's really

E: that so you so you so you think women are taking over the world

S: erm yeah I think that they are

In (6) the candidate has been asked to comment on the contemporary roles of women in society and, in doing so, he indirectly refers to the social stereotypes typically associated with men. To convey this idea, he employs *and whatever* to express a vague generalisation about gender roles. This use is stance-oriented for at least three interrelated reasons: (i) it does not convey concrete factual information, (ii) it signals the speaker's casual and detached attitude – suggesting that he does not wish to elaborate, or commit himself too strongly to, the specific details of traditional male roles – and (iii) it serves an interpersonal and pragmatic function, signalling that the candidate is not making a rigid claim, but rather he is giving a broad comment on the topic at issue. In this sense, *and whatever* produces a casual, stance-oriented tone, whereas a more formal general extender (e.g., *and so on, et cetera*) would have produced a factual, content-oriented tone.

The second aspect that emerges from the analysis of the concordance lines is tightly connected to the first aspect mentioned above. To put it simply, since all instances of the retrieved general extenders are used to refer to facts and to add concrete, additional contents (as exemplified in 3 to 5), all vague expressions were classified as cases of informational vagueness, with speakers mainly using general extenders to list items. Furthermore, examining the context of each occurrence of general extender revealed that these expressions can also be classified as instances of discourse-related vagueness. In other words, speakers use these expressions to manage the constraints of online production: since they have limited time to plan their utterances ahead, they rely on these devices either to buy time or to conclude their turns smoothly. Such strategies can be observed in examples (7) and (8), where general extenders are accompanied by pauses, repetitions, and hesitation devices, which clearly signal the cognitive pressures involved in real-time speech.

(7)

it's the reason why it's considered the masterpiece as he as he er painted er this erm this space with a with a erm er the fond er er eyes mouth *and all* the erm also the landscape and er the colours are real so

(8)

we need money to have our goods er so medicines er erm for food er school also for travelling erm *and so forth* state for the tax and services erm

Before turning to the next Section, two more aspects that emerged from the analysis of the concordance lines are worth commenting on. First, all retrieved general extenders can be classified as multifunctional, meaning that they can serve multiple pragmatic purposes at the same time. In the present corpus, for instance, general extenders perform functions associated with both informational vagueness (e.g., listing items, indicating approximation, or leaving details unspecified) and discourse-related vagueness (e.g., signalling hesitation, holding the floor, or marking a transition). The multifunctional nature of these pragmatic devices has also been previously noted by other scholars (e.g., Overstreet, 1999; Cheshire, 2007; Tagliamonte & Denis, 2010).

Second, no instances of relational vagueness were found in the Italian subcomponent of the TLC – apart from example (6), which has already been discussed above. This may suggest that EFL learners may not be fully aware of the range of nuanced functions that general extenders – and that of other pragmatic markers (Aijmer, 2009) – can serve and may therefore limit their use to the most basic ones.

3.3 The Students' Questionnaire

Before addressing the questionnaire results, it is important to provide an overview of the participants' profiles. The participants consisted of 209 Italian EFL learners from both secondary and university-level institutions, with a balanced gender distribution and an age range primarily between 18 and 24 years. Moreover, while all students had studied English for several years, only a minority of them held an official language certification, most

commonly at the intermediate level (B1–B2).

Against this background, the questionnaire responses offer valuable insights into the students' experiences with vague language throughout their educational journey; in order to gather such data, this section of the questionnaire combined closed and open-ended questions – with both types being presented in Italian – to allow students to express themselves more freely.

First, students were asked whether vague language had been explicitly taught in the English courses they had attended up to that point. The vast majority (83.7%) responded negatively, indicating that vagueness and related expressions were not included in their English language programs. In contrast, 16.3% reported having received explicit instruction on this linguistic phenomenon; however, most of these students (73.6%) were enrolled in a university course, while the remaining (26.4%) were still attending high school. Such answers are visually summarised in Figure 1 below.

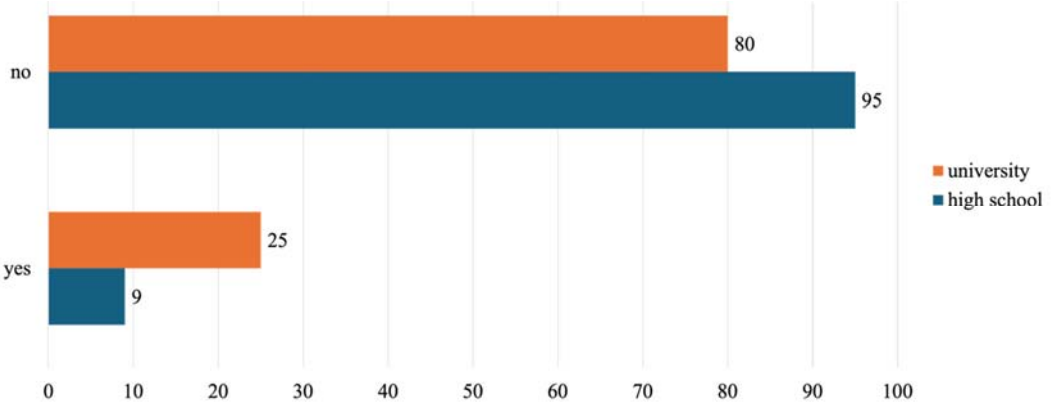

Figure 1. Students' responses regarding explicit instruction on vague language in their English courses

In the following item, students had to indicate, through a multiple-choice task (allowing more than one answer), how they got familiar with vague expressions apart from direct teaching. First, four predefined options were provided. They corresponded to the most common informal learning sources – mostly associated with media-based exposure – and included: (i) reading books in English, (ii) listening to English songs, (iii) engaging with screen-based entertainment (both movies and TV series), and (iv) online communication through social media. Additionally, the multiple-choice task also offered an open-ended section, where participants could specify other ways in which they had encountered the English vague expressions at issue. In this regard, the third and fourth categories were the most frequently selected, with 130 and 127 responses, respectively; then, the *learning-through-songs* category was chosen by 74 students, while the *learning-through-books* option was the least selected, with only 29 responses. Such results are visually summarised in Figure 2 below.

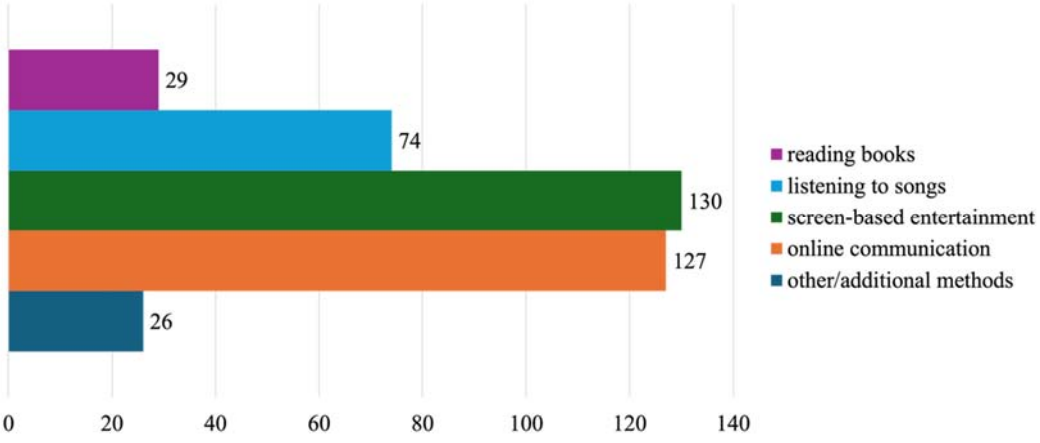

Figure 2. Students' reported sources of informal exposure to English vague expressions. Multiple selections were allowed

Interestingly, 26 students also added their own responses under the *other* option and provided further details on their learning experience with vague language; unfortunately, not all answers could be categorised, as some students used the open-ended section to leave brief or incomplete comments. Consequently, the valid responses could be classified into four categories, according to the different type of exposure, or learning strategy, indicated by the students. Along these lines, the respondents got familiar with vague expressions through (i) social exposure (e.g., talking to English native speakers, either during their time abroad or through contact with English-speaking friends and relatives), (ii) private school attendance (where the exposure to English tends to be more immersive and advanced than in formal classroom settings), (iii) other media-based content (e.g., playing videogames), and (iv) self-guided learning (i.e., by grasping their meaning intuitively or translating similar Italian expressions).

Some of the students also used the open-ended section to underline the fact they were not aware of this existence of this linguistic phenomenon; for instance, *Cage4* (an 18/19-year-old male student), jokingly wrote that "È la prima volta che le sento ma tanto domani non le ricordo perché ho risposto a caso perché non ci capisco nulla <3" [It's the first time I've heard them but I won't remember them tomorrow anyway because I answered at random because I can't make sense of it <3 (my translation)].

4. Discussion

The present analysis has shed light on the habits of Italian EFL learners regarding their use of vague language; in this respect, both the TLC and the students' questionnaire proved to be particularly fruitful for examining this phenomenon.

In particular, the analysis revealed that Italian students are generally aware of the existence of vague language, and for some, it even forms part of their linguistic repertoire. However, the data also suggests that those who employed vague expressions in their speech were not fully aware of their nuanced meanings or pragmatic potential. This is evident from the fact that all retrieved instances of general extenders were predominantly used either for informational purposes (e.g., to refer to factual content, list items, or indicate approximation within discourse) or to address discourse-related constraints (e.g., managing turn-taking, buying time during speech planning, or concluding an utterance smoothly).

This pattern differs significantly from the use of vague expressions by native – or more proficient – speakers and it points to a relatively limited awareness of the multifunctional nature of these forms. While general extenders can also serve important interpersonal functions (e.g., mitigating claims, expressing shared knowledge, or conveying informality; Overstreet, 2005; Tagliamonte & Denis, 2010), the Italian learners in the corpus tend to overlook such uses and rely on such expressions in a more literal and content-oriented way instead (Note 2). This limited awareness is further reflected in the forms that students prefer; in this regard, most students tend to resort to the same fixed expressions (i.e., *and so on*, *et cetera*, and *or something like that*), often at the expense of more creative or varied alternatives.

A possible explanation for this tendency may be related to the fact that vague language is rarely addressed in the Italian educational system; this may lead learners to underuse the general extenders' full potential, and the variety of expressions offered by the English language.

The limited attention given to vague language in the Italian educational system is also confirmed by the students' questionnaire, where the substantial majority of respondents reported that vague expressions were not included in their English curricula. Rather, they became familiar with such linguistic items primarily through informal learning sources, such as TV series, YouTube videos, music, and social media platforms. These findings are consistent with previous research on Italian EFL learners' learning habits, which shows that most students engage with English through audiovisual input and social media in informal, self-directed ways rather than through formal instruction (Cappelli & Simi, 2024; Pavesi & Bianchi, 2024; Ghia, 2024).

Interestingly, such informal exposure seems to play a compensatory role in the absence of explicit teaching; in this sense, engaging with English-language media allows students to frequently encounter vague expressions in authentic, naturalistic contexts. However, because these encounters occur without pedagogical guidance, learners often develop only a partial understanding of the interpersonal and pragmatic functions of such expressions. Seen from a wider perspective, such findings seem to highlight the dual nature of informal learning. On the one hand, it provides learners with valuable exposure to authentic English input; on the other hand, because such exposure is not supported by the right pedagogical feedback, it does not necessarily promote a deep awareness of how these expressions contribute to the management of social relationships and conversational tone.

Taking everything into account, a more explicit teaching about vague expressions – as well as other pragmatic

devices – should be encouraged in order to provide EFL learners with a wider and more pragmatically aware repertoire of expressions and communicative tools. Teachers, especially in more advanced courses, can play a crucial role in guiding learners to notice how vague expressions function in authentic communication, reflecting Aijmer's claim that "the formal and functional properties of pragmatic markers need to be 'talked about' in the classroom" (2013, p. 112).

References

Aijmer, K. (2009). "so er I just sort I dunno I think it's just because…": A corpus study of I don't know and dunno in learners' spoken English. In A. H. Jucker, D. Schreier, & M. Hundt (Eds.), *Corpora: Pragmatics and discourse* (pp. 151–168). Amsterdam/New York: Rodopi. https://doi.org/10.1163/9789042029101_009

Aijmer, K. (2013). *Well I'm not sure I think…* The use of *well* by non-native speakers. In G. Gilquin, & S. De Cock (Eds.), *Errors and disfluencies in spoken corpora* (pp. 93–116). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/bct.52.04aij

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written language*. London, England: Longman.

Buysse, L. (2014). 'We went to the restroom or something'. General extenders and stuff in the speech of Dutch learners of English. In J. Romero-Trillo (Ed.), *Yearbook of corpus linguistics and pragmatics 2014* (vol. 2, pp. 213–237). Dordrecht, Germany: Springer. https://doi.org/10.1007/978-3-319-06007-1_10

Cappelli, G., & Simi, N. (2024). Instagram English microlessons. A new genre for non-formal EFL learning. *Lingue e Linguaggi*, *66*, 185–218. https://doi.org/10.1285/i22390359v66p7

Channell, J. (1994). *Vague language*. Oxford, England: Oxford University Press.

Cheshire, J. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics*, *11*(2), 155–193. https://doi.org/10.1111/j.1467-9841.2007.00317.x

Cole, J., & Vanderplank, R. (2016). Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning. *System*, *61*, 31–42. https://doi.org/10.1016/j.system.2016.07.007

Cutting, J. (Ed.). (2007). *Vague language explored*. Basingstoke, England: Palgrave Macmillan. https://doi.org/10.1057/9780230627420

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *BELL: Belgian Journal of English Language and Literatures*, *2*, 225–246.

Dressman, M., & Sadler, R. W. (Eds.). (2020). *The handbook of informal language learning*. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119472384

Gablasova, D., Brezina, V., & McEnery, T. (2019). The Trinity Lancaster corpus development, description and application. *International Journal of Learner Corpus Research*, *5*(2), 126–158. https://doi.org/10.1075/ijlcr.5.2

Ghia, E. (2024). What drives L2 viewers' informal access to English? A case study on audiovisual input. *Lingue e Linguaggi*, *66*, 43–73. https://doi.org/10.1285/i22390359v66p43

Ghiselli, S. (2024). World-wide English. The Internet as a language learning tool. *Lingue e Linguaggi*, *66*, 75–103. https://doi.org/10.1285/i22390359v66p75

Gilquin, G., & Granger, S. (2022). Using data-driven learning in language teaching. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (2nd ed., pp. 430–442). London, England: Routledge. https://doi.org/10.4324/9780367076399-30

Lancaster University, The ESRC Centre for Corpus Approaches to Social Science (CASS), & Trinity College London. (n.d.). *Help Documentation for the Trinity Lancaster Corpus*. Trinity Lancaster Corpus Hub (TLC Hub). Retrieved October 8, 2025, from http://corpora.lancs.ac.uk/trinity/assets/help.pdf

Machetti, S. (2006). *Uscire dal vago*. Bari, Italy: Laterza.

Manca, E. (2024). Social media and online informal learning of English university language students' preferences, attitudes and perceptions. *Lingue e Linguaggi*, *66*, 219–239. https://doi.org/10.1285/i22390359v66p219

Metsä-Ketelä, M. (2012). Frequencies of vague expressions in English as an academic lingua franca. *Journal of English as a Lingua Franca*, *1*(2), 263–285. https://doi.org/10.1515/jelf-2012-0019

Neary-Sundquist, C. (2014). The use of pragmatic markers across proficiency levels in second language speech. *Studies in Second Language Learning and Teaching*, *4*(4), 637–663. https://doi.org/10.14746/ssllt.2014.4.4.4

Overstreet, M. (1999). *Whales, candlelight, and stuff like that: General extenders in English discourse*. New York, NY: Oxford University Press. https://doi.org/10.1093/oso/9780195125740.001.0001

Overstreet, M. (2005). And stuff und so: Investigating pragmatic expressions in English and German. *Journal of Pragmatics*, *37*(11), 1845–1864. https://doi.org/10.1016/j.pragma.2005.02.015

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130–149. https://doi.org/10.1017/s0267190512000098

Parvaresh, V., Tavangar, M., Rasekh, A. E., & Izadi, D. (2012). About his friend, how good she is, and this and that: general extenders in native Persian and non-native English discourse. *Journal of Pragmatics*, *44*(3), 261–279. https://doi.org/10.1016/j.pragma.2011.12.003

Pavesi, M. (2024). The informalisation of English language learning. A first national research project. *Lingue e Linguaggi*, *66*, 17–42. https://doi.org/10.1285/i22390359v66p17

Pavesi, M. (2025). The informal use and learning of English: A new era. In M. Pavesi, F. Bianchi, S. Bruti, & R. Zago, *In contact with English. Informal media-driven learning among Italian university students* (pp. 21–46). Rome, Italy: Carocci.

Pavesi, M., & Bianchi, F. (2024). 'In the presence of English': Informal English language learning. An introduction. *Lingue e Linguaggi*, *66*, 7–16. https://doi.org/10.1285/i22390359v66p7

Pavesi, M., Ghia, E., Bianchi, F., Bruti, S., & Simi, N. (2025). Experiencing English informally through the media. In M. Pavesi, F. Bianchi, S. Bruti, & R. Zago, *In contact with English. Informal media-driven learning among Italian university students* (pp. 79–115). Rome, Italy: Carocci.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards, & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–226). London, England: Longman.

Peirce, C. S. (1902). "Vagueness". In M. Baldwin (Ed.), *Dictionary of philosophy and psychology II*. London, England: Macmillan.

Sabet, P. G., & Zhang, G. Q. (2015). *Communicating through vague language: A comparative study of L1 and L2 speakers*. Basingstoke, England: Palgrave Macmillan. https://doi.org/10.1057/9781137486387

Samigoullina, A. D. (2020). Vague expressions in teaching English as a foreign language. *Russian Linguistic Bulletin*, *1*(21), 15–18.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1–22). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.9

Tagliamonte, S. A., & Denis, D. (2010). The stuff of change: General extenders in Toronto, Canada. *Journal of English Linguistics*, *38*(4), 335–368. https://doi.org/10.1177/0075424210367484

Toffoli, D., Sockett, G., & Kusyk, M. (Eds.). (2023). *Language learning and leisure: Informal language learning in the digital age*. Berlin/Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110752441

Voghera, M. (2012). Chitarre, violino, banjo e cose del genere. In M. Voghera, & A. M. Thornton (Eds.), *Per Tullio De Mauro. Studi offerti dalle allieve in occasione del suo 80° compleanno* (pp. 341–364). Rome, Italy: Aracne. https://doi.org/10.4399/97888XXXXXXXX18

Voghera, M., & Collu, L. (2017). Intentional vagueness: A corpus-based analysis of Italian and German. In M. Napoli, & M. Ravetto (Eds.), *Exploring intensification: Synchronic, diachronic and cross-linguistic perspectives* (pp. 371–390). Amsterdam/New York: John Benjamins. https://doi.org/10.1075/slcs.189.19vog

Wagner, J. (2015). Designing for language learning in the wild: Creating social infrastructures for second language learning. In T. Cadierno, & S. W. Eskildsen (Eds.), *Usage-based perspectives in second language learning* (pp. 75–102). Berlin; Germany: Mouton de Gruyter. https://doi.org/10.1515/9783110378528-006

Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, *20*(1), 1–15. https://doi.org/10.18806/tesl.v20i1.935

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Notes

Note 1. In this regard, Paquot and Granger (2012) emphasise the advantages of using learner corpora to investigate the set of formulaic sequences in learner language. These corpora (i) contain continuous stretches of discourse rather than isolated items, (ii) arise from tasks that allow learners to choose their own wording (e.g., essays, interviews) and (iii) are available in electronic format, facilitating the use of automated tools for analysis.

Note 2. In this respect, De Cock (2004) also observed that the EFL speakers in her study tended to use certain pragmatic markers (e.g., *sort of* and *kind of*) in patterns more typically associated with their literal, non-pragmatic use (p. 239).