

Learning English-Italian False Friends in the Digital Age: Does Informal Online Exposure to English Help?

Silvia Masi¹ & Sabrina Noccetti¹

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

Correspondence: Silvia Masi, Department of Philology, Literature and Linguistics, University of Pisa, via Santa Maria 67, 56126, Pisa, Italy. Tel: +39 050 2215853. E-mail: silvia.masi@unipi.it

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

doi:10.5539/ijel.v16n3p55 URL: <https://doi.org/10.5539/ijel.v16n3p55>

Abstract

Research in EFL has long highlighted the problematic nature of false friends (FF), which are words in different languages that are formally similar but semantically divergent and often lead to misunderstandings. However, studies on how such items are effectively learned remain limited, particularly in the context of English as learned by Italians. Although pedagogical translation and an explicit focus on form are commonly considered effective in addressing lexical areas prone to cross-linguistic interference, recent evidence suggests that variables related to formal instruction alone do not fully account for the acquisition of English–Italian FF, which are also largely absent from curricular materials. This raises the question of how learners encounter and master such items, and whether informal digital exposure plays a role. This study investigates the relationship between the level of knowledge of FF among a sample of young Italian university students, as assessed through a vocabulary test specifically designed for this purpose, and the students' informal contact with English, as revealed by a questionnaire on their digital media use habits. Results show that media exposure – measured in terms of quantitative variables such as frequency and duration, and also including platform diversity, film viewing and subtitle use – does not emerge as a primary predictor of performance, while competence in this lexical set tends to correlate more significantly with overall English proficiency and general lexical competence. The disaggregation of results by grammatical category of target items also highlights uneven sensitivity across word classes, with adjectives (as opposed to nouns and verbs) showing the only statistically significant correlation. These findings point to the need for further research into qualitative dimensions of learning this hybrid and elusive lexical set, with a refined methodology and greater attention to individual differences and word-class effects.

Keywords: EFL, English–Italian false friends, incidental learning, informal learning, L2 acquisition, vocabulary learning

1. Introduction

False friends (henceforth FF) are lexical items that are phonetically similar across languages but differ in meaning as a result of accidental development or of historical language contact followed by semantic change. Different classifications have been proposed in the relevant literature. Following Chamizo-Domínguez (2008), two basic types can be identified: chance FF and semantic FF. The former share the same form but have different etymologies and completely different meanings in different languages. With reference to the English–Italian language pair, an example could be that of *estate*, which in English means 'property', while in Italian it identifies one of the four seasons (i.e., 'summer'), as the two forms derive from Medieval and Classical Latin, respectively. Semantic FF, by contrast, share an etymological origin (they are therefore cognates), and although their meanings differ across languages, semantic relations between them can still be identified (also see Chamizo-Domínguez & Nerlich, 2002). As a matter of fact, the meanings of individual lexical items may diverge over time, undergoing processes such as, *inter alia*, narrowing and broadening (O'Neill & Casanovas Català, 1997), alongside the development of distinct connotations and figurative extensions cross-linguistically (Chamizo-Domínguez & Nerlich, 2002).

An example of a semantic FF is *fastidious* (Chamizo-Domínguez, 2008), which in English means 'very attentive to detail' or 'hard to please because of high standards'. The Italian *fastidioso*, by contrast, means 'annoying' or 'irritating' and carries a more negative evaluative polarity. Historically, the English form also shared the negative sense of their common Latin root, but this meaning has become obsolete, with the word undergoing semantic amelioration (Note 1) over time.

English has borrowed extensively over the centuries, with Latin constituting one of its primary sources of lexical influence. In fact, FF have long been recognised as a major source of difficulty in foreign language learning, especially among Romance language speakers learning English, contributing to an increased likelihood of deceptive cross-linguistic mappings and interference in L2 acquisition. Indeed, Lado (1957) associates FF with the highest degree of difficulty based on interlinguistic relations of form and meaning. Among semantic FF, those with completely different meanings are referred to as absolute or total FF (e.g., English *fastidious* and Italian *fastidioso*), while those sharing part of their meanings cross-linguistically are partial FF (e.g., English *abusive* and Italian *abusivo*, see section 2). Absolute FF have been found to create less confusion than partial ones. This is especially true in the early stages of learning, although incomplete resemblance resulting from only partial semantic overlap may pose difficulties even at higher levels of proficiency (O'Neill & Casanovas Català, 1997). Gouws, Prinsloo, and De Schryver (2004, p. 804) talk about a “semantic continuum of false friends” and argue for adequate lexicographic treatment of items which typically require disambiguation.

Research on L2 vocabulary learning has shown that translation and explicit focus on form are the best approaches to dealing with lexical areas prone to cross-linguistic interference (see for example Kerr, 2014; Chacón Beltrán, 2006). Experimental results, that is, demonstrate that form-focussed instruction leads to better learning outcomes for FF than simple exposure to enhanced input (Chacón Beltrán, 2006), thereby underscoring the effectiveness of explicit pedagogical intervention and the need for targeted instructional methodology. Such results are congruent with a Plurilingual Turn in language teaching (Carreres, Noriega-Sánchez, & Pintado Gutiérrez, 2021), and with advances in neuroscience that show that humans possess a single, integrated linguistic system. In more detail, bilinguals access their linguistic repertoire non-selectively, relying on their L1 when processing an L2 (Abutalebi, 2008; Starreveld, De Groot, Rossmark, & Van Hell, 2014; Libben, 2017). This appears to be all the more natural and plausible when learning takes place in a foreign environment. Such a tendency, combined with learners' expectation for interlinguistic equivalence – unless contradictory evidence is encountered (Chacón Beltrán, 2006; Swan, 2017), can result in negative transfer (Selinker, 1972) in case of false transparency (Laufer, 1997), particularly with partial FF, a trend that is especially pronounced at elementary and intermediate levels of proficiency but which can be found among highly advanced users as well (Chacón Beltrán, 2006; Anissimova, 2015).

Although several studies on FF have examined English in combination with other languages – such as Spanish, Catalan, French, and German, to name a few (O'Neill & Casanovas Català, 1997; Chamizo-Domínguez & Nerlich, 2002; Chacón Beltrán, 2006), Italian has received comparatively limited attention (see Aliyar, 2020 esp. on Italian and Spanish). In addition to the lexicographic resource by Browne, Mendes and Natali (1995) (also see Hill, 1982), more recent studies of particular relevance include Marini (2024) and Masi and Noccetti (2025).

Marini (2024) examines the challenges posed by FF for Italian learners of English across different educational levels. The study employs a questionnaire-based design combining background questions on proficiency, exposure, and motivation with translation tasks, multiple-choice exercises, and isolated word-meaning recognition. The results indicate a general progression across educational levels, with university and postgraduate participants achieving higher overall accuracy, despite some difficulties. Proficiency levels are determined through self-assessment and performance on the test tasks rather than through an external standardised proficiency measure. Consequently, FF performance is interpreted as reflecting general competence, even though this lexical set is known to be exceptionally sensitive to cross-linguistic influence and has been found to be problematic even at higher levels of proficiency. In addition, the analysis relies on descriptive and interpretive comparisons rather than inferential statistical procedures. Overall, however, the findings are consistent with previous research highlighting the role of explicit instruction, while also suggesting that formal education alone may not be sufficient and that sustained and varied exposure to the target language plays an important role in learners' development.

Masi and Noccetti (2025) explored how FF knowledge correlates with educational background. The participants in the experimental study were Italian undergraduate students in the humanities. Relevant findings were derived from statistical correlations between their responses to a questionnaire on their habits of contact with English and results from a specifically designed vocabulary test on FF. The Oxford Placement Test (OPT), i.e., an adaptive online general English assessment which includes grammar, vocabulary, and listening comprehension tasks to assign learners to six proficiency bands (from A1 to C2), was also preliminarily administered to ascertain students' proficiency level independently of their knowledge of FF. The variables of the educational background focussed on in this phase of the research were the type of high school attended before accessing university, frequency of travel to English-speaking countries and attendance at English courses abroad. The results showed that knowledge of FF among the Italian learners who participated in the study was limited and weakly to moderately correlated with their overall English proficiency, which, in turn, was associated with the type of high school attended

(especially Linguistic, Classical, and Scientific schools). Differences among B2, C1, and C2 proficiency levels were minimal, indicating that advanced proficiency does not guarantee native-like competence in FF. Furthermore, both the frequency of travel to English-speaking countries and attendance at language courses abroad displayed negligible correlations with general proficiency and FF competence. Overall, then, the study suggests that the variables of the educational background mentioned above hardly predict a good command of this lexical set, so that other factors are likely to be involved in its acquisition.

Admittedly, FF are heterogeneous with respect to meanings, part of speech, frequency of occurrence, degree of formality, and the proficiency levels with which they are associated, which do not easily allow for systematic accounts in formal instruction. Moreover, they are elusive in the sense that they tend to be excluded from standard ELT materials, which are usually intended for linguistically heterogeneous learning contexts without addressing the specific needs of learners' first languages (Roca-Varela, 2015). When tackled, their instruction is dominated by traditional grammar-translation practices, with limited scope for originality or pedagogical experimentation. As an alternative, the use of more diversified approaches is advocated, covering game-based tasks, visual materials, blogs and mobile language-learning applications, which are intended to make learning more motivating, while at the same time encouraging learners' autonomy.

Autonomy in learning takes on particular importance in the contemporary landscape, shaped by globalisation, pervasive multilingualism, and continuous media exposure. English now circulates extensively across both traditional and digital media, greatly increasing the opportunities for informal language learning, i.e., the acquisition of a language that occurs outside formal instructional settings through unstructured, self-directed, and often incidental exposure to the target language (see Dressman, 2020; Pavesi & Ghia, 2020; Amengual-Pizarro & Alonso Alonso, 2024; Pavesi, 2024). With regard to online platforms and language-learning applications, these activities are often classified as non-formal or informal, since they are pedagogically planned yet entirely managed by the learner (Amengual-Pizarro & Alonso Alonso, 2024). Additionally, recent research indicates that a substantial portion of second-language vocabulary acquisition takes place beyond formal instructional settings, through non-formal – and often incidental – exposure to new lexical items in digital environments (Lee & Dressman, 2018; Pavesi, Bianchi, Bruti, & Zago, 2025; Webb, Uchihara, & Yanagisawa, 2023). This situation prompts reflection on the extent to which such forms of exposure might actually influence the acquisition of FF, and whether they can offset, or at least complement, their sparse and non-systematic treatment in standard instructional materials.

In other words, since the results of previous research on variables closely related to a formal contact with English did not yield sufficiently strong correlations with FF knowledge, and given the scant coverage of such lexical items in standard teaching resources, it is worth asking how learners actually encounter them and may succeed in mastering them. The present contribution addresses this question by resuming and expanding on Masi and Noccetti (2025) while focussing on the following issue: in today's highly connected society, saturated with countless opportunities for digital exposure, might learners' habits of informal – and presumably incidental – contact with English have an impact on the learning of FF?

The next section outlines the methodology underpinning the study, while the subsequent sections present the findings, followed by their discussion and concluding remarks.

2. Methodology

Building on Masi and Noccetti (2025), the data for the present study were mainly collected through a questionnaire on learners' background and habits of exposure to English and the False Friends Vocabulary Test (henceforth FFVT). These tools were administered to a sample of 87 Italian undergraduate students in their early twenties, enrolled in BA programmes in the humanities at our university.

Data collection began with the questionnaire, which consisted of several sections designed to elicit information on participants' family background, linguistic profile, prior education, motivation for learning English, and patterns of media use and reading. To limit the impact of potential confounding factors, both students who were fully bilingual (i.e., possessing native-like competence in a second language) and those who were diagnosed with specific learning disorders were excluded from the study. The factors considered at this stage of the analysis focussed primarily on quantitative variables of informal exposure to English through digital media, esp. social media such as Instagram, Tik Tok, Facebook, the online video-sharing platform of YouTube, but also subscription-based video-on-demand services such as Netflix, Disney+, and Amazon Prime Video.

Participants were then asked to complete the FFVT, specifically developed to assess their knowledge of a selected set of FF, although this purpose was not disclosed to avoid influencing responses. For the selection, different lexicographic resources (Note 2) were consulted and cross-checked to ensure accurate meaning disambiguation.

The resulting selection includes words ideally associated with different levels of lexical proficiency in an EFL context, in order to be representative of the targeted lexical set as a whole. At this stage of the research, the test items mainly centred on absolute or full FF, which tended to exclude semantic overlap between English and Italian. For instance, in the case of English *actual* – ‘real’, ‘existing’ – and Italian *attuale* – ‘present’, ‘current’, none of these meanings coincide, while in the case of English *abusive* – ‘violent’, ‘insulting’, ‘illegal’ – and Italian *abusivo* – ‘unauthorised’, ‘illegal’, semantic resemblance is partial and may require more nuanced disambiguation. For this reason, *abusive* was not included in our set of FF.

The FFVT was administered via the university’s Moodle platform and was limited to 40 minutes, with backward navigation disabled. The test comprised 70 lexical items, yielding a total of 76 questions organised into three main sections according to part of speech: 30 nouns, 26 adjectives (also covering a few adverbs), and 14 verbs, with distractors included for each category. Table 1 presents the target items, excluding distractors, by part of speech.

Table 1. Selected target items and corresponding Italian FF subdivided by part of speech (excluding distractors)

Nouns	Adjectives and adverbs	Verbs
Affluence – Affluenza	Actual – Attuale	Accommodate – Accomodare
Antics – Antichità	Capacious – Capace	Appoint – Appuntare
Argument – Argomento	Conceited – Concitato	Arrange – Arrangiare/si
Canteen – Cantina	Consistent – Consistente	Attend – Attendere
Confetti – Confetti	Contemptuous – Contento	Confront – Confrontare
Delusion – Delusione	Definitely – Definitivamente	Conjure – Congiurare
Emergence – Emergenza	Educated – Educato	Demand – Domandare
Estate – Estate	Eventual – Eventuale	Injure – Ingiuriare
Fabric – Fabbrica	Fastidious – Fastidioso	Pretend – Pretendere
Factory – Fattoria	Genial – Geniale	Quarrel – Querelare
Furniture – Fornitura	Gracious – Grazioso	Recover – Ricoverare
Ingenuity – Ingenuità	Inhabited – Inabitato	
Lecture – Lettura	Lurid – Lurido	
Library – Libreria	Morbid – Morbido	
Mansion – Mansione	Sensible – Sensibile	
Mess – Messa	Sensitive – Sensitivo	
Physician – Fisico	Spirited – Spiritato	
Preservative – Preservativo	Terrific – Terrificante	
Recipient – Recipiente	Ultimately – Ultimamente	
Rumour – Rumore	Voluble – Volubile	
Sympathy – Simpatia		

The structure of the answer options was inspired by Nation’s Vocabulary Test (Nation, 1997), and for each target word, four Italian alternatives were provided:

- 1) the correct translation,
- 2) the Italian FF,
- 3) an option semantically related to the correct answer (e.g., through synonymy or other sense relation),
- 4) a word semantically related to the Italian FF.

While the correct translation and the semantically related option to it awarded one point and half a point, respectively, the other two options did not receive any points.

Each of the three sections contained a couple of preliminary diagnostic questions aimed at eliciting participants’ self-evaluation of their familiarity with the target words. They required students to select from a list the words they believed they had previously encountered and identify those whose meanings they recognised with confidence. By way of example, the following preliminary questions introduce the noun FF:

- *Quali dei seguenti nomi inglesi hai già incontrato in passato? Riporta nella casella sottostante solo quelli di cui NON sei sicura/o di conoscere il significato* (Which of the following English nouns have you encountered before? In the box below, list only those whose meaning you are NOT sure you know).

- *Indica quali dei seguenti nomi inglesi conosci. Seleziona solo quelli di cui conosci per certo il significato, altrimenti lascia in bianco* (Indicate which of the following English nouns you know. Select only those whose meaning you are certain of; otherwise, leave them blank).

They were followed by the actual query listing the target items to be associated with the correct Italian equivalent. Below is the one on the subset of noun FF,

- *Seleziona la parola italiana che, tra le opzioni date, si avvicina maggiormente al significato dei nomi inglesi che seguono* (Select the Italian word that, among the given options, most closely matches the meaning of the following English nouns).

Table 2 shows representative examples of English target words and their corresponding Italian options for each of the three main parts of speech of FF.

Table 2. Examples of target words and respective options from the FFVT

Target word	Correct Italian translation	Italian FF	Option semantically related to the correct translation	Option semantically related to the Italian FF
Affluence	<i>ricchezza</i> (wealth)	<i>affluenza</i> (flow, turnout)	<i>benessere</i> (well-being)	<i>partecipazione</i> (participation)
Actual	<i>effettivo</i> (real)	<i>attuale</i> (current)	<i>autentico</i> (authentic)	<i>moderno</i> (modern)
Appoint	<i>nominare</i> (nominate)	<i>appuntare</i> (sharpen, pin, note down)	<i>incaricare</i> (entrust)	<i>affinare</i> (finish)

For illustrative purposes, the table follows the order of alternatives (1–4) described above and includes back-translations; however, in the FFVT the options were randomised and the translations were obviously excluded.

The FFVT results were first correlated with FF frequency as independently determined by reference to corpora data. Subsequently, associations with questionnaire responses were examined, focussing on the diversity of media platforms used and on other quantitative variables related to the frequency and duration of exposure, with additional attention to film viewing and subtitle use. Finally, a subset of informants' data was further analysed for general proficiency level, using the OPT (<https://elt.oup.com/feature/global/oxford-online-placement/>), and general lexical competence, using the Vocabulary Size Test (VST, cf. Nation & Beglar, 2007). More detailed information on the analyses is provided in the next section.

3. Results

As previously stated, the objective of this research is to investigate whether exposure to digital media input can facilitate the implicit learning of the lexical set known as FF. This stage of the study specifically aims to identify the exposure variables that most significantly influence learning. These variables are correlated with participants' performance in the FF assessment. Following the correlation of the average frequency of all FF included in the study with the FFVT, as well as the frequency of FF disaggregated by grammatical category (nouns, verbs, and adjectives) to determine whether more familiar FF tend to occur more frequently in digital media input, the focus then shifted to the investigation of additional variables. The objective was to identify which factors exert the greatest influence on the learning process. The variables to be considered include diversity of digital media, duration of media exposure, subtitle usage, and general vocabulary knowledge.

3.1 Frequency of FF and FFVT Results

The primary objective of this analysis was to determine whether a quantifiable relationship exists between the frequency of specific FF within digital English corpora and university students' performance on the FFVT involving the same lexical items. By comparing the mean frequency values retrieved from two digital corpora (see Table 3), *YouGlish.com* and the *Sketch Engine English Web 2020 (enTenTen)* (Note 3), with experimental test scores, this study seeks to investigate the extent to which informal exposure to web-based language data influences the acquisition or recognition of FF. From a theoretical perspective, a strong positive correlation would suggest that higher levels of environmental exposure contribute significantly to lexical accuracy. Conversely, a weak or non-existent correlation would indicate that exposure alone is insufficient to overcome the cognitive interference typical of FF, suggesting that learning does not follow a direct linear relationship based solely on input frequency.

Table 3. Distribution of FF frequencies across the *Youghlish* and *EnglishWeb* corpora, and mean frequency values

FF	Cat	Frequency on <i>Youghlish</i>	Frequency on <i>Englishweb2020</i>	Mean Frequency values
Attend	v	20531	4728626	2374578.5
Library	n	67387	3726115	1896751
Actual	adj	96436	2996599	1546517.5
Argument	n	56064	2996019	1526041.5
Definitely*	adv.	171905	2411285	1291595
Estate	n	24466	2340730	1182598
Appoint	v	2634	2255628	1129131
Ultimately*	adv.	79382	1814503	946942.5
Demand	V	51758	1802855	927306.5
Recover	V	14463	1838023	926243
Arrange	V	4420	1733016	868718
Consistent	adj	35252	1672238	853745
Physician	N	14234	1457807	736020.5
Factory	N	17645	1417904	717774.5
Sensitive	adj	25127	1235422	630274.5
Lecture	N	38579	1212579	625579
Fabric	N	16186	1162668	589427
Recipient	N	8114	995271	501692.5
Injure	V	1039	869674	435356.5
Confront	V	8263	794381	401322
Mess	N	25659	582669	304164
Pretend	V	14185	573760	293972.5
Inhabited	adj	2110	385440	193775
Sympathy	N	4304	380117	192210.5
Emergence	N	4969	358667	181818
Sensible	adj	5997	316661	161329
Terrific	adj	12546	309697	161121.5
Mansion	N	3682	287202	145442
Eventual	adj	2591	278621	140606
Rumour	N	345	201644	100994.5
Delusion	N	1865	154394	78129.5
Gracious	adj	4025	144588	74306.5
Conjure	V	1024	138825	69924.5
Ingenuity	N	2812	98783	50797.5
Spirited	adj	1590	95422	48506
Educated	adj	17456	62545	40000.5
Preservative	N	347	71326	35836.5
Canteen	N	571	68899	34735
Morbid	adj	942	50625	25783.5
Quarrel	V	635	47333	23984
Affluence	N	587	28959	14773
Confetti	N	768	28223	14495.5
Lurid	adj	216	27300	13758
Genial	adj	133	26817	13475
Accommodate	V	83	24527	12305
Contemptuous	adj	193	24021	12107
Fastidious	adj	201	14687	7444
Conceited	adj	230	13518	6874
Capacious	adj	290	11069	5679.5
Furniture	N	10233	118	5175.5
Voluble	adj	33	4203	2118
Antics	N	707	2381	1544

To assess the relationship between word frequency in English web corpora and student test results, the Pearson Correlation Coefficient (r) was employed (Note 4). This parametric test was selected given the continuous nature of the variables and the assumption of linearity. The analysis yielded a statistically significant positive correlation, $r = .49$ ($p = .0002$, $\alpha = .05$). This result indicates that FF appearing more frequently in web-based corpora tend to be associated with higher accuracy scores among student participants. However, the strength of the relationship is

best described as weak to moderate, as the association does not reflect a strong linear dependency (see Figure 1). The coefficient of determination ($R^2 = 0.24$) indicates that approximately 23.8 % of the variance in test scores can be explained by word frequency, leaving 76.2% of the variance attributable to other factors. While statistically significant, the limited explanatory power of frequency constitutes a critical finding, suggesting that word frequency alone is not the dominant factor in mastering FF. This supports the view that informal exposure may be insufficient to resolve the semantic interference inherent in FF, stemming from mismatches between the L1 and L2 lexical mappings. Moreover, exposure through digital media does not guarantee encounter with the specific FF under investigation, given the vast and heterogeneous nature of online content. Consequently, the likelihood of encountering particular FF remains probabilistic rather than systematic.

Following the examination of general trends, the analysis was extended by disaggregating the data according to grammatical categories (nouns, verbs, and adjectives). This approach allows for an assessment of whether informal exposure operates uniformly across the lexicon or whether certain word classes are more sensitive to frequency effects. The analysis of nouns revealed a weak positive correlation ($r = .36$), with a coefficient of determination of $R^2 = .13$, indicating that approximately 13% of the variance in test scores can be attributed to frequency. However, this relationship did not reach statistical significance ($p = .11$, $\alpha = .05$). The corresponding scatter plot (Figure 2) shows a dispersed pattern with limited linearity, suggesting that the observed trend may be attributable to random variation rather than a systematic effect of exposure. In contrast, the analysis of verbs produced a moderately positive correlation ($r = .54$), with frequency accounting for nearly 30% of the variance in test performance ($R^2 = .30$). Despite the relatively stronger association, the correlation did not reach statistical significance ($p = .08$), likely due to the reduced size of the verb subset. Nevertheless, this result suggests that verbs may be more sensitive to environmental exposure than nouns (see Figure 3). The most robust findings emerged from the analysis of adjectives (Figure 4). This category exhibited the strongest correlation ($r = .57$) and was the only grammatical class to reach statistical significance ($p = .009$, $\alpha = .05$). Approximately 32.3% of the variance in test performance was explained by adjective frequency, indicating that adjectives are particularly responsive to input frequency. It is therefore plausible that the significant effects observed in the overall analysis were primarily driven by adjectives, with verbs contributing to a lesser extent.

Overall, these results demonstrate that the effects of informal learning are not uniformly distributed across grammatical categories. While frequency appears largely irrelevant for nouns and remains inconclusive for verbs, it emerges as a reliable predictor for adjectives. One possible explanation is that adjectives receive less explicit instructional focus and may therefore benefit more from repeated exposure to authentic input. By contrast, nouns and verbs are often central to explicit instruction and may already be familiar to learners, reducing the observable impact of frequency-based exposure and increasing variability in the data.

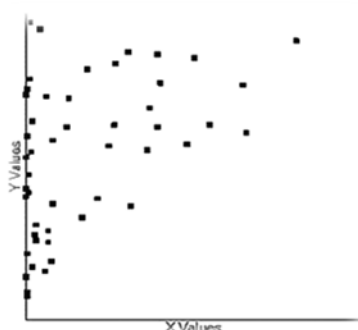


Figure 1. FF frequency and FFVT results

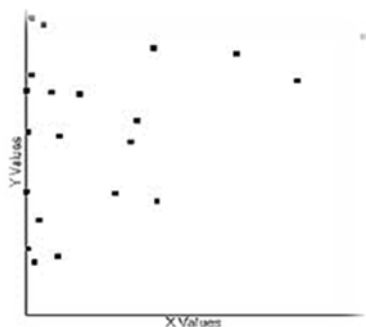


Figure 2. Noun frequency and FFVT results

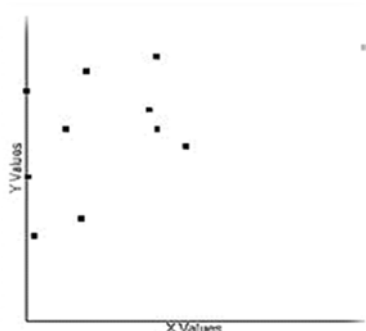


Figure 3. Verb frequency and FFVT results

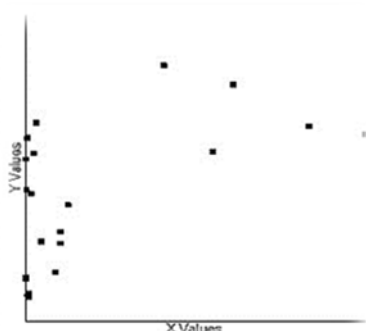


Figure 4. Adjective frequency and FFVT results

3.2 Diversity of Digital Media and FFVT Results

Having identified a positive, albeit moderate, correlation between the web frequency of FFs and test performance, with statistical significance levels for the selected FF as a whole, notable variations across nouns, adjectives, and verbs persisted. Given these findings, and notwithstanding the likely influence of extraneous factors, a more detailed investigation into linguistic exposure to false friends is necessary. The initial phase of this analysis examined whether higher proficiency in FF correlates with the diversity of digital media platforms used and, by extension, the variety of genres encountered.

A Pearson correlation coefficient was computed to assess the relationship between the diversity of media exposure and performance on the FFVT. The analysis revealed a negligible negative correlation between the two variables $r = -.03$, $n = [87]$, $p = .783$.

The coefficient of determination ($R^2 = .0009$) indicates that media variety accounts for only 0.09% of the variance in FFVT scores. Consequently, the remaining 99.91% of the variation is attributable to other factors or inherent

sampling error. Because the p-value ($p = .783$) significantly exceeds the standard alpha level of .05, the null hypothesis cannot be rejected. These results suggest that there is no statistically significant relationship between the breadth of media consumption and proficiency in learning FF. The data points demonstrate a random distribution (see Figure 5), indicating that an increase in media variety does not reliably predict test performance.

3.3 Weekly Frequency of Exposure to Digital Media and FFVT Results

The time spent on digital media can be considered a variable that affects the performance as it offers more chances to meet the FF. A Pearson correlation analysis was conducted to examine the relationship between weekly time spent on media exposure and performance on the FFVT. The results indicated a weak positive correlation between the two variables, $r = .15$, suggesting that higher frequencies of media exposure were associated with only a slight increase in FFVT scores. However, the strength of this association was negligible. The coefficient of determination was low ($R^2 = .02$), indicating that approximately 2.13% of the variance in FFVT performance could be explained by frequency of media exposure, while the remaining variance was likely due to other factors. Importantly, the correlation did not reach statistical significance ($p = .18$), thus, the observed trend may be attributable to random variation within the sample rather than a reliable effect. Although this correlation was marginally stronger than that observed for media variety ($r = -.03$), it nevertheless fails to meet the criteria for statistical significance and does not support the conclusion that increased media exposure frequency is a valid predictor of FFVT outcomes. Visual inspection of the scatterplot (see Figure 6) further corroborated these findings, as both distributions reflected a pattern consistent with the null hypothesis, showing no discernible linear trend.

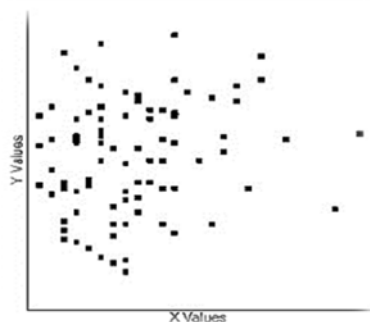


Figure 5. Diversity of media exposure and FFVT Results

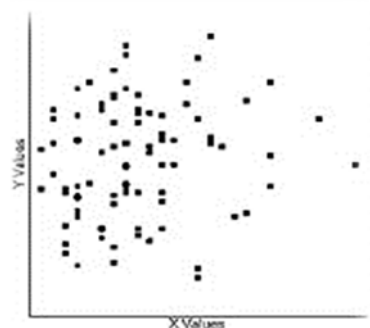


Figure 6. Weekly time of media exposure and FFVT Results

3.4 Weekly Hours of Exposure to Digital Media and FFVT Results

The number of hours spent on digital media each week may have a stronger effect on FF knowledge, as it exposes users to repeated and intensive contact with linguistic input and provides a quantitatively measurable variable that is more reliable than the frequency with which media platforms are accessed (see section 3.3).

A Pearson correlation analysis was conducted to investigate the relationship between time spent on media and performance on the FFVT. The analysis revealed a weak positive correlation between time of exposure and test

outcomes, $r = .16$, indicating that longer exposure was associated with only a minimal increase in FFVT scores. The coefficient of determination was low ($R^2 = .03$), suggesting that approximately 2.5% of the variance in test performance could be attributed to differences in exposure time, while the vast majority of the variance was explained by other factors not captured by this measure (for example, learners' cognitive differences, their motivation and engagement with the content, and the quality of input, such as the contextual saliency of the specific FF encountered). Crucially, the correlation did not reach statistical significance, $p = .14$, consequently, the results do not support the assumption that increased time spent on digital media reliably leads to higher FFVT scores (see Figure 7).

A closer inspection of the data ($N = 87$) helps to explain the weakness and non-significance of this relationship. Several outliers deviated markedly from the expected linear pattern, including participants who achieved high scores with minimal exposure time and others who performed poorly despite extensive exposure. This pattern suggests that learning outcomes in this sample were not linearly related to time spent on digital media and that factors such as prior lexical knowledge, learning efficiency, and the depth of cognitive engagement may have played a more substantial role than duration alone. Moreover, the analysis by exposure-time indicated a non-linear (see Figure 7), threshold-like pattern, with performance improving only after a certain level of exposure and declining again at the highest exposure levels.

Taken together, these results suggest that increasing the amount of media exposure does not necessarily lead to proportional gains in test performance. After a certain point, additional exposure time appears to provide limited or no benefit. Differences in how learners engage with the material, rather than the duration of exposure alone, are likely to play a more important role. Accordingly, the quantity of media exposure by itself is not a reliable predictor of FF test outcomes.

As correlations are suitable for measuring consistent linear directions, but often fail to capture curvilinear relationships where progress reaches a plateau or begins to decline, in addition to Pearson correlations, a series of one-way analyses of variance (ANOVAs) were conducted to examine whether non-linear patterns in the data, such as threshold or plateau effects, might account for the weak linear associations observed. To address this limitation, participants were grouped according to exposure intensity and media use, allowing for a comparison of mean test scores across discrete levels of input quantity. The ANOVA compared the four groups of participants according to the intensity of their exposure to digital media, where group A was the less intensive exposure cohort and group D the most intensive exposure group. Mean scores showed a gradual increase across groups, with the highest average performance observed in the most intensive exposure group D; however, this effect did not reach statistical significance, $F(3, 83) = 1.73$, $p = .17$. Although the descriptive pattern suggests a potential plateau effect, where gains appear to level off beyond approximately 18-22 hours of exposure, the high within-group variability reduced the stability of this trend and prevented the detection of significant mean differences. This outcome is consistent with the correlational findings and indicates that time spent on input, by itself, does not reliably differentiate learner performance.

Parallel ANOVAs were conducted for media variety and weekly digital media use. No significant differences emerged across groups defined by the number of media platforms followed, $F(2, 84) = 0.48$, $p = .62$, nor across light, moderate, and heavy weekly users, $F(2, 84) = 0.89$, $p = .41$. Collectively, these results converge with the correlational analyses in demonstrating that neither exposure duration, usage frequency, nor platform variety significantly predicts FFVT performance. The convergence of non-significant findings across both linear and group-based analyses supports the conclusion that lexical learning in this sample is driven more by individual differences and qualitative aspects of engagement than by the sheer quantity of media-based input.

3.5 Watching Films on Digital Media and FFVT Results

Since videos, reels, and blog posts are generally brief and may lack a sufficiently rich context to disambiguate false friends, a further correlation was established between the frequency of watching English-language films via digital media and performance on the FFVT. Indeed, films offer narrative development that may facilitate both the interpretation and the acquisition of this specific lexicon.

The analysis revealed a small positive correlation between the two variables, $r = .12$. However, the strength of the association was weak, indicating that higher frequency of watching films in English was only minimally associated with higher FFVT scores.

The coefficient of determination was low ($r^2 = .014$), suggesting that approximately 1.4% of the variance in FFVT scores can be explained by frequency of watching films in English. This relationship was not statistically significant, $p = .27$, ($p < .05$). Even these results indicate that, in this sample, frequency of watching films in English was not a meaningful predictor of FFVT performance.

In this instance as well, a brief inspection of the dataset reveals the presence of several outliers that may have influenced the observed weak and non-significant correlation. In particular, these outliers are a small number of participants that reported very high frequencies of watching films in English (e.g., values above 15), yet obtaining only average or below-average FFVT scores. These cases represent high-leverage outliers on the X variable that do not follow the expected positive trend (Figure 8).

Additionally, some outliers were observed at the extremes of the FFVT scores. A few participants achieved very high FFVT scores despite low or minimal exposure to films in English, while others obtained very low FFVT scores across varying levels of exposure, including moderate to high frequencies. Overall, these outliers increase data dispersion and suggest that factors other than film-watching frequency may play a substantial role in FFVT performance, thereby attenuating the strength of the correlation.

The scatter plot of these values (Figure 8) displays a wide dispersion of data points, indicating a weak association between X values (frequency of watching films in English) and Y values (FFVT scores). Although a slight upward tendency can be observed, the overall pattern does not suggest a strong linear relationship. Most data points are clustered toward the lower end of the X-axis, with Y values spreading across a broad range, which suggests substantial variability in FFVT performance among participants with low exposure to films in English.

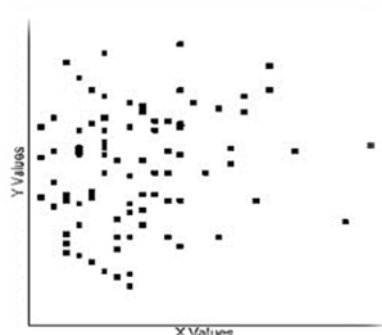


Figure 7. Weekly hours of media exposure and FFVT results

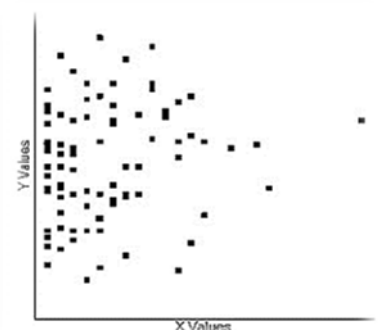


Figure 8. Film-watching frequency and FFVT results

As X values increase, the distribution of Y values remains heterogeneous rather than converging toward higher scores. Several observations at higher X values are associated with only moderate or low Y scores, while some relatively high Y scores appear at low X values. This pattern, consistent with the presence of outliers and high variability, likely attenuates the strength of the correlation. Overall, the analysis indicates a weak and non-significant relationship between the two variables which suggests that frequency of watching films in English alone does not adequately explain variation in FFVT performance. Even when the time spent watching films is cross-referenced to the results of the FFVT the same scenario emerges (see Figure 9). The analysis revealed a weak positive non-significant correlation between the two variables, $r = .10$, $p = .35$ ($p < .05$) indicating a weak association. The coefficient of determination showed that only a small percentage of the variance in FFVT scores

($r^2 = .01$) can be related to the amount of time spent watching films in English.

3.6 Subtitles Usage on Digital Media and FFVT Results

The relationship between the frequency of usage of subtitles (in English or Italian) when watching contents on digital media was related to the results of the test on FF. The analysis revealed a weak negative correlation, $r = -.18$ and showed that higher values on the X variable (subtitle usage) were associated with slightly lower values on the Y variable (FFVT results). However, the strength of this relationship was minimal and did not reach statistical significance, $p = .09$ ($p < .05$). As displayed by Figure 10, the data points are widely dispersed, with no clear linear trend, and appear clustered at lower X values with substantial variability along the Y-axis, confirming that the two variables are not related.

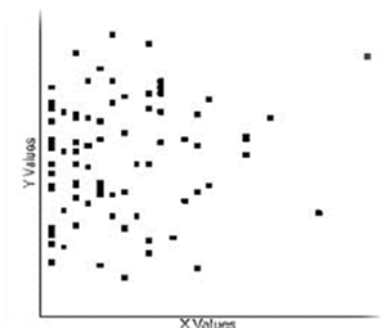


Figure 9. Film-watching time and FFVT results

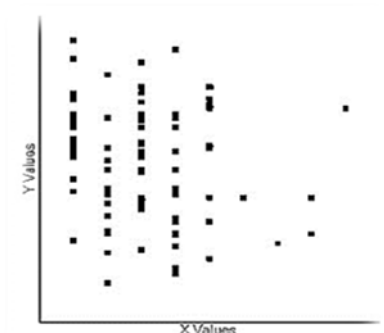


Figure 10. Subtitles usage and FFVT results

3.7 General Vocabulary Knowledge, Proficiency Level in General English and FFVT Results

Correlational analyses were conducted with the aim of identifying possible associations between the diversity and quantity of digital media exposure; however, these analyses did not yield robust results. Instead, only weak correlations were observed, accounting for a minimal proportion of the variance in the data and highlighting the presence of outliers that introduce noise into the analysis. It is therefore likely that individual characteristics play a more substantial role in shaping competence in this specific lexical set. Given that greater general lexical competence may facilitate the comprehension of input encountered on digital media and support the acquisition of FF, an additional correlation was carried out between performance on a general English vocabulary test, the VST, and scores on the FFVT. This analysis was conducted on a subgroup of participants ($N = 69$) who had completed the VST. The correlation revealed a weak but statistically significant positive association between general vocabulary knowledge and FFVT performance, $r = .27$. The coefficient of determination was modest ($r^2 = .07$), indicating that approximately 7.0% of the variance in FFVT scores can be explained by general lexical competence. Despite the limited explanatory power, the relationship reached statistical significance, $p = .03$, ($p < .05$). Overall, these findings suggest that general vocabulary knowledge is related to FF competence, albeit to a limited extent, and that additional factors are likely to contribute to performance in this lexical set.

If general vocabulary knowledge affects performance on the FFVT, overall English proficiency may likewise be related to both lexical breadth and FF competence (see Masi & Noccetti, 2025). To examine these relationships, Spearman's rank-order correlations (Note 5) were conducted for this group of students ($N = 69$), exploring the associations among English proficiency level, ascertained through the OPT, FFVT performance, and VST. English proficiency was operationalised on a seven-point ordinal scale ranging from A1 (1) to C2 (7) (Note 6).

The analysis revealed a moderate positive correlation between English proficiency level and FFVT scores ($\rho = .49$, $p < .001$), indicating that learners with higher overall proficiency tended to show greater competence in FF knowledge. A positive correlation was also found between English proficiency level and general lexical knowledge ($\rho = .32$, $p = .008$), suggesting that increases in global proficiency are accompanied by broader vocabulary knowledge. Crucially, FFVT scores were themselves moderately correlated with VST ($\rho = .41$, $p < .001$), indicating that learners with a richer general lexicon were associated with better performance on the FFVT. These findings indicate that FF competence is not an isolated skill but is embedded within broader dimensions of linguistic knowledge. While the proportion of explained variance remains modest, the results suggest that a well-developed general lexicon may provide the semantic scaffolding necessary to integrate and disambiguate lexically misleading items. In this sense, FF knowledge appears to be more strongly associated with underlying language proficiency and lexical breadth than with informal exposure variables alone, reinforcing the view that individual learner characteristics play a central role in mediating FF acquisition.

4. Discussion

The present study investigated whether exposure to English-language input through digital media facilitates the informal acquisition of FF, a lexical set characterised by cross-linguistic semantic interference. Overall, the findings suggest that digital media exposure, when considered in terms of frequency, duration, and platform variety, exerts only a limited influence on FF competence. While some statistically significant correlations emerged, their explanatory power was consistently modest, indicating that informal exposure alone is insufficient to account for learners' performance in this specific lexical area.

The analysis of FF frequency in web-based corpora revealed a statistically significant positive correlation with test performance, suggesting that more frequently occurring FF (see Table 3) are generally associated with higher levels of accuracy. However, despite its significance, frequency explained less than one quarter of the variance in test scores. This finding highlights a key limitation of frequency-based accounts of lexical learning: while repeated exposure may enhance familiarity, it does not necessarily lead to accurate semantic mapping, especially for FF, where form–meaning correspondences conflict across languages. Moreover, informal exposure neither guarantees that learners will attend to the items nor that they will encounter them. In this respect, FF seem to require learners to actively resolve entrenched L1–L2 interference rather than simply accumulate exposure. The disaggregation of results by grammatical category further refines this interpretation. The fact that adjectives displayed the strongest and only statistically significant correlation suggests that informal learning effects are not uniformly distributed across the lexicon. Adjectives may be particularly amenable to frequency-based learning because they are less systematically taught in formal instructional settings and often acquire meaning through contextual nuance rather than direct translation. In contrast, nouns and verbs, especially high-frequency instructional items, such as *library* and *to attend*, may already be familiar to learners through explicit teaching, reducing the observable impact of additional exposure and introducing greater variability into the data. This uneven sensitivity across word classes highlights the importance of considering lexical categories when assessing informal learning effects. Crucially, variables traditionally associated with input quantity, such as weekly frequency of exposure and hours spent on media platforms, failed to predict FFVT performance. Both correlational and ANOVA-based analyses converged in showing weak, non-significant effects and high within-group variability. These results challenge the assumption that increased exposure necessarily leads to proportional gains in lexical competence. Instead, the data suggest a non-linear learning trajectory, where gains may plateau beyond a certain threshold or even decline at higher exposure levels. Such patterns align with language learning models –most notably Schmidt's (2012) *Noticing Hypothesis* – which emphasise the necessity of attention, deeper processing, and active engagement, over mere exposure duration. Specifically, the data lend support to the idea that the quality of exposure is more consequential than its sheer quantity. The degree of cognitive engagement during viewing, including focussed attention and intrinsic motivation to acquire new lexical items, might be a more reliable predictor of success in developing advanced L2 competence than passive time accumulation.

Similarly, the absence of a relationship between media variety and FF performance suggests that diversification of platforms does not, in itself, enhance lexical learning. While exposure to multiple platforms may increase the quantity of input, it does not guarantee repeated encounters with specific lexical items, nor does it ensure the contextual richness required to disambiguate FF. The random dispersion observed in the scatter plots reinforces

the notion that FF learning depends less on the breadth of input sources and more on how learners interact with and process linguistic material.

The analyses concerning film watching and subtitle usage further support this interpretation. Despite the richer narrative context offered by films, frequency and duration of film exposure showed only weak and non-significant correlations with FF performance. The presence of numerous outliers (high achievers with minimal exposure and low achievers with extensive exposure) suggests that individual differences play a decisive role. Similarly, subtitle usage was weakly and negatively associated with test scores, although not significantly. This may reflect differences in cognitive processing strategies, where reliance on subtitles could reduce attention to lexical form–meaning mapping in the target language rather than facilitate it.

Finally, the most consistent predictor of FF competence was general vocabulary knowledge, which showed a weak but statistically significant positive correlation with FFVT performance. This relationship was further contextualised by the Spearman correlation analysis including English proficiency level, which revealed that FF performance was moderately associated with both overall proficiency and lexical knowledge, highlighting the interdependence of these dimensions. Although the proportion of explained variance remained limited, this result underscores the importance of individual learner characteristics. A richer general lexicon may provide a stronger semantic network, enabling learners to integrate new or conflicting lexical information more effectively. In this sense, FF competence appears to be scaffolded by broader lexical proficiency rather than driven directly by informal exposure to digital media input.

Taken together, these findings suggest that the informal learning of FF through digital media is a constrained and highly variable process. Exposure may contribute to learning under certain conditions, particularly for specific lexical categories, but it is neither sufficient nor uniformly effective. Instead, FF acquisition appears to be mediated by qualitative aspects of engagement, existing lexical knowledge, general proficiency in English, and cognitive factors that extend beyond measurable exposure. These results call for a more nuanced understanding of informal learning and point to the continued relevance of explicit instruction and metalinguistic awareness in addressing persistent areas of lexical interference.

5. Concluding Remarks

This study investigated the possible impact of informal contact with English through digital media on the learning of a selection of FF among a sample of Italian university students. Results show that media exposure – measured in terms of quantitative variables such as frequency and duration, and also including platform diversity, film viewing and subtitle use – does not emerge as a primary predictor of performance, while competence in this lexical set tends to correlate more significantly with overall English proficiency and general lexical knowledge. Additionally, the disaggregation of results by grammatical category of target items highlights uneven sensitivity across word classes, with adjectives (as opposed to nouns and verbs) showing the only statistically significant correlation.

If anything, the non-conclusive findings of research conducted so far emphasise the problematic nature of the items at issue. On the one hand, learning FF involves interference from learners' first language, which may call for deliberate instructional settings that encourage attention on form–meaning mismatches – conditions more typically associated with formal learning environments. On the other hand, the widespread presence of English in contemporary social and digital spaces has multiplied opportunities for informal exposure and for both incidental and intentional learning. In practice, these different learning environments are likely to coexist, and their combined effect may be responsible for greater learning gains (Sok & Han, 2020).

In fact, our results point to the need for further research into the qualitative dimensions of FF learning, employing a more refined methodology and paying greater attention to individual differences and word-class effects. More specifically, the hybrid nature of FF calls for a more restricted and homogeneous set of target items, while their elusive quality and co-dependence on broader dimensions of linguistic knowledge require a revised protocol with fewer but more targeted questions concerning the circumstances and actual content of exposure, in order to better identify individual learning patterns and the possible presence and salience of evidence that may counteract entrenched cross-linguistic interference.

Acknowledgments

Although the present work stems from a joint collaboration, the primary authorship of sections 1, 2, and 5 is attributed to Silvia Masi, with Sabrina Noccetti contributing sections 3 and 4.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*,

- 128(3), 466–478. <https://doi.org/10.1016/j.actpsy.2008.03.014>
- Aliyar, M. (2020). A corpus-based analysis of linguistic interference of Italian native speakers in English and Spanish written production as a foreign language. *Journal of Second Language Teaching and Research*, 8(1), 103–137.
- Amengual-Pizarro, M., & Alonso Alonso, R. (2024). *Informal language learning: Expanding the English learning horizons beyond the classroom*. Bern, Switzerland: Peter Lang. <https://doi.org/10.3726/b22104>
- Anissimova, A. (2015). *Optimization of the teaching-learning process: A case study of “Translator’s false friends”* (pp. 1–7). Proceedings of the 8th International Conference of Education, Research and Innovation, ICERI2015.
- Browne, V., Mendes, E., & Natali, G. (1995). *More and more false friends. Bugs & bugbears. Dizionario di ambigue affinità e tranelli nella traduzione fra inglese e italiano*. Bologna, Italy: Zanichelli.
- Carreres, Á., Noriega-Sánchez, M., & Pintado Gutiérrez, L. (2021). Introduction: Translation and plurilingual approaches to language teaching and learning. *Translation and Translanguaging in Multilingual Contexts*, 7(1), 1–16. <https://doi.org/10.1075/ttmc.00066.int>
- Chacón Beltrán, R. (2006). Towards a typological classification of false friends (Spanish-English). *RESLA Revista Española de Lingüística Aplicada*, 19, 29–39.
- Chamizo-Domínguez, P. J. (2008). *Semantics and pragmatics of false friends*. New York/London: Routledge.
- Chamizo-Domínguez, P. J., & Nerlich, B. (2002). False friends: Their origin and semantics in some selected languages. *Journal of Pragmatics*, 34(12), 1833–1849. [https://doi.org/10.1016/S0378-2166\(02\)00024-3](https://doi.org/10.1016/S0378-2166(02)00024-3)
- Dressman, M. (2020). Introduction. In M. Dressman, & R. W. Sadler (Eds.), *The handbook of informal language learning* (pp. 1–12). Chichester, England: Wiley Blackwell. <https://doi.org/10.1002/9781119472384.ch0>
- Gouws, R., Prinsloo, D. J., & De Schryver, G.-M. (2004). *Friends will be friends: True or false. Lexicographic approaches to the treatment of false friends* (pp. 797–806). Proceedings of the 11th EURALEX International Congress.
- Hill, R. J. (1982). *A dictionary of false friends*. London, England: Macmillan Education.
- Kerr, P. (2014). *Translation and own-language activities*. Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/9781009024815>
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.
- Laufer, B. (1997). What’s in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140–155). Cambridge, England: Cambridge University Press.
- Lee, J. S., & Dressman, M. (2018). When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly*, 52(2), 435–445. <https://doi.org/10.1002/tesq.422>
- Libben, M. (2017). Non-selective language activation and bilingualism as the default mental lexicon. In M. Libben, M. Goral, & G. Libben (Eds.), *Bilingualism* (pp. 103–122). Amsterdam, The Netherlands: John Benjamins. <https://doi.org/10.1075/bpa.6.05lib>
- Marini, E. (2023/24). *False friends: A study of the challenges for Italian students*. Unpublished BA thesis. Università degli Studi di Padova, Italy. Retrieved from https://thesis.unipd.it/retrieve/a0f1b4de-fc35-400c-aff8-ed1b9609ca1d/Marini_Elena.pdf
- Masi, S., & Noccetti, S. (2025). *Getting to know false friends: A case study from the perspective of Italian learners of EFL* (pp. 2388–2397). Proceedings of the 18th Annual International Conference of Education, Research and Innovation, ICERI2025. <https://doi.org/10.21125/iceri.2025.0800>
- Nation, I. S. P. (1997). L1 and L2 use in the classroom: A systematic approach. *TESL Reporter*, 30(2), 19–27.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- O’Neill, M., & Casanovas Catalá, M. (1997). False friends: A historical perspective and present implications for lexical acquisition. *BELLS Barcelona English Language and Literature Studies*, 8, 103–115. <https://raco.cat/index.php/Bells/article/view/102791>
- Pavesi, M. (2024). The informalisation of English language learning. A first national research project. *Lingue e*

Linguaggi, 66, 17–42.

- Pavesi, M., Bianchi, F., Bruti, S., & Zago, R. (2025). *In contact with English: Informal media-driven learning among Italian university students*. Rome, Italy: Carocci.
- Pavesi, M., & Ghia, E. (2020). *Informal contact with English. A case study of Italian postgraduate students*. Pisa, Italy: ETS.
- Roca-Varela, M. L. (2015). English false friends at hand! Learning difficult English words through games. *Docencia e Investigación*, 25(2), 47–65.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209–241. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In W. M. Chan, K. N. Chin, S. Bhatt, & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27–49). Boston/Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9781614510932.27>
- Sok, S., & Han, Z. H. (2020). A study of L2 vocabulary acquisition under incidental and intentional conditions. *VIAL – Vigo International Journal of Applied Linguistics*, 17, 113–140. <https://doi.org/10.35869/vial.v0i17.1467>
- Starreveld, P. A., De Groot, A. M., Rossmark, B. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition*, 17(2), 258–276. <https://doi.org/10.1017/S1366728913000308>
- Swan, M. (2017). *History is not what happened: The case of contrastive analysis. A follow-up to Claire Kramsch's review of Linguistics across cultures*. Retrieved from <https://mikeswan.net/wp-content/uploads/2017/08/History-is-not-what-happe.pdf>
- Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching*, 56(2), 161–180. <https://doi.org/10.1017/S0261444822000507>

Notes

Note 1. Semantic amelioration is a type of semantic change in which a word's meaning becomes more positive, elevated, or socially valued over time. The process is, for instance, discussed by Chamizo-Domínguez (2008, p. 7) in relation to the English noun *topic* and its Spanish cognate *tópico*. While *topic* in English has developed a neutral or positively valued meaning referring to the subject of discourse, Spanish *tópico* has followed a different semantic trajectory, commonly denoting a cliché or overused idea.

Note 2. For example, resources covered glossaries of FF found on several webpages offering English courses. The various meanings of relevant items were compared by reference to dictionaries such as the *OED*, the *Oxford Advanced Learner's Dictionary*, *Collins*, *Webster*, *Macmillan*, and *Treccani* for Italian, alongside Hill (1982) and Browne et al. (1995).

Note 3. *YouGlish.com*, a specialised pronunciation and spoken-language database, indexes over 100 million captioned YouTube videos, providing access to a vast repository of real-world oral discourse. *English Web 2020 (enTenTen20)* corpus via Sketch Engine comprises approximately 36.5 billion words (over 42 billion tokens) sourced from diverse web domains.

Note 4. The Pearson Correlation Coefficient (r) is a bivariate parametric statistic used to measure the strength and direction of a linear relationship between two continuous variables. Ranging from -1 to +1, the coefficient quantifies the degree to which changes in one variable (e.g., word frequency) are associated with proportional changes in another (e.g., test scores). In this study, r serves to determine whether the link between linguistic input and lexical acquisition follows a predictable linear trend, assuming both normality in distribution and the consistency of variance across the variables.

Note 5. Spearman's Rank Correlation Coefficient (ρ) was employed as a non-parametric measure to assess the strength and direction of the rank-ordered relationship between word frequency and test performance. Unlike Pearson's r , Spearman's ρ does not assume normality or linearity, as it operates on the rank order of the data rather than raw values. This approach was selected to mitigate the influence of potential outliers and to account for the non-linear distribution often characteristic of lexical frequency data in digital corpora.

Note 6. The B1+ intermediate level was included alongside the six standard CEFR levels, as it is a mandatory requirement for the degree programmes of some participants in the study.

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).