A corpus-based analysis of linguistic interference of Italian native speakers in English and Spanish written production as a foreign language

Mahnaz Aliyar

Abstract

The purpose of the present study is to investigate the language learning process of Italian native speakers through the analysis of errors found in their English and Spanish written performance. In order to accomplish this goal, essays written by Italian university students were collected, typed and used to build two computer-readable corpora. Furthermore, competence errors retrieved from each corpus were classified according to their sources into intralingual and interlingual errors. Then, interlingual errors were classified according to linguistic levels and form of errors. Finally, the created taxonomies were analysed to simulate, subsequently, the students' interlanguage systems. Thus, by comparing the results of the analysis, and examining the L1 negative transfers encountered in the interlanguage systems, the study has explored whether native speakers of the same L1 will develop the same learning process towards different languages. The results of the research endeavour have demonstrated that L1 has a very strong influence on foreign language acquisition and use. Moreover, the empirical research evidence has shown that in most aspects of students' interlanguage systems, L1 influence is greater than other factors that lead to intralingual errors such as incomplete rule application and overgeneralisation. Specifically, the level of similarity between L1 and TL does not affect the considerable degree of L1 influence.

Keywords: *Second Language Acquisition; Corpus Linguistics; Error Analysis; English; Spanish; Italian*

Introduction

Second language acquisition (SLA) is a relatively young subfield of applied linguistics, and its study has significantly expanded and developed in the past five decades (Gass and Selinker, 2008). Thus, SLA research is a field of linguistic enquiry which is interested in the processes underlying the development of foreign languages among non-native speakers (Myles, 2005). Such processes are influenced by numerous factors, among which native language (L1) can be considered as one of the most influential (Foley and Flynn, 2013). Not surprisingly, many researchers have addressed the relationship between L1 and the acquisition of a foreign language, and the role of L1 in second language (L2) learning has extensively been discussed (see Ellis, 2015; Krashen, 1981; Mitchell et al., 2004).

Despite the youth of SLA research, great efforts have been made to understand and describe learners' interlanguage, that is the linguistic system used by L2 learners within the process of language acquisition of a target language (TL) (Selinker, 1972, 214). Central to the notion of interlanguage is L1 transfer which its role has been one of the most controversial issues in SLA research. According to Selinker, L1 transfer is a major cognitive process in L2 learning. He differentiates between positive transfer, that is the processes whereby L1 knowledge facilitates the acquisition of a TL, and negative transfer, i.e. the processes whereby L1 knowledge negatively impact L2 learning (Selinker, 1983). As an example with relation to L1 transfer, an Italian native speaker who is learning Spanish might incorrectly produce:

*Hoy la situación ha completamente cambiado.

In this example, the L1 word order interferes with L2 rules; thus, L1 transfer results in something incorrect. This incorrect utterance is known as negative transfer. In this regard, it is worth mentioning that, while L1 transfer helps shape interlanguage rules, it is not the only cognitive process involved in the development of interlanguage. Factors such as overgeneralization of target language rules, literacy level and social context also do influence the development of interlanguage (Tarone, 2018). Nonetheless, the discussion of these factors goes beyond the purpose of this paper (see Bayley et al., 1996; Gass et al., 1989; Han et al., 2014; Tarone et al., 2006).

Furthermore, error analysis, an area of research tied in with SLA research, is concerned with the analysis of erroneous structures made by language learners in the target language. The current research, thus, is grounded in the field of SLA and relies mainly on error analysis theories. In particular, the study makes use of theories developed in the field of error analysis, such as those involved with the classification of errors and investigation of their causes (Bussmann, 1996; Dulay et al., 1982; Granger, 2003). Furthermore, current research can also be contextualised in the area of modern contrastive linguistics considering it will carry out a comparison between students' English/Spanish written production with their mother tongue, that is Italian.

With regard to L2 writing, first language transfer is interpreted both as a learning device and as a strategy to solve communication problems. Specifically, L2 learners may resort to their L1 to compensate for their deficiencies in the L2 knowledge. According to SLA researchers, when L2 learners attempt to compose a written piece, they might use transfer as a means to convey their meaning and their complex ideas. Thus, errors may occur if the learner

inappropriately transfers a linguistic form from L1 to the TL (see Gilquin, and Paquot, 2008; Manchón, 2012; Ramón Torrijos, 2009; Wolfersberger, 2003).

Thus, this research sets out to explore the process of foreign language learning of Italian native speakers. In essence, the aim of the present study is to investigate whether native speakers of the same L1 will develop the same learning process towards different languages. Relying on a corpus-based analysis, the research strives to provide solid empirical support regarding L1 influence on L2 development. By employing methods of error analysis, the study will identify patterns of interference errors in Italian students' English and Spanish written performance. The overall goals of this research, thus, are to uncover the frequency and nature of a wide range of interlanguage features and to yield more insights into foreign language development and use.

Materials and Methods

Steps of the study

In order to investigate the study hypothesis, the research will go through the following steps:

Data collection and the creation of corpora

Identification of errors caused by L1 influence

Classification of interference errors

Examination and comparison of the created error taxonomies

Exploration of interlanguage systems to grasp their predominant features

Evaluation of the extent of L1 influence on TL production

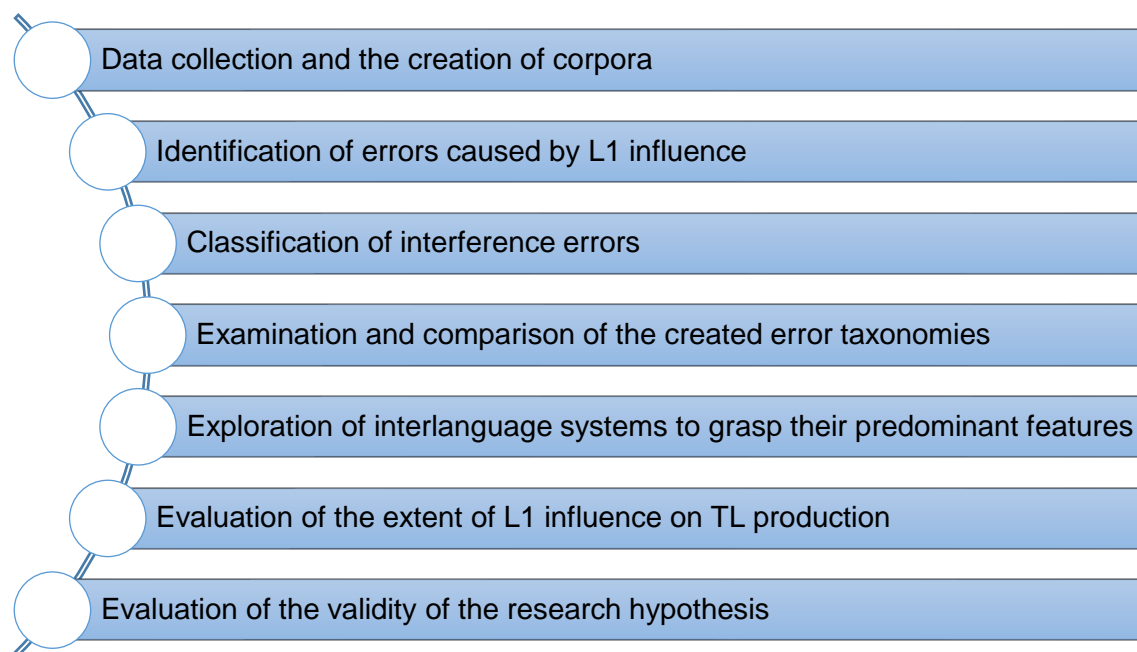Evaluation of the validity of the research hypothesis

Figure 1. *Steps of the research endeavour*

As outlined in Figure 1, the study has firstly collected the data, that is 100 English/Spanish essays written by 93 Italian university students and has created two learner corpora. Then,

errors caused by the L1 influence have been identified with the aid of corpus-based error analysis. Interference errors have been then classified according to different linguistic criteria, namely, cause of errors, linguistic levels and form of errors. Then, through the examination of the created error taxonomies, students' interlanguage systems have been simulated, and their prevailing characteristics have been identified. Lastly, by analysing instances of L1 negative transfer, the extent of the L1 negative influence on the TL production has been evaluated and conclusions have been drawn from the acquired results. These steps will be meticulously presented in the following parts.

Data Collection

Analysing language production of advanced students is one of the key components of second language research which serves to reveal persistent interlanguage errors. As stated previously, this study aims to investigate the process of language learning of Italian native speakers through the analysis of students' errors when writing in a foreign language. Thus, the subjects of this study were Italian students of the Master's degree in "*Lingua, Società e Comunicazione*" of the University of Bologna. The sample consists of 93 Italian students (7 students are either present in both groups or have taken the exam twice) who are mostly in their 20s and are approximately 10% male and 90% female.

These students were chosen due to their high level of language proficiency required by course regulations. To be specific, students' knowledge of both English and Spanish should be level C1 according to the European framework upon entry, and C2 at the end of their degree course. It is to be illustrated that, students' language proficiency level has been proved prior to the admission to the programme either by having passed certain university credits or by providing a recognised international language certificate.

Furthermore, the study has taken into account the "History of the English Language" and the "Spanish Linguistics 1" courses held in 2012-2013 and 2013-2014 academic years at the University of Bologna. The data consist of 100 essays of approximately 500-600 words each which are of different types, including argumentation and discussion aimed to analyse and evaluate a range of given topics. More precisely, a total of 50 essays for each target language, i.e. English and Spanish, written by the students during their language course final exams were collected and typed. The essays were then used to build two computer-readable corpora (one for each language). The English corpus of essays consists of 27,551 tokens and the Spanish corpus includes 29,116 tokens.

The essays produced for exams were selected for the compilation of the corpora firstly to ensure equal conditions for all the students. This means that all essays were supervised and timed (90 minutes allowed), and students did not have access to reference materials except for monolingual dictionaries (in the case of English groups). Secondly, students were given specific topics to write about. This given topics, thus, made it possible to obtain essays which represent the same level of formality i.e. formal academic essays.

Learner Corpora and Interlanguage Development Research

It has been proved, over more than thirty years of research in corpus linguistics, that learner corpora are of great help for studying learners' interlanguage (Granger et al., 2013; Hyland et al., 2012). Furthermore, the fact that learner corpora are considerably beneficial to foreign language acquisition research is now widely acknowledged (Aijmer, 2009; Granger et al., 2015). According to Granger, learner corpora are greatly beneficial for both researchers and learners. First of all, they are useful for researchers considering that a high number of SLA studies are aimed to determine learners' interlanguage knowledge as well as to describe how this interlanguage system develops over time. Secondly, a detailed description of learner errors, deduced from learner corpora, can aid learners' L2 development by helping them to achieve a high level of accuracy in the language (Granger, 2003).

The development of recent corpus linguistics tools such as concordancers, parsers and NLP (Natural Language Processing) toolkits has revolutionised SLA studies. For instance, corpus annotation, that is associating linguistic information such as word classes, grammatical structure, semantic and pragmatic features with the content of a corpus, is considered an essential part of learner data analysis. Hence, many corpus annotation tools have been developed in recent years (AnCoraPipe; Atomic; UAM Corpus Tool to name just a few). Moreover, NLP, which is concerned with the automated processing of human language, is now widely used to analyse learner language (Meurers, 2012).

Accordingly, the present study made use of AntConc (Anthony, 2018), a toolkit for concordancing and text analysis, to analyse the corpora in order to find the occurrence and the frequency of language errors. This software displays the data in keyword-in-context concordance lines presenting the searched word/phrase as a node at the centre of the screen, with a few words before and after it. In addition, TagAnt, a Part-Of-Speech tagger (Anthony, 2016), was used in order to build the annotated corpora according to word classes, which facilitated the process of analysis.

Furthermore, the Uppsala Student English corpus, USE, (Axelsson and Berglund, 1999-2001) and the Corpus Escrito del Español Como L2, CEDEL2, (Lozano and Georgopoulos, 2006-2016) were used as the reference corpora in order to generate keyword lists. The USE is a publicly available corpus compiled by Margareta Westergren Axelsson and Ylva Berglund in 1999-2001. It can be accessed on the Internet from the [Oxford Text Archive](#) and can be used for research and educational purposes. The USE corpus consists of 1,489 essays written by Swedish university students of English, and contains a total number of 1,221,256 words. The aim of the creation of the USE was to build a powerful tool for research into the process and results of foreign language teaching and acquisition, as manifest in the written English of Swedish university students (Axelsson, 2003). Thus, this corpus has so far been used in many studies (e.g. Berglund and Oliver, 2002; Pravec, 2002).

The Corpus Escrito del Español como L2 is a written L1 English/Greek - L2 Spanish corpus compiled by Cristóbal Lozano and Athanasois Georgopoulos. The CEDEL2 data collection started in 2006 and is still continuing. Its primary aim was to investigate the role of the interfaces (lexicon-syntax and syntax-discourse) as a potential source of observed deficit in the development of learners' interlanguage grammars. This corpus is also a relatively large learner corpus of Spanish including 571,448 words (excluding the Spanish natives subcorpus) coming from 1,782 English and Greek native speakers who are learners of the Spanish language (Lozano and Mendikoetxea, 2013). The CEDEL2 is publicly available [online](#) and many studies have hitherto used it as a source of data (e.g. Lozano, 2016; Vázquez Veiga, 2016).

Although the two reference corpora consist of written production from learners representing different mother tongue backgrounds i.e. Swedish, Finnish, English and Greek, they are still suitable to generate keyword lists, considering they are large-scale corpora that contain a wide range of topics. Both reference corpora cover topics of different types, which ensure the inclusion of a wide array of linguistic structures and lexical items which facilitate the comparison of the interlanguage systems. More specifically, the fact that none of the most frequent interlanguage errors committed by Italian students stands in a high rank in the keyword lists (neither English nor Spanish corpus) can be evidence that these errors are peculiar to Italian students. However, it is to be specified that, CEDEL2 consists of compositions rather than essays which can arguably affect results of the comparison. Nonetheless, apart from this limitation, CEDEL2 suits the purpose of this study.

Data Analysis

The Standard Deviation

In order to start the analysis from a general point of view, the standard deviation (SD) of sentence length was assessed. In other words, the deviation of sentence length from the mean (expected value i.e. average sentence length) was calculated for each set of data. Then the results were analysed in order to figure out whether the students' interlanguage system is structurally affected by the L1 i.e. Italian. For example, a higher deviation of sentence length from the English standard average sentence length can be explained in the light of the L1 influence on learners' language use. Italian learners of English, under the influence of the rules governing their L1 sentence production, might produce relatively longer sentences in English. This deviation can support the influence of L1 on interlanguage as it evidences the manifestation of L1 structure in L2 production.

Research Approach

Combining qualitative and quantitative forms of research is nowadays a common practice in SLA research. It is largely acknowledged that "a mixed methods inquiry offers a potentially more comprehensive means of legitimizing findings than do either QUAL or QUAN methods alone by allowing investigators to assess information from both data types" (Dörnyei, 2007).

Accordingly, this research has integrated both quantitative and qualitative analyses in order to bring out the best of both paradigms. By using qualitative research exploratory nature, the study has evaluated learners' interlanguage and described the types of errors. Furthermore, involving quantitative method has served to find the number and the frequency of errors as well as to identify the most significant types of errors according to the purpose of the study. In particular, the current study made use of qualitative analysis to focus on in-depth investigations of interlanguage errors in order to reveal the factors that underlie learners' language use by observation, description and explanation of naturally occurring language; whereas quantitative analysis was used to evaluate learners' language use and to examine the validity of qualitative findings.

Moreover, given the purpose of this research, the study integrated a 'deductive approach' (Myles, 2007) towards the learner corpora in order to explore the data to test the study hypothesis about the nature of learners' language acquisition process. In specific, a deductive-based research investigates learner corpora to (dis)confirm a specific hypothesis about the nature of learner language and to answer questions generated through introspection, SLA

theories or as a result of another analysis of experimental data (Barlow, 2005). Accordingly, by analysing the interlanguage errors retrieved from the corpora and comparing the results of the analysis obtained from the two sets of data, the current study will examine its research hypothesis.

In terms of the steps of the error analysis, to conduct an adequate enquiry, a series of stages is suggested in the literature. In this regard, Corder (1974) proposes the following stages:
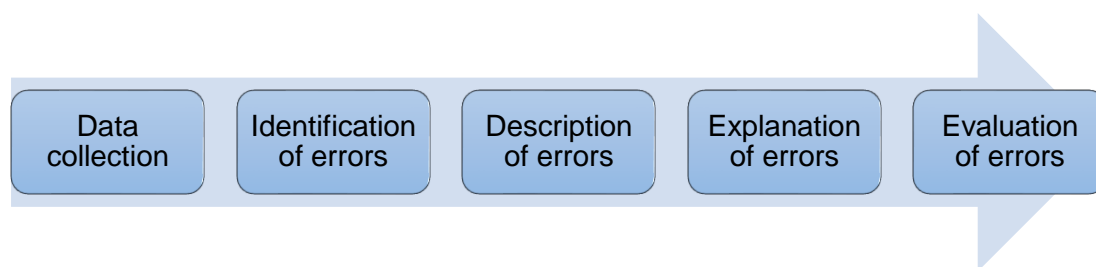

Figure 2. *Corder's five stages of error analysis (1974)*

Thus, the sequence of steps implemented in this study, began with collection of a sample of learner language and identification of the errors in their context, followed by classification and description of the errors. As for explanation, considering the purpose of the research, sources of errors were studied, and error types were explored. The findings were then evaluated in order to analyse the significance of errors from an SLA perspective.

Towards a Classification of Learners' Errors

Language learners' errors reflect a significant aspect of the learning process. They are considered of great importance in order to gain insights into the processes that govern SLA (James, 1998; Myles, 2005). Moreover, reflecting on erroneous structures present in learners language allows drawing conclusions on the nature and process of language production in general. It is, thus, of no surprise that error analysis has always been among the main subjects of SLA studies. In this regard, efforts have been directed towards providing an error taxonomy of satisfactory depth, and many scholars have proposed various classifications considering errors from different perspectives.

Many scholars are in favour of integrating various error classifications in order to achieve multi-dimensional taxonomies which allow conducting deeper levels of analysis (Granger, 2003; James, 1998). Accordingly, due to the aim of the present research, the study carried out error analysis through three kinds of error taxonomies. Since the study concentrates on the influence of L1 on the acquisition of foreign languages, only interlingual/interference errors (James, 1998), i.e. the incorporation of features or elements of the L1 into the L2 leading to

erroneous items, are considered. In order to achieve maximum objectivity and precision in the identification of interlingual errors, two major paradigms were taken into account: first, transfer errors, that are errors in the TL reflecting the L1 structure; second, literal translation, i.e. errors committed by learners in the L2 as a result of a word-for-word translation of their L1 phrase/sentence or expression into the TL.

In addition, the linguistic categories (considering particular linguistic constituents i.e. noun, verb, adjective, etc.) and the classification based on surface structure alteration (e.g. omission, addition, misformation and misordering) proposed by Dulay et al. (1982) are employed in order to conduct further analyses.

Non-native-like language production can be referred to by various terms, including mistake, that is an error of performance, and error i.e. an error of competence (Corder, 1981). In addition, Corder (1981) also distinguishes between "covert" and "overt". He explains that "covert idiosyncratic" errors are thoroughly grammatical at sentence level; however, they cannot be interpreted within their context. "Overt idiosyncratic" errors, in contrast, are ungrammatical at the sentence level. An overt error is easy to identify as there is a clear deviation in form. A covert error, however, occurs in utterances that are apparently grammatical, but which are incorrect in the context:

> *Finally in the 19th and the 20th centuries it reached respectively economical and cultural power. For these reasons, it must be assumed that, thanks to its glorious history, it has achieved a great success all over the world becoming the most influential language.

In this script taken from the English corpus, *economical* is an overt error as it is an incorrect form. The second sentence is superficially well-formed, nevertheless it is an instance of a covert error, as the utterance does not mean what the learner intended it to mean, i.e. *it can be assumed* […].

The same goes for the following Spanish script:

> *[…] los hispanohablantes en EEUU que dejan el español, la lengua de herencia, y abrazan la lengua inglesa, porque es vista como la llave por el éxito, la lengua de la economía y de la política, así hay el abandono del español.

In this example, *por* is an overt error since it is ungrammatical at the sentence level. *Hay*, on the contrary, is a covert error, because it is apparently correct but does not have a proper meaning in the context, i.e. *se produce*.

Therefore, the first step towards error classification was to collect "overt" errors. Then, the study excludes errors of performance and just accounts for all the errors of competence in order to obtain the statistical data. Finally, only interlingual errors are considered for qualitative analysis. The research target errors are shown schematically in figure 3.
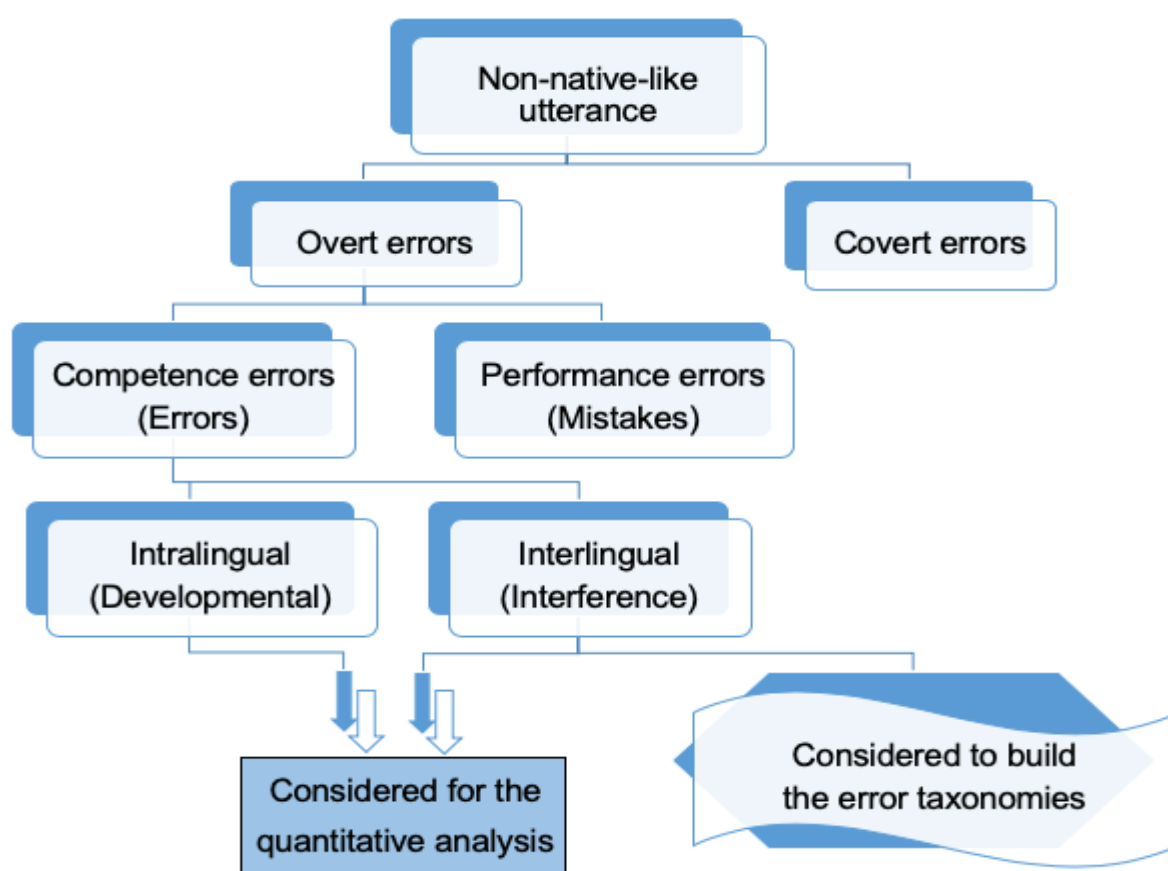

Figure 3. *The research target errors*

Examples of different types of errors are as follows:

> *We have seen how difficult is to enter in the global market for the countries from Expanding Circle. Some aspects linked to language, different ways of living, different social and cultural problems, make a possible growth, impossible.

In this utterance, the omission of the pronoun *it* (how difficult … is), the misuse of the preposition *in* (to enter <u>in</u> the global market), and the misuse of the indefinite article *a* (<u>a</u> possible growth) are examples of competence errors, i.e. noticeable deviations from the grammar which reflects the competence of the learner (Brown, 2000). Differently, the missing *the* (<u>…</u> Expanding Circle) is an example of performance errors that is a failure to utilise a known system correctly (Brown, 2000). It can be considered a 'slip' considering the learner has used *the* correctly in the same phrase throughout the whole essay. In addition, the misused *in* and the omitted *it* are instances of interlingual errors as the former is caused by L1 interference and the latter reflects L1 structure. The misused *a*, on the other hand, is an example of intralingual errors, as it results from learner's lack of knowledge of the TL rule, i.e. *growth* is an uncountable noun.

> *De todas maneras, Manuel Alvar ha recientemente demostrado que esta variable depende más del estilo de vida de una persona. Otros estudios sociolingüísticos han definido el habla de las mujeres conservado y sensible, mientras que él de los hombres sería independiente, competitivo y jerárquico. Lo que es seguro, es que las mujeres se acercan más a la norma prestigiosa y al prestigio abierto; en contras, los hombres están asociados con el vernáculo y el prestigio encubierto.

Similarly, in this part of an essay taken from the Spanish corpus, the present word order error (ha recientemente demostrado), and the *en contras* incorrect form are instances of competence errors. While the misuse of *él* instead of the correct form *el* is a performance error since the student has shown a complete understanding of the L2 rules underlying the use of *el* and *él* as well as their differences throughout the essay. Moreover, **ha recientemente demostrado* is an interlingual error caused by the L1 influence; whereas **en contras* is an intralingual error which represents structures of neither the L1 nor the L2.

In order to facilitate the comparison between the English and the Spanish error taxonomies, interference error rates were computed. More specifically, the total number of competence errors and interference errors was firstly counted within each linguistic category of each set of data. Then, the ratio of interference errors to competence errors was calculated separately within the English and the Spanish linguistic categories. The interference error rates per se could determine the extent of the L1 influence on L2 learning and use. In addition, by comparing the English error rates with the Spanish error rates, the study endeavours to evaluate such influence according to different target languages.

Results and Discussion

A General Overview of the Results

The current research set out to study the learning process of ninety-three university students, all native speakers of Italian, regarding different target languages. To do so, the research has explored English and Spanish essays written by a group of Italian university students for evidence of errors in general, and L1 negative transfers to the TL in particular. In light of the research materials and methods outlined in the Materials and Methods section, the results of the error analysis will be presented and discussed.

General Findings

To start the discussion from a general point of view, the SD of sentence length will be presented. The deviation of sentence length from the average sentence length calculated for each set of data shows 1.06 SD for the English essays and 0.06 SD for the Spanish essays.

According to the results, it can be said that Italian students tend to write relatively longer sentences compared to the English standard average sentence length. While in English short sentences are recommended, in Italian long sentences are normally produced and generally accepted. Hence, 1.06 SD for English written production is not surprising since it reveals these students' writing habits.

Thus, very small SD for the Spanish data occurs because the standard sentence length for both Italian and Spanish is approximately the same. Therefore, it can be deduced that the students' L2 production is structurally affected by their L1. Furthermore, after classifying interlanguage errors retrieved from the corpora according to criteria described in the Meterials and Methods section, the ratio of transfer errors to competence errors was computed within each linguistic category. The results are shown in Table 1 below:

Table 1: *Interference error rates*
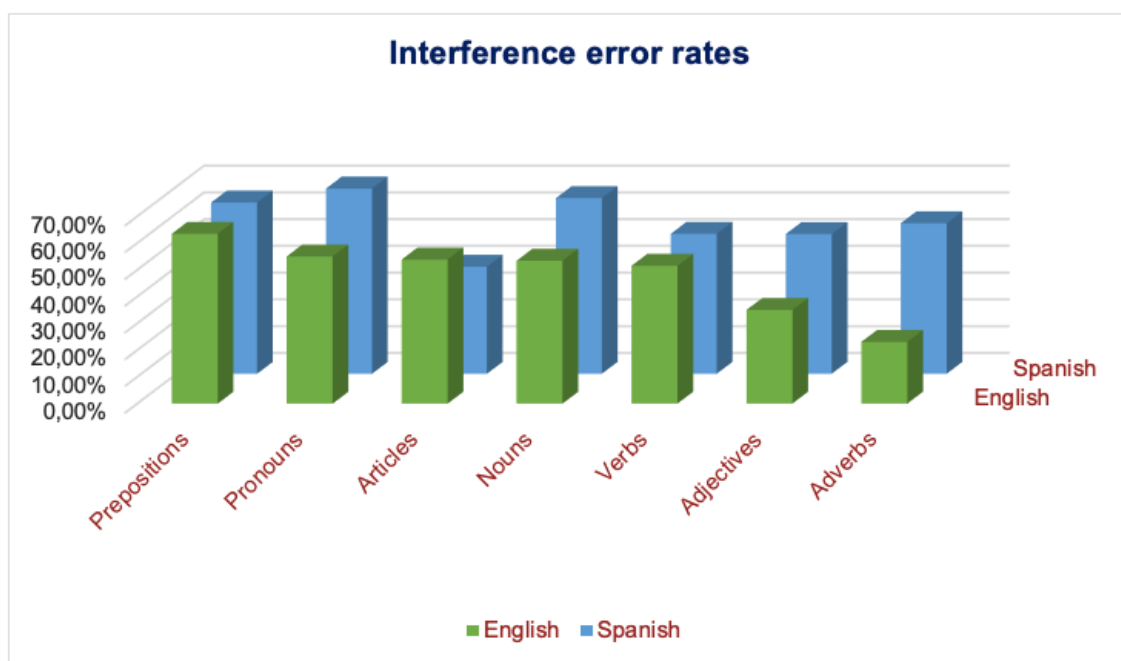
	Prepositions	Pronouns	Articles	Nouns	Verbs	Adjectives	Adverbs
English	63.46%	55%	53.84%	53.42%	51.51%	35%	23.07%
Spanish	64%	69.23%	40%	65.65%	52.32%	52.23%	56.25%

A general overview of the results demonstrates that for both target languages, most of the competence errors are caused by the influence of the L1. In fact, Adjective, Adverb (in English language production) and Article (in Spanish language production) are the only linguistic categories in which intralingual errors outnumber interlingual errors, with 35%, 23.07% and 40% interference error rates respectively.

These results confirm the significant role of the mother tongue in the acquisition and the use of foreign languages, regardless of how close the linguistic systems of L1 and TL are. In this regard, the high percentage of interference error rates for the Spanish set of data is clear evidence to prove such a role in case of close languages. Accordingly, it can be deduced that the apparent similarity between the Italian and the Spanish linguistic systems leads to a series of errors caused by the distinct and complex nature of TL, i.e. the subtle nuances of meaning and the diversity between TL and L1 rules. Moreover, these numbers reveal how the negative influence of L1 is persistent and remarkable even concerning high-proficiency learners.

Spelling Errors

On the basis of corpus-based interlanguage analysis, the study has identified specific linguistic features and has obtained a comprehensive description of Italian learners' interlanguage. The overall result of the study shows a total number of 746 errors in the English set of data and 783 errors in the Spanish set of data (without counting orthographical, spacing and stylistic errors). In addition, in the Spanish corpus, a total of 937 spelling errors has been identified, of which 346 errors are due to faulty use of diacritical marks, while the English corpus contains a smaller number of spelling errors, that is 415 errors.

The overall higher number of errors in the Spanish corpus can have different explanations. Firstly, since the Spanish corpus, compared to the English corpus, contains a larger number of word types (4331 vs. 2485), a higher number of errors can be expected, especially in the case of spelling errors.

Secondly, the structural similarities between the Italian and the Spanish linguistic systems induce many overgeneralisation errors (Bailini, 2016). In other words, students tend to extend the application of language rules to other contexts where such rules should not be applied. For instance, there are many words which differ from Italian to Spanish through the substitution of the phoneme /t/ for phoneme /d/ (i.e. *Lato* (it) vs. *Lado* (es)); however, this cannot be regarded as a general rule since there are many words which violate this scheme. However, instances of a faulty application of this framework are numerous in the Spanish written performance of the Italian learners. *tradado*, *esdándar*, *desdinado*, *endiende* and *inmediados* are a few examples of this overgeneralisation error encountered in the Spanish corpus.

Lexis

The common origin shared by Italian and Spanish, that is the Latin language, has led to many errors provoked by so called 'false friends', i.e. words that look similar but differ significantly in meaning. Instances of persistent incorporation of incorrect linguistic forms into the TL are numerous in the Spanish corpus. A few examples are as follows:

Table 2: *Some examples of Spanish false friends*

Erroneous form	Correct form
subir e.g. es la importante simplificación que ha <u>subido</u> e.g. desde la conquista del América, el español ha <u>subido</u> muchos cambios lingüísticos	**experimentar/sufrir** e.g. es la importante simplificación que ha <u>experimentado</u> e.g. desde la conquista de América, el español ha <u>sufrido</u> muchos cambios lingüísticos
suportar e.g. a veces <u>suportado</u> por elementos no lingüísticos e.g. los hechos tienen que ser <u>suportados</u>	**sostener** e.g. a veces <u>sostenido</u> por elementos no lingüísticos e.g. los hechos tienen que ser <u>sostenidos</u>
exprimir e.g. <u>exprimir</u> las propias opiniones sobre un asunto e.g. se <u>exprimía</u> por medio de un antilenguaje	**expresar** e.g. <u>expresar</u> las propias opiniones sobre un asunto e.g. se <u>expresaba</u> por medio de un antilenguaje
haber e.g. el hecho de <u>haber</u> una propria historia e.g. los jóvenes <u>han</u> la tendencia a hablar con un lenguaje particular	**tener** e.g. el hecho de <u>tener</u> una propia historia e.g. los jóvenes <u>tienen</u> tendencia a hablar con un lenguaje particular
utilizo e.g. es frecuente el <u>utilizo</u> de vulgarismos e.g. el <u>utilizo</u> de conectores argumentativos	**uso** e.g. es frecuente el <u>uso</u> de vulgarismos e.g. el <u>uso</u> de conectores argumentativos

False friends, as the name itself implies, have appeared as a result of faulty incorporation of rather similar words which have a different meaning in each language. *Subido*, for instance, is the past participle of the verb *subir* which means to go up/to raise; in Italian, however, the verb *subire* means to experience/to suffer. Thus, the very similar forms of these two verbs has led the learners to make such an error. The same applies to *suportar* (to bear) and *supportare* (to support) as well as *exprimir* (to squeeze) and *esprimere* (to express). Indeed, the case of *haber* is rather divergent. *Haber* and the Italian verb *avere* have the same etymology but their meaning and use have evolved differently in both languages. In Italian, it functions both as an auxiliary verb and a transitive verb meaning to have/to possess. In Spanish, however, it has lost the latter function and is mainly used as an auxiliary verb to form tenses. Therefore, this

divergence has led Italian students to use incorrectly *haber* instead of *tener* or *poseer*. As for the last example, the word *utilizo* which is not even a noun, but the first-person singular of the verb *utilizar* (to utilise/to use) conjugated in the present indicative, is inappropriately employed in the TL due to its resemblance to the Italian noun *utilizzo* (utilisation/use).

Word Order

Word order errors found in the TL production, which reflect the structure of the Italian language, are further manifestation of the L1 influence on TL use (see Bailini, 2016; Calvi, 2003; Pierini, 2012; Salmasi, 2001; San Vicente et al., 2014, for discussions on contrastive analysis). Nevertheless, it is not surprising that the number of word order errors is more noticeable in English production (21 vs 4). This happens, firstly, because in English the meaning of a sentence is mainly dependent on the order of its words, whereas both Italian and Spanish allow greater variation in word order. Secondly, the sentence structure of Italian mostly coincides with that of Spanish and only varies slightly in some minor structures. The following are some of the examples extracted from the corpora:

Table 3: *Some examples of word order errors from the English corpus*

Erroneous form	Correct form
Misordering	
Besides a foreign language, nowadays for education is fondamental the informatic competence.	Nowadays, besides a foreign language, the informatic competence is fundamental for education.
We can take as example Singlish.	We can take Singlish as an example.
Non-native english speakers learn at school grammar, pronunciation and standard vocabulary.	Non-native English speakers learn grammar, pronunciation and standard vocabulary at school.
It is spoken when occurs a contact between people of different mother tongue […].	It is spoken when a contact between people of different mother tongues occurs […].

The examples reported in Table 3 show some of the processes underlying the students' word order errors. Although Italian and English share some similarities including definite and indefinite articles, singular and plural forms of nouns, perfect and progressive verb forms, there are some features that are completely different between these two languages. For

instance, Italian is characterised by its flexible word-order patterns. Thus, in some structures there are different word-order possibilities: in some cases, these alternative positionings do not carry different semantic values but indicate stylistic features (the first example in the table above); whereas in others, word order varies depending on the word on which emphasis is going to be placed (Singlish in the example above). In addition, the positioning of subject, adjectives and adverbs has proved to be problematic for the Italian learners (the last two examples above).

Table 4: *Some examples of word order errors from the Spanish corpus*

Erroneous form	Correct form
Misordering	
Hoy la situación ha completamente cambiado.	Hoy la situación ha cambiado completamente.
[…] ha recientemente demostrado como los hablantes negros y blancos, en condición de igualdad, no presentan diferencias en el habla.	[…] ha demostrado recientemente cómo los hablantes negros y blancos, en condición de igualdad, no presentan diferencias en el habla.
[…] hay el difundido temor que el "Expanding English" pueda "sofocar" no solo la lengua española […].	[…] hay un temor difuso que el "Expanding English" pueda "sofocar" no solo la lengua española […].
[…] y al final afierma que hay el uso de particulares rasgos como: […].	[…] y al final, afirma que se da un uso de rasgos particulares como: […].

With regard to word order, Italian and Spanish have numerous features in common; however, they are not completely devoid of dissimilarities. The first difference to be noted is regarding the structure of compound tenses, i.e. present perfect, past perfect, conditional perfect, etc. While in Italian auxiliary and main verbs are allowed to be split by adverbs, Spanish does not accept the division of these, that is the auxiliary verb must be followed only by the past participle of the main verb. This difference could lead to word order errors such as the aforementioned first two examples in which the adverbs *completamente* and *recientemente* are mispositioned.

Furthermore, in Italian, the adjective can be positioned before or after the noun to which it

refers, whereas in Spanish the adjective is normally placed after the noun. Therefore, errors such as the last two examples happen in the Spanish language production of Italian learners as a result of their L1 influence.

So far, the findings have confirmed that students' interlanguage is structurally affected by their L1. Moreover, the explained results have demonstrated remarkable consistency with the study hypothesis. In the following part of this section, instances of the most frequent and significant negative transfers retrieved from the corpora will be presented and discussed in order to confirm this hypothesis.

Linguistic Categories

Prepositions

To present the findings of the investigation in more details, the first category of errors to be discussed is preposition errors since it holds the highest interference error rate for the English language production and the third position for the Spanish language production (see Table 1).

Generally, using prepositions correctly is one of the most challenging aspects of learning a new language, since the use of prepositions varies considerably between different languages. Despite the fact that the choice of prepositions does not closely coincide within different languages, language learners tend to translate prepositions from their mother tongue to TL regardless of the degree of similarity between TL and L1 rules. Thus, the subjects of this study are no exception, even if they possess a high level of language proficiency. The most numerous errors are those related to the misformation, that is the incorrect selection of prepositions, followed by addition and finally omission errors:

Table 5. *Some examples of preposition errors from the English corpus*

Erroneous form	Correct form
Omission	
listen people speaking English	listen to people speaking English
Addition	
almost 1 billion of people	almost 1 billion people
to enter in the global market	to enter the global market

D. Cristal answers <u>to</u> this question	D. Crystal answers this question
Misformation	
the most powerful economy <u>of</u> the world	the most powerful economy <u>in</u> the world
good proficiency <u>of</u> Standard English	good proficiency <u>in</u> Standard English
the main reasons <u>of</u> its success are	the main reasons <u>for</u> its success are
the increasing need <u>of</u> business interpreters	the increasing need <u>for</u> business interpreters
a threat <u>for</u> mutual intelligibility	a threat <u>to</u> mutual intelligibility
English was imposed by colonizers <u>to</u> local population	English was imposed by colonizers <u>on</u> the local population

Table 6. *Some examples of preposition errors from the Spanish corpus*

Erroneous form	Correct form
Omission	
hablar viendo el propio interlocutor	hablar viendo (mirando) <u>al</u> propio interlocutor
es difícil encontrar alguien	es difícil encontrar <u>a</u> alguien
Addition	
se pueden enumerar <u>en</u> seis	se pueden enumerar seis
impone <u>de</u> ser	impone ser
nos permite <u>de</u> recibir noticias	nos permite recibir noticias
Misformation	
hablar entorno <u>de</u> un tema	hablar en torno <u>a</u> un tema
las personas que partecipan <u>al</u> debate	las personas que participan <u>en</u> el debate
poner <u>en</u> relievo	poner <u>de</u> relieve
<u>en</u> manera particular	<u>de</u> manera particular
se escribe <u>en</u> mayúscula	se escribe <u>con</u> mayúsculas
un pequeño titulo <u>por</u> su intervención	un pequeño título <u>para</u> su intervención

Furthermore, errors with the same value committed in both target languages are particular testimony to the validation of the study hypothesis:

Table 7. *Examples of preposition errors*

Erroneous form	Correct form
<u>at</u> the end[1]	<u>in</u> the end
<u>a</u> la primera fase	<u>en</u> la primera fase
key <u>for</u> material success	key <u>to</u> material success
la llave <u>por</u> el éxito	la llave <u>para</u> el éxito
to be understood <u>from</u> native speakers	to be understood <u>by</u> native speakers
era conocido solo <u>de</u> los expertos	era conocido solo <u>por</u> los expertos
would be surely left aside <u>from</u> many students	would be surely left aside <u>by</u> many students
ser aceptado <u>de</u> un grupo	ser aceptado <u>por</u> un grupo

Pronouns

The pronoun category is the one with the highest interference error rate for the Spanish set of data and the second in the ranking for the English production. However, even if most of the errors found in this category are caused by the influence of Italian, the total number of errors is not very high which means, generally speaking, that students do not show any particular difficulty in the use of English and Spanish pronouns. Nonetheless, English pronoun errors are very peculiar as they predominantly reflect the structure of Italian (being a more inflected language) which allows to waive the subject pronoun. Furthermore, the omission of the antecedent to the relative pronouns encountered in both the English and the Spanish corpora is a frequent error which happens due to the nature of the Italian pronouns *chi* and *che* that do not need an antecedent.

---

[1] used incorrectly to conclude the essay (*At the end, one cannot agree with the fact that English is an obligatory subject at a very young age.). It is to be noted that, this error, being a covert error, has not been considered for the computation of the error rates nor the creation of the error taxonomy. It is reported here for the sole purpose of comparison with its Spanish equivalent error which is instead an overt and competence error.

Table 8. *Some examples of pronoun errors from the English corpus*

Erroneous form	Correct form
Omission	
for who wants	for <u>someone</u> who wants
who speaks it correctly	<u>a person</u> who speaks it correctly
because is the most widly understood one	because <u>it</u> is the most widely understood one
is obvious that not all the English speakers really speak the same English	<u>it</u> is obvious that not all English speakers really speak the same English

Table 9. *Some examples of pronoun errors from the Spanish corpus*

Erroneous form	Correct form
Omission	
es la proximidad comunicativa que permite que el usuario percibe el intercambio con los demás	es la proximidad comunicativa <u>la</u> que permite que el usuario percibe el intercambio con los demás
no es el color de la piel sino la cultura y la procedencia que pueden determinar una variación	no es el color de la piel sino la cultura y la procedencia <u>las</u> que pueden determinar una variación
en una situación de igualdad social entre blancos y negros no es la raza que influencia sus usos lingüísticos	en una situación de igualdad social entre blancos y negros no es la raza <u>la</u> que influencia sus usos lingüísticos

Articles

The article category holds the third position among interference error rates for the English set of data, while for the Spanish set of data, it is the only category whose interference error rate is under 50%, thus representing mostly intralingual errors.

The reason why there is an interference error rate of 40% for the Spanish articles is the fact that the use of articles in Italian and Spanish almost coincides. Nevertheless, all erroneous applications of articles found in the Spanish corpus reflect divergence between Italian and

Spanish, such as the existence of two singular masculine definite articles in Italian (*il/lo*) and application of article before geographical names:

Table 10. *Some examples of article errors from the Spanish corpus*

Erroneous form	Correct form
Addition	
ambos <u>los</u> lenguajes	ambos lenguajes
la conquista de<u>l</u> América	la conquista de América
Misformation	
es necessario recordar <u>lo</u> estudio	es necesario recordar el estudio

Conversely, the use of articles in Italian differs from its usage in English. There are many situations in which the application of articles does not coincide in English and Italian. To be specific, Italian requires the employment of an article in relatively more situations. In other words, it can be said that in Italian, apart from some exceptions, the article must always be applied. For instance, unlike in English, in the Italian language an article is used before names of continents, countries, years, and possessive pronouns (with some exceptions).

Therefore, it is not surprising that the subjects of this study, like many learners of English, have problems with the use of English articles. However, the students' overuse of the definite article stands as testimony to the influence of L1 on TL. The results of the analysis of the English corpus show that the ratio of article addition to its omission is almost 15.5%. While most cases of article omission are intralingual, that is caused by either the complexity of the rules governing the English articles or by the students' lack of knowledge of these rules, many cases of article addition are interference errors. In other words, article addition happens in situations where the Italian grammar requires the use of an article. Some examples of interlingual erroneous employment of articles are as follows:

Table 11: *Some examples of article errors from the English corpus*

Erroneous form	Correct form
Addition	
<u>a</u> 70% has a general, not specific, awareness	70% has a general, not specific, awareness
<u>a</u> 50% of the world	50% of the world
majority of <u>the</u> people	majority of people
<u>the</u> comprehension is fundamental	comprehension is fundamental
<u>the</u> childhood	childhood
<u>the</u> tolerance	tolerance

A serendipitous finding of the investigation is the misuse of the English indefinite article *an* before words starting with /h/ followed by a vowel. Considering the letter H is always silent in Italian, these errors thoroughly reflect an Italian mindset that unconsciously ignores the letter H and consequently contributes to errors such as *an homogenous*, *an heritage* and *an high* […].

Nouns

The next category to deal with is the noun errors. The nature of errors appearing in the corpora differs dramatically between English and Spanish due to distinct structures and rules commanding each language. To be specific, while erroneous uses of nouns found in the English corpus are of different types, including omission of subject and misuse of gerund, errors encountered in the Spanish corpus constitute mostly lexical creation along with incorrect selection, addition or omission of a single vowel/consonant:

Table 12: *Some examples of noun errors from the English corpus*

Erroneous form	Correct form
Omission	
in job interviews <u>is</u> required to speak in English	in job interviews <u>candidates are</u> required to speak in English
Misformation	

the <u>employing</u> of one language	the <u>employment</u> of one language
the <u>spreading</u> of new medias	the <u>spread</u> of new media
such as the <u>developing</u> of	such as <u>development</u> of

Table 13. *Some examples of noun errors from the Spanish corpus*

Erroneous form	Correct form
Misformation	
en el óptica del <u>superamiento</u> de la dicotomía	en la óptica de la <u>superación</u> de la dicotomía
los ámbitos de la tecnología y del <u>divertimiento</u>	los ámbitos de la tecnología y de la <u>diversión</u>
no obstante el <u>intervento</u> de algunos grupos	no obstante la <u>intervención</u> de algunos grupos
analizó las fases de <u>aprendimiento</u> del inglés estandar	analizó las fases de <u>aprendizaje</u> del inglés estándar
tiene nada que ver con la <u>ridución</u> del tamaño	tiene nada que ver con la reducción del tamaño
con la <u>indipendencia</u> de las colonias	con la <u>independencia</u> de las colonias

Regarding the English language production, an error that particularly reflects one of the grammar rules of Italian is the use of singular form of the word *variety* instead of its plural form. This happens because in Italian the equivalent word *varietà* is invariant in singular and plural as it ends with a stressed vowel:

Table 14. *Examples of the incorrect use of singular form of the word variety*

Erroneous form	Correct form
Misformation	
one of the countless <u>variety</u> of English	one of the countless <u>varieties</u> of English
there are many other interesting <u>variety</u>	there are many other interesting <u>varieties</u>
pointing out that these <u>variety</u>	pointing out that these <u>varieties</u>
other <u>variety</u> of English have	other <u>varieties</u> of English have

Verbs

Verb is a category that entails a considerably high number of errors in both languages. However, as the interference error rates imply, factors contributing to the creation of these errors emanate from both intralingual and interlingual groups. In other words, even though the majority of errors are due to the influence of mother tongue, there is also considerable evidence of intralingual errors which occur because of faulty acquisition/lack of knowledge of TL rules or incomplete application of such rules.

As far as negative transfer is concerned, many verb errors found in the English corpus are examples of direct translation of Italian verbs/structures. For instance, the students resorted to a word-for-word translation of the Italian structures *entrare in contatto* or *fare un errore* which has led them to formulate incorrect English utterances:

Table 15. *Some examples of verb errors from the English corpus*

Erroneous form	Correct form
Misformation	
to <u>enter</u> in contact	to <u>come</u> into contact
errors non-native speakers <u>do</u>	errors non-native speakers <u>make</u>
the last one is <u>contrasted</u> by the government	the last one is <u>opposed</u> by the government
<u>is</u> no longer strictly <u>conform</u>	no longer strictly <u>conforms</u>

Errors of this kind, i.e. direct translation from Italian, are also encountered in the Spanish corpus:

Table 16. *Some examples of direct translation of Italian verbs into Spanish*

Erroneous form	Correct form
Misformation	
ahora <u>hacen</u> parte de nuestra cotidianidad	ahora <u>forman</u> parte de nuestra cotidianidad
los que <u>cubren</u> altos cargos institucionales	los que <u>tienen</u> altos cargos institucionales

In Spanish, the verb *estar* is used to describe location of something as well as any condition

that is temporary. However, the Italian verb *stare* does not have the same usage of *estar,* and these concepts are mostly expressed by the verb *essere.* Despite the emphasis focused on this distinction from the earliest stages of language teaching, errors of this kind are numerous in the language performance of Italian learners. This is also true of the subjects of this study, who made many errors of this type:

Table 17. *Some examples of verb errors from the Spanish corpus*

Erroneous form	Correct form
Misformation	
<u>son</u> relacionados	<u>están</u> relacionados
el registro <u>es</u> marcado	el registro <u>está</u> marcado
los jovenes <u>son</u> en una situación	los jóvenes <u>están</u> en una situación

Another type of interference error retrieved from the Spanish corpus is the erroneous use of the verb *ser* as an auxiliary verb in active voice. This occurs because in Italian, there are two auxiliary verbs, *avere* and *essere*, which are used to form compound tenses (broadly speaking, transitive verbs mainly take *avere* and all reflexive verbs and most intransitive verbs take *essere*). However, in Spanish only the auxiliary verb *haber* is used to form compound tenses, regardless of the nature of the main verb, and *ser* (the equivalent of *essere*) is used only as an auxiliary verb to form passive voice. In addition, in Italian when the auxiliary verb *essere* is used, the past participle needs to agree with the gender and the number of the subject, whereas in Spanish, the past participle remains invariable (except for passive voice):

Table 18. *Some examples of auxiliary verb errors from the Spanish corpus*

Erroneous form	Correct form
Misformation	
la frase no <u>es</u> terminad<u>a</u>	la frase no <u>ha</u> terminad<u>o</u>
cuando los ordenadores <u>son entrados</u> en las vidas de la mayor parte de las personas las cosas cambiaron	cuando los ordenadores <u>entraron</u> en las vidas de la mayor parte de las personas las cosas cambiaron

Adjectives

With regard to adjectives, while errors found in the Spanish corpus represent principally incorrect selection, addition or omission of a single vowel/consonant, errors occurring in the English corpus are of various kinds, such as incorrect word formation and pluralisation of adjectives:

Table 19. *Some examples of adjective errors from the English corpus*

Erroneous form	Correct form
Misformation	
<u>Italians</u> children	<u>Italian</u> children
<u>Africans</u> countries	<u>African</u> countries
<u>others</u> languages	<u>other</u> languages
the <u>politic</u> and economic domain	the <u>political</u> and economic domain
<u>educative</u> institutes	<u>educational</u> institutes
<u>fondamental</u>	<u>fundamental</u>

Table 20. *Some examples of adjective errors from the Spanish corpus*

Erroneous form	Correct form
Misformation	
textos <u>scientíficos</u>	textos <u>científicos</u>
estudios mas <u>recentes</u>	estudios más <u>recientes</u>
su <u>propria</u> imagen	su <u>propia</u> imagen
lenguaje <u>especialistico</u>	lenguaje <u>especializado</u>

In this regard, it is worth mentioning that errors such as *fondamental*, *scientífico*, *propria* and *recentes*, are not simply spelling errors, as each of them has been repeated various times in the corpora and they can be traced in the Italian language as well. For instance, *fondamentale* is the Italian word for *fundamental*, which explains the reason why *fondamental* is encountered 4 times in the corpus. Similarly, *propria* is the Italian word for *propia* which has appeared 6 times in the corpus by its Italian spelling rather than its correct form. Accordingly, the study has opted to include such errors under their linguistic category rather than a broad spelling-mistake category in order to highlight the L1 influence that has led to these errors.

Adverbs

Adverb is the last category to be discussed. Compared to other categories, adverbs account for a smaller number of errors. These errors are mainly made through direct translation from Italian to the TL:

Table 21. *Some examples of adverb errors from the English corpus*

Erroneous form	Correct form
Misformation	
a <u>hardly</u> debated point	a <u>hotly</u> debated point
to be <u>strongly</u> used	to be <u>greatly</u> used
<u>first</u> you learn it, <u>first</u> you improve it	the <u>sooner</u> you learn it, the <u>sooner</u> you improve it

Table 22. *Some examples of adverb errors from the Spanish corpus*

Erroneous form	Correct form
Misformation	
<u>vera y propria</u>	<u>propiamente dicha</u>
<u>todavía</u>, hay che tener en cuenta	<u>a pesar de ello</u>, hay che tener en cuenta
<u>todavía</u> la CMO ofrece	<u>sin embargo,</u> la CMO ofrece

The interlanguage analysis discussed so far has demonstrated the significant effect of negative transfer on the acquisition and the use of foreign languages. In other words, L1 has a very strong influence on foreign language development and greatly affects the process of language learning regardless of the level of similarity between L1 and TL.

Conclusion

The study has investigated whether native speakers of the same L1 will develop the same learning process towards different languages. The research, thus, set out to find patterns of errors in Italian learners' English and Spanish language production by studying two corpora consisting of essays produced by these learners. The results of the investigation have

demonstrated great consistency with the study hypothesis. Indeed, they have shown that learners' L1 plays an essential role in learners' L2 development and affects the process of language learning, regardless of the degree of kinship of the L1 and the TL. In essence, the level of similarity between L1 and TL does not change the degree of L1 influence, even if it affects types of error.

The study has examined two corpora which include writing data from Italian native speakers, for evidence of negative transfers from Italian to L2 writing. Moreover, through a close examination of the data, the research has strived to simulate the students' interlanguage systems and classify significant errors found in their foreign language written performance. Thus, the interlanguage analysis has revealed the decisive effect of negative transfer, which hinders learners' foreign language performance. In other words, the empirical evidence provided by the research supports that L1 has considerable influence on foreign language acquisition and use. Additionally, L1 influence has manifested itself approximately to the same extent in both the English and the Spanish language production, that is interference error rates vary minimally between the two target languages.

To be specific, interlingual errors outnumber intralingual errors in 11 out of 14 groups of error taxonomies. Moreover, prepositions have proven to be one of the most challenging aspects of language learning and, along with verbs, constitute the category with the highest number of errors. In addition, the study has identified numerous examples of direct translations, false friends, and word order errors that yield insights into the process of Italian's L2 development and use through illustration of characteristics and features of their interlanguage systems.

As highlighted earlier, in the process of learning a foreign language, L1 is a crucial basis for language acquisition as it helps learners to classify the language input, improve their language learning ability or simply reduce their anxiety. In other words, L1 is a source of knowledge that contributes to the development of foreign languages, and learners rely on it quite heavily, both consciously and unconsciously, to facilitate their language learning task. Thus, L1 can be regarded as a psychological tool that enables learners to understand the TL and retrieve language from memory. While at the beginner level this cognitive tool can provide support for students in their effort to achieve learning task; for high-proficiency learners, like the sample of this study, it serves to elaborate more sophisticated ideas and expand more elaborate content.

Therefore, according to the findings of this study, it can be concluded that L1 is a mediating tool in L2 learning, and regardless of the level of proximity/distance of L1 and TL, it helps students create and develop their interlanguage. Therefore, students create a parallel interlanguage system which relies on their mother tongue and reflects the L1 structures and rules. Thus, it is not surprising that empty spaces of interlanguage structure, caused by either lack of knowledge or complexity of TL, get filled with L1 elements. Nonetheless, in the early stages of language learning the interlanguage system is much closer to L1. Later, through enhancement of language proficiency, this system will greatly assimilate to TL and distance from L1 yet not totally independent.

As concerns the limitations of this study, comprehending and recognising the real source of language errors has always been challenging and problematic. It is a hard task and, arguably, in many cases, causes of errors cannot be identified with 100% certainty. This study is no exception. Although the best effort has been made to obtain the most acceptable and correct taxonomy, the reasons why some errors were attributed to L1 influence may not be universally shared. Furthermore, the current research has focused much attention on studying the role of negative transfer on the development of L2 and did not investigate the influence of other factors, such as positive transfer on the development of the interlanguage.

While, in the last decades, a prolific body of contrastive research has been conducted regarding the acquisition of Spanish by Italian native speakers, there are far less studies that address the acquisition of English by Italian native speakers from a contrastive perspective. Therefore, an important implication of this research is that, apart from its investigation into the learning of the English language, it has endeavoured to take a step further, by comparing Italian native speakers' interlanguage systems regarding two different languages.

As for the practical application of this research in L2 teaching and learning, the findings of this study might be used to make learners with the same profile aware of the major interference problems and the persistent interlingual errors. Additionally, considering this research provides information on common difficulties faced by Italian students in L2 production, language instructors might employ the findings as an aid to ease and improve the students' learning experience and outcome. L2 instructors can integrate the insights drawn from the authentic L2 data to warn learners against attested pitfalls. In particular, the findings can be

used to develop teaching materials, including remedial exercises targeting error-prone items or structures for this specific learner target, i.e. Italian students. For instance, the attested erroneous use of prepositions, the incorrect word formations, or false friends can easily be turned into choose-the-correct-option, word-building or vocabulary-enhancement exercises respectively.

Moreover, the fact that learner corpora pave the way for a more comprehensive and reliable qualitative analysis is now widely acknowledged. Consistently, this study fits into corpus linguistics research applied to SLA, a field of research that has yielded salient results with regard to learners from different mother tongues, but that as far as Italian EFL learning is concerned, is still largely unexplored. Considering the present study has provided new insights into the role of L1 in L2 learning, this research will hopefully open a new path to investigate Italian's EFL learning. Nevertheless, this study can be considered only the beginning of this enquiry, and there is a definite need to explore the effects of learners' L1 on their L2 writing in more depth. In this regard, a complementary study aiming to examine the results of the present research with the aid of elicitation tasks would be of great importance.

Biodata

Mahnaz Aliyar is a Persian native-speaker linguist fluent in English, Italian and Spanish. She received her bachelor's degree in Italian Language from the University of Tehran, Iran, and her master's degree in Modern Languages and Linguistics from the University of Bologna, Italy.

Holding a TESOL certificate, she has worked several years as a language tutor (teaching English, Italian and Persian). She is currently involved in a research project investigating the cross-linguistic influence in SLA. Her research interests include applied linguistics, second language acquisition, foreign language teaching and learning, and language development and use.

References

Aijmer, K. (Ed.). (2009). *Corpora and language teaching*. Amsterdam: John Benjamins.

Anthony, L. (2016). *TagAnt* (Version 1.2.0) [Computer software]. Waseda University. Tokyo, Japan. Retrieved January 06, 2018, http://www.laurenceanthony.net/software/tagant/

Anthony, L. (2018). *AntConc toolkit* (Version 3.5.6) [Computer Software]. Waseda University. Tokyo, Japan. Retrieved April 29, 2018, from http://www.laurenceanthony.net/software/antconc/

Axelsson, M. W. (2003). *Manual the Uppsala student English corpus (USE)*. Uppsala University. Retrieved February 13, 2018, from http://ota.ox.ac.uk/desc/2457

Axelsson, M. W. & Berglund, Y. (1999-2001). *The Uppsala Student English corpus (USE)*. The department of English, Uppsala University. Retrieved February 13, 2018, from http://purl.ox.ac.uk/ota/2457

Bailini, S. (2016). *La interlengua de las lenguas afines. El español de los italianos, el italiano de los españoles*. Milan: LED Edizioni Universitarie.

Barlow, M. (2005). Computer-based analysis of learner language. In Ellis, R. & Barkhuizen, G. P., *Analysing learner language* (pp. 335-357). Oxford: Oxford University Press.

Bayley, R., & Preston, D. (Eds.). (1996). *Second language acquisition and linguistic variation*. Amsterdam: John Benjamins.

Berglund, Y. & Oliver, M. (2002). The influence of external factors on learner performance. In Kettemann, B. & Marko, G. (Eds.), *Teaching and learning by doing corpus analysis* (pp. 205-215). Amsterdam: Rodopi.

Brown, H. D. (2000). *Principles of language learning and teaching*. (4th ed.). New York: Longman.

Bussmann, H. (1996). *Routledge dictionary of language and linguistics*. London: Routledge.

Calvi, M. V. (2003). Lingüística contrastiva de español e italiano. *Mots Palabras Words, 4*, 17-34. Retrieved February 16, 2018, from https://www.ledonline.it/mpw/allegati/mpw0403calvi.pdf

Corder, S. P. (1974). Error analysis. In Allen, J. P. B. & Corder, S. P. (Eds.), *Techniques in applied linguistics*. London: Oxford University Press.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.

Dulay, H. C., Burt, M. K. & Krashen, S. D. (1982). *Language two*. New York: Oxford University Press.

Ellis. R. (2015). *Understanding second language acquisition*. (2nd ed.). Oxford: Oxford University Press.

Foley, C., & Flynn, S. (2013), The role of the native language. In Herschensohn, J. & Young-Scholten, M. (Eds.), *The Cambridge handbook of second language acquisition* (pp. 97-113). Cambridge: Cambridge University Press.

Gass, S. M., Madden, C., Preston, D. & Selinker, L. (Eds.). (1989). *Variation in second language acquisition: Psycholinguistic issues* (Vol. 2). Clevedon: Multilingual Matters.

Gass, S. M. & Selinker, L. (2008). *Second language acquisition: An introductory course.* (3rd ed.).   Now York: Routledge.

Gilquin, G. & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, *1*(1), 41-61. Retrieved October 28, 2017, from http://hdl.handle.net/2078.1/70006

Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, *20*(3), 465-480. Retrieved November 01, 2017, from https://www.researchgate.net/publication/228602144_Error-tagged_learner_corpora_and_CALL_A_promising_synergy

Granger, S., Gilquin, G. & Meunier, F. (Eds.). (2015). *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.

Granger, S., Gilquin, G. & Menuier, F. (Eds.). (2013). *Twenty years of learner corpus research. Looking back, moving ahead*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Han, Z. H. & Tarone, E. (Eds.). (2014). *Interlanguage: 40 years later*. Amsterdam: John Benjamins.

Hyland, K., Chau, M. H. & Handford, M. (Eds.). (2012). *Corpus applications in applied linguistics.* London: Continuum.

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.

Krashen, S. (1981). *Second language acquisition and second language learning*. California: Pergamon Press.

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In Alonso Ramos, M. (Ed.), *Spanish learner corpus research: State of the art and perspectives* (pp. 236-265). Amsterdam: John Benjamins.

Lozano, C. & Georgopoulos A. (2006-2016). *CEDEL2: Corpus Escrito del Español como L2.* Universidad de Granada. Retrieved February 15, 2018, from http://cedel2.learnercorpora.com

Lozano, C. & Mendikoetxea, A. (2013). Learner corpora and SLA: The design and collection of CEDEL2. In Díaz-Negrillo, A., Ballier, N. & Thompson, P. (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 65-100). Amsterdam: John Benjamins.

Manchón, R. M. (Ed.). (2012). *L2 writing development: multiple perspectives*. Berlin: Mouton de Gruyter.

Mitchell, R. & Myles, F. (2004). *Second language learning theories*. (2nd ed.). London: Hodder Arnold.

Meurers, D. (2012). Natural language processing and language learning. In Chapelle, C. A. (Ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell Publishers. Retrieved October 17, 2019, from http://tjure.sfs.uni-tuebingen.de/~dm/papers/meurers-11.pdf

Myles, F. (2005). Interlanguage corpora and second language acquisition research. *Second Language Research*, *21*(4), 373-391. Retrieved October 30, 2017, from https://hal.archives-   ouvertes.fr/hal-00572085/document

Myles, F. (2007). Using electronic corpora in SLA research. In Ayoun, D. (Ed.), *Handbook of French applied linguistics* (pp. 377-400). Amsterdam: John Benjamins.

Pierini, P. (2012). *Comparing Italian and English. An introduction.* (2nd ed.). Rome: Aracne.

Pravec, N. A. (2002). A survey of learner corpora. *ICAME Journal*, *26*, 81-114. Retrieved November 10, 2017, from http://nora.hd.uib.no/icame/ij26/pravec.pdf

Ramón Torrijos, M. del M. (2009). Effects of cross-linguistic influences on second language acquisition: A corpus-based study of semantic transfer in written production. *Revista de Lingüística y Lenguas Aplicadas*, *4*, 147-159. Retrieved March 07, 2018, from https://doi.org/10.4995/rlyla.2009.741

Salmasi, S. (2001). *English-Italian contrastive analysis. Five points of syntax*. Bologna: CLUEB.

San Vicente, F. & Morillas, E. (Eds.). (2014)*. Oralidad contrastiva español-italiano: Aspectos gramaticales, discursivos y textuales. Cuadernos AISPI, 4*. Milan: Ledizioni.

Selinker, L. (1972). Interlanguage. International Review of Applied Linguistics in Language Teaching, 10(3), 209-231.

Selinker, L. (1983). A new account of language transfer. In Gass, S. M. & Selinker, L. (Eds.), *Language transfer in language learning* (pp. 98-111). Rowley: Newbury House.

Tarone, E. (2018). Interlanguage. In Chapelle, C. A. (Ed.), *The encyclopedia of applied linguistics*.   Oxford: Wiley-Blackwell Publishers. Retrieved September 13, 2019, from   https://doi.org/10.1002/9781405198431.wbeal0561.pub2

Tarone, E., Swierzbin, B., & Bigelow, M. H. (2006). The impact of literacy level on features of interlanguage in oral narratives. *Rivista di Psicolinguistica Applicata, 6*, 65-77. Retrieved September 17, 2019, from this Link.

Vázquez Veiga, N. (2016). Discourse markers in CEDEL2 and SPLLOC corpora of learner Spanish: Analysis of some lexical-pragmatic failure. In Alonso Ramos, M. (Ed.), *Spanish learner corpus research: State of the art and perspectives* (pp. 267-297). Amsterdam: John Benjamins.

Wolfersberger, M. (2003). L1 to L2 writing process and strategy transfer: A look at lower proficiency writers. *TESL-EJ*, *7*(2). Retrieved March 16, 2018, from http://www.tesl-ej.org/wordpress/issues/volume7/ej26/ej26a6/