

Article

Not peer-reviewed version

Lexical Transfer in Written Corpus of Learners of Portuguese as a Foreign Language

[Alessandra Baldo](#) *

Posted Date: 18 February 2025

doi: 10.20944/preprints202502.1362.v1

Keywords: Lexical deviant forms; lexical transfer; Portuguese as foreign language; lexical borrowings; neologisms



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Lexical Transfer in Written Corpus of Learners of Portuguese as a Foreign Language

Alessandra Baldo

Federal University of Pelotas, Brazil; alessabaldo@gmail.com

Abstract: This article presents a study on lexical deviations due to borrowings and neologisms made by 15 proficient learners of Portuguese whose native language (L1) was Italian, based on the analysis of ninety texts. The main objectives were twofold: (i) to identify the lexical deviations with the highest number of occurrences and (ii) to verify whether the language most frequently used in the deviations was the L1 or a previously acquired foreign language (FL). The data analysis resulted in the identification of 28 lexical deviations, mostly neologisms, as well as the use of the L1 as the preferred source language. Research findings challenge Kellerman's (1977) psychotypological hypothesis of languages, which asserts that the language typologically closest to L1, rather than L1 itself, is most frequently employed in lexical deviations. Explanations to the study outcomes will be based on research findings from Garcia-Lecumberri & Gallardo (2003) and Llach (2010).

Keywords: Lexical deviant forms; lexical transfer; Portuguese as foreign language; lexical borrowings; neologisms

Introduction

This study, whose general theme is bilingual lexical acquisition, aims to identify lexical deviations produced by a group of proficient Portuguese learners, speakers of Italian as L1, based on the analysis of their written productions. While there is extensive research on foreign language vocabulary acquisition, there has not yet been a study where all participants were native Italian speakers and fluent in the target language (TL). Thus, we understand that the data obtained from this study may be useful in attempting to understand the processes of lexical acquisition of foreign languages in advanced learners.

The article has five sections. The initial section presents the theoretical concepts of the study, while the second section describes the research methodology. The third section offers a detailed analysis of the data, and the fourth, a synthesis of the findings. This leads to the closing section, which discusses the contributions and limitations of the study, as well as potential areas for future research.

1. Theoretical Underpinning

1.1. FL Acquisition

In order to understand the nature of the linguistic acquisition of bilingual speakers, the notion of interlanguage, inseparable from linguistic transfer and initially introduced by Selinker (1972) is, more than necessary, fundamental: For the linguist, it is about the "existence of a linguistic system of its own based on the observable result that results from the learner's attempts to produce a norm of the target language" (1972: 214).

The concept is used to describe the distinct phases of learning a FL and assumes that the acquisition of a foreign language is better understood through the analysis of the different phases of learning development rather than through contrastive analysis (CA) of errors. The interlanguage is modified when new information allows the learner to develop hypotheses about the target language system, including its syntactic, phonological, morphological, and pragmatic rules.

According to Selinker (1972: 215), there are five mental processes that constitute the interlanguage. The first three are linguistic transfer, training transfer, which occurs when a previous learning method affects the learner's performance in FL acquisition, and learning strategies, which refer to the learner's attempts to develop their communicative competence in a FL. The remaining two processes are communication strategies in foreign languages, related to the ways in which the learner usually communicates in the target language, and the overgeneralization process of FL rules, which occurs when the learner uses a rule in an inappropriate context.

1.2. Loans/Borrowings and Neologisms

Regarding the phenomenon of linguistic transfer, among the most investigated in the lexical-semantic plane of language are loans, referred here by the broader word "borrowings", code-switching, and the creation of neologisms. In Haugen's (1950) classic text on loans, the author divides them into three categories: loanwords, the most general category that encompasses the importation of phonemic form and meaning, although the substitution of native phonemes may be more or less complete, without morphological substitution; loanblends, which include cases where, in addition to the substitution of sounds and inflections of the "model" language, there is also partial substitution of native morphemes; and loanshifts, related to loans with modifications at the level of meaning and at the level of form as well.

A reframing of Haugen's typology of loans is put forward by Grant (2020). The author starts by differentiating loanwords from loan creations, defined as new coinages from loan material. He also defines pure loanwords, which correspond to Haugen's loanwords, and loanblends, in which only part of the form of the lexical item is imported.

Grant (2020) also subdivides loanshifts into three categories: pure loan translations, defined previously as calques; loan renditions, which can be seen in more free and less literal translations of foreign words; and semantic loans¹, that refer to the extension of the meaning of a foreign word, mostly due to interferences from the L1 or from other previously learned FLs.

Besides loans and neologisms, Haugen (1950) cites code-switching as another concept used to describe linguistic contacts that involve alternations, both voluntary and involuntary, from L1 to the target-language (TL)². Unintentional alternations, the author explains, are designated as translinguistic transfers or interferences and occur when a specific lexical item is not available in the learner's mental FL lexicon, causing them to resort to available resources in the L1 lexicon and/or in the lexicon of an already known FL.

These "translinguistic interferences" refer not only to the influence of L1 on code-switching but also to any other foreign language known by the learner, a phenomenon identified by Kellerman (1977)³ and confirmed by a series of subsequent studies (cf. Ringbom, 2007; O Laoire and Singleton, 2009; Estrela and Antunes, 2017), which became known as the psychotypological hypothesis. What these studies have shown is that the main motivation for the loan is not necessarily the learner's native language or the foreign language of greater proficiency, but rather the one perceived by the learner as typologically closer to the target language.

Most studies on the topic have suggested that the influence of L1 decreases as proficiency in FL increases (Olsen, 1999; Herwig, 2001; Naves et al., 2005).⁴ Benatti (2024), in a recent study on Romanian speakers learning Italian as an L3 and having English as an already acquired FL, arrived

¹ The author also uses the alternate expression "semantic calques" to refer to "semantic loans".

² In this study, TL refers to the language being acquired at the time of the study, and FL to any other previously acquired foreign languages, or acquired in concomitance with the TL.

³ See section 3 for a more thorough discussion on the hypothesis.

⁴ A very similar finding was reported in studies on transfer of morphological structures, on section 1.4: "At initial stages, the L1 influence seems to be more prevalent while L1-related differences tend to flatten out with increased TL proficiency".

at the same conclusion: according to his analysis, transfer due to the typological effect is stronger within learners at the initial levels, especially at the lexical level.

A discordant result, however, is found in Garcia-Lecumberri & Gallardo (2003), who argue that transfer is the main strategy for all learners, with the difference being not in its use but in the motivations for doing so. According to the study data, in the initial stages of acquisition, L1 was used as a reference model for the development of new grammatical structures and vocabulary incorporation, and the number of L1 borrowings was greater. On the other hand, as the learner became more proficient in the FL, the cases of lexical borrowings from L1 decreased quantitatively, while the formation of words by coining and calquing increased qualitatively.

Llach (2010, p. 6) seeks to justify the study data of Garcia-Lecumberri and Gallardo (2003) based on the following reasoning: more proficient learners do not need to directly borrow words from L1, being able, in the case of calquing, to outperform literal translation and semantic extension from L1 to the FL. In the case of coinages, similarly, these learners were also able to use the linguistic system of the targeted language – more specifically, graphophonic and morphological rules – instead of the L1 structures.

1.3. Semantic Neologisms

In addition to coining and calquing, semantic neologisms are of particular interest in our study. According to Ferraz (2006, 221), they constitute the “expansion of meanings of lexical units already existing in the language, with new meaning”. Grant (2020), as previously mentioned, refers to them as “semantic loans”, explaining that they occur when “a native word undergoes extension of its meaning on the model of a foreign counterpart” (2020:24.)

It should be noted, however, that when it comes to research in foreign language vocabulary acquisition, the concept stirs some controversy. Leiria (2006), for instance, uses the term “lexical inadequacy” instead of semantic neologism, stating that the semantic boundaries of a word are in relation to other words and involve paradigmatic and syntagmatic relations, forming a network of connections.

Thus, argues the author, the peculiarity of FL acquisition is that such boundaries may not coincide in L1 and in FLs, and should, therefore, be reconceptualized, which “implies the definition of new semantic boundaries within each microsystem but may also imply a restructuring of the conceptual structure associated with L1” (Leiria, 2006: 257-258). It is exactly in the processes of reconceptualization and restructuring, the author explains, that learners incur in lexical deviations, employing a generalization or approximation strategy to the most appropriate word in the FL, which results in inappropriate lexical choices.

1.4. Recent Studies on Psychotypological Hypothesis

Not all recent research based on the psychotypological hypothesis posed by Kellerman (1977) mentions Kellerman, as there has been a switch of expressions, the most important being the use of “language distance” instead of “typological similarity”.

In some of the most recent studies, Eibenstein (2019), Diaubalick et al (2020) and Vallerossa (2021) have looked into the transfer of aspect from an L1 to a TL in learners who were proficient in another foreign language. According to Vallerossa (2021), these studies have identified the typological similarity/difference between the FL and the TL and the proficiency in the FL as the two main relevant factors influencing the target language performance. Studies have shown that knowing a FL similar to the TL helps transfer knowledge not only between languages of the same family (Diaubalick et al, 2020), but also between languages that come from different language groups but codify aspect similarly (Eibenstein 2019; Vallerossa et al, 2021).

Concerning proficiency, Eibenstein (2019) and Vallerossa et al. (2021) found that proficiency was a necessary condition for transfer of aspectual knowledge in English. In the studies, highly proficient learners of English did much better in their aspectual choices than the low-proficient ones.

Another topic of recent studies is the prototype factor of the FL. Salaberry (2005, 2020), Diaubalick et al. (2020), and Vallerossa et al. (2011) found out that “knowledge of a Romance language facilitates the acquisition of prototypical associations” in the TL, another Romance language in the studies, “while the L1 still constrains the acquisition of non-prototypical meanings” (Salaberry, 2020).

1.4. *The Psychotypological Hypothesis Revisited*

Over the last years, Kellerman’s typological hypothesis has been challenged on different grounds. Yan Lee (2022), for instance, points out that the “conceptualizations of what qualifies as “typologically similar” or “distant” are not well-established in bilingualism research” (2022: 3335). She argues that it is necessary to better consider the similarity-difference degrees and dimensions, and how bilingual demands are linked to the nature of language tasks in FL – i.e, receptive versus productive skills. Above all, she emphasizes the need to meticulously interpret bilingual neural activity and neurostructural changes before conclusions on the influence of typological distance can be reached.

After commenting on some of the most common dimensions of typological distances - writing systems and scripts, degree of lexical roots (Shook & Marian, 2013; Perfetti, Sun & All, 2015), morphosyntactic structures (Hope, 2020; Padovani et. all, 2015)⁵, and cognitive language barriers, Lee (2022) also mentions hypotheses put forward by scholars on basis of results of neuroimaging studies on bilinguals.

One of such hypotheses is made by Antoniou and Wrights’s (2017), who suggest an “interference inhibition effect” when comparable linguistic items between typological similar L1 and FL compete in the brain. As the authors argue that all languages are activated simultaneously, no matter the target language, that results in extra work to restrain unrelated forms in the non-target language. Lee correlates the authors’ findings with results derived from brain images, such as Hilchey and Klein’s (2011)’bilingual inhibitory control advantage hypothesis and Abutaeli and Green (2016)’s adaptative control hypothesis.

Furthermore, some scholars have studied the objective and the perceived effects of language distance on FL acquisition, be it a second, third or fourth language, be it acquired simultaneously with other FL or not. Ng and Min (2024), for instance, have investigated the role played by the perceived language distance among Korean, Cantonese and English in Korean learners whose Cantonese was their L1, and English, their L2, as well as the interplay of factors such as metalinguistic awareness and level of L3 proficiency. The main result was that language learning was more effective amongst the learners who perceived the L2 (English) and the L3 (Korean) as more distant than the ones who perceived them as less distant languages.

3. Methodology

In this section, we present the selection criteria for the informants and the corpus that allowed us to answer the two main questions of the study: (i) Is there a specific category of lexical deviations— borrowings or neologisms—that stands out for the number of occurrences?; and (ii) Which language is most used as the source of lexical and/or borrowings neologisms, the L1 or an additional FL? We begin with data related to the informants, and in the second part of the section, we focus on describing the characteristics of the written corpus.

3.1. *The Informants*

The informants to the study were 15 students enrolled in the course “Portuguese and Brazilian Literature: The South Atlantic is here” of the master’s program in Literature at the University of Bologna, which took place between February and April 2022, with a total of 4 hours and 30 minutes of weekly classes. The students voluntarily agreed to participate in the study, granting access to their

⁵ See also previous section.

opinion texts based on the topics studied during the course. Email requests for text usage and for information related to students' background on Portuguese learning were sent to them, and their acceptance and responses were saved.

The classification of the learners as proficient in Portuguese came from two instruments: from the assessment of the professor responsible for the course, the students' time dedicated to learning Portuguese, and the analysis of their textual productions, in which the use of appropriate Portuguese for the requested literary genre—i.e., opinion text in an academic context, with a low percentage of lexical and grammatical deviations—was verified.

Table 1 presents information related to the variant of Portuguese which the students reported to be more familiar with, obtained through a questionnaire sent via email to the participants. As for the Portuguese variant, nine of the fifteen participants declared that Brazilian Portuguese (BrPt.) was the predominant one, and the remaining six, the European variant (EuPt.). In addition, information related to years of Portuguese learning are in the third column, and knowledge of other foreign languages, in the fourth. The time dedicated to learning Portuguese refers to the course load of a university program in Portuguese Language Literature.

Table 1. Learners' profile.

Student	Pt. variants	Learning Time	Foreign Language(s)
1	BrPt.	6 yr.	Spanish, English
2	BrPt.	4 yr. 6 mo.	Spanish, English
3	EuPt. & BrPt.	4 yr 6 mo.	English
4	BrPt.	4 yr. 6 mo.	Spanish, English, French
5	EuPt.	4 yr. 6 mo.	Spanish
6	EuPt. & BrPt.	4 yr.6 mo.	Spanish, English
7	EuPt. & BrPt.	4 yr. 6 mo.	Spanish
8	EuPt. & BrPt.	6 yr.	French
9	EuPt.	4 yr. 6 mo.	English
10	BrPt.	5 yr.	Spanish, English, French
11	EuPt. & BrPt.	6 yr.	English
12	BrPt.	10 yr.	English
13	BrPt.	5 yr.	Spanish, French
14	EuPt. & BrPt.	4yr. 6 mo.	Spanish, English
15	BrPt.	4 yr. 6 mo.	Spanish, English

As far the knowledge of foreign languages (FLs) is concerned, the first information to note is that the most spoken FL is English (11speakers), followed by Spanish (10 speakers). Table 1 also shows that two subjects knew three languages in addition to Portuguese, French, Spanish, and English, and that seven subjects had knowledge of only one other FLs besides Portuguese: four claimed to know English; two, Spanish; one, French.

Transforming the numbers into percentages, we have that 60% of the informants declared being more familiar with the Brazilian variant of Portuguese, while 40% with the EuPt. variant. Concerning the time dedicated to study Portuguese, 60% had studied it for 4 and a half years, with only one of them (7%) had studied for 4 years. The remaining 40% had been studying Portuguese for no less than 5 years⁶.

3.2. The Corpus

The corpus consists of texts written by fifteen students from the master's program in Literature at the University of Bologna, produced during the course "Portuguese and Brazilian Literature II: The

⁶ One of the subjects was taught by her Brazilian mother on a non-systematically basis since they were 10 years old.

South Atlantic is Here" in 2022. The course was conducted online, and the texts were made available on the specific blog of the course. Each student wrote six opinion texts, based on the course syllabus content, totaling 90 texts. The average number of words per text was 544, and the texts totaled 48,958 words.

4. Presentation and Analysis of Data

There are two parts to this section, organized as follows: in the first, we describe and analyze the lexical deviations resulting from borrowings, either from L1 or any previously learnt FL(s) of the learners; in the second, subdivided into three parts, we first describe the interlinguistic and intralinguistic neologisms and then turn to semantic neologisms. The categories of borrowings and neologisms identified in the texts were adapted from the classification proposed by Leiria (2006). Table 2 shows the categories of deviations used in the analysis of our corpus, with adaptations from Leiria's proposal and the insertion of the category "semantic neologisms".

Table 2. Classification of Borrowings and Neologisms.

1 Borrowings	2 Interlinguistic Neologisms	3 Intralinguistic Neologisms	4 Semantic Neologisms
1a Borrowings from L1	Neologisms based on L1 Neologisms based on TL Extension of word meaning in TL		
1b Borrowings from FL(s)	Neologisms based on FL(s)		

In addition to the insertion of semantic neologisms in our analysis, another modification concerns the terminology adopted for the classification of neologisms: we adopted the terms interlinguistic and intralinguistic neologisms. While interlinguistic neologisms encompass two subcategories, those formed based on the L1 and those formed based on a FL, intralinguistic neologisms comprise only lexical creations entirely based on the TL system. Data identified in learners' texts are organized into tables. Borrowings are in Table 4, while neologisms are in following three tables. Table 5 presents the interlinguistic neologisms in which both the L1 and a FL are used; Table 6, the interlinguistic neologisms originated only using FLs systems; and Table 7, the semantic neologisms.

Firstly, we present the total number of deviations, separated by category, to guide us in the following sections.

Table 3. Number of Deviant Lexical Forms by Category.

Borrowings		Neologisms			
From L1	From FL(s)	Interlinguistic Neologisms L1-TL	Interlinguistic Neologisms FL(s) – TL	Intralinguistic Neologisms	Semantic Neologisms
3	-----	19	2	----	4
	3				25
Total					28

Table 3 shows, firstly, the small number of lexical deviations in relation to the total number of texts and words that comprise the corpus. Amongst the ninety analyzed texts, which totaled almost forty-nine thousand words (48,958, precisely), only 28 deviant lexical forms corresponding to the use of borrowings and neologisms were located. This data can be explained, on the one hand, due to the specificity of the deviations analyzed by this study, and on the other, to the learners' proficiency level. Another data that can be observed in Table 3 is the difference between the number of occurrences of deviations classified as borrowings and those classified as neologisms, a topic to which we will return later. Furthermore, the absence of FL borrowings and intralinguistic neologisms, i.e., those entirely

based on linguistic structures of FL(s), is noteworthy. Also noteworthy is the significant difference between the number of neologisms created from L1 and those created from FL.

4.1. Borrowings

Table 4 presents the identified borrowings in the corpus, placed in their immediate context, that is, in the sentences or sentence segments in which they occur in the texts. The neologisms are in italics, and the words in foreign languages (Portuguese, Italian and Spanish) are between quotation marks. Furthermore, the translation in English to the foreign words appear in the tables, immediately following the neologisms (and, in Table 4, the borrowings).

Table 4. Borrowings.

No.	Words in Context	Borrowings	
		from L1	from FL
1	...compares the slave ships to hell, which seems to be the only term of <i>paragone/comparison</i> in this mirror game...	Paragone	
2	... the <i>declino/decline</i> of a generation of men and women who participated in something they perhaps didn't understand.	Declino	
3	Most comments in Portuguese express longing for that <i>prosperoso/prosperous</i> time.	Prosperoso	

Only three borrowings were identified, and they all derived from the learners’ L1, the first one, “paragone” is probably the most distant from the corresponding Portuguese word, “comparação”. As for borrowings 2 and 3, *declino e prosperoso*, they are so similar with their Portuguese counterparts that it seems quite clear that the learners meant to write “declínio” and “próspero”.

Without undervaluing the analysis of borrowings from L1, another significant point presented in Table 4 is the absence of borrowings from foreign languages. If we refer to the FLs known by the authors of the texts, summarized in Table 1, we can see that English and Spanish were the most spoken languages —eleven learners knew English, and ten learners knew Spanish—and this data alone could justify the use of these FLs as a source of lexical borrowings. Additionally, as the same table shows, two of them also knew French, in addition to Spanish and English.

The first question that arises, thus, is the reason for the exclusive choice of L1 as the source language for borrowings. According to the literature, the expected outcome would be the use of borrowings from foreign languages as well, or at least from foreign languages closely related to the TL, but this was not what the data showed. As seen elsewhere, according to the psychotypological hypothesis in FL acquisition (Kellerman, 1977), the language perceived as typologically closest to the TL is the one most often used for linguistic borrowing, rather than the L1 or the FL in which the learner is most proficient.

Within that context, Leiria (2006)’s research findings based on the analysis of a written corpus of Learners of Portuguese from various L1s, are relevant. The scholar states that the language chosen for lexical borrowings was the one they understood as typologically closest to the TL, and that learners would resort to their L1s only in two situations: (i) when they were unfamiliar with the corresponding lexical item in either the TL or the nearest FL, and/or (ii) when they believed the words had the same spelling and meaning in the languages in contact.

Naturally, the informants in this study were aware of the typological proximity between Portuguese, Italian, and Spanish, as all are Romance languages, and knew that Spanish is even closer to Portuguese than Italian. Considering that, and assuming that Leiria’s (2006) results are correct, we can infer that the use of “paragone”, *declino*, and *prosperoso* was due to either the unfamiliarity with the corresponding lexical items in Portuguese and Spanish, or the belief that the words were written the same way in both Italian and Portuguese. Given the methodology of this research, however, it is not possible to categorically affirm that this was indeed the motivation for the use of L1, nor to differentiate which of the two situations raised by the author—i.e., unfamiliarity with the words in

the TL and Spanish or belief in identical spelling in L1 and TL—led the informants to use their L1s. In fact, it could simply be a momentary lack of access to an already acquired lexical item, which would invalidate the first two hypothesis.

In the next section, we observe the cases of neologisms, subdivided into three categories: interlinguistic neologisms between L1-TL (Italian - Portuguese), neologisms between a FL-TL (foreign language - Portuguese), and semantic neologisms.

4.2. Interlinguistic Neologisms

As Table 1 shows, there are 21 lexical deviations classified as interlinguistic neologisms. Of this total, nineteen are L1-TL interlinguistic neologisms, constructed by combining the morphological system of the L1 with the morphological system of the target language (TL), and two are -TL interlinguistic neologisms, constructed by combining the morphological structures of a FL and the TL. Thus, the prevalence of the L1 found in cases of borrowings also applies to neologisms, which naturally leads to the non-confirmation of the psychotypological hypothesis for this second type of lexical deviation.

Table 5 presents ten of the interlinguistic neologisms built by the amalgamation of the L1-TL linguistic systems, as the remaining nine follow the same pattern as the ones analyzed here.

Table 5. Interlinguistic neologisms.

No.	Words in Context	Interlinguistic Neologisms L1-TL
1	The Portuguese were the first to <i>esfrutare</i> /take advantage of the South Atlantic as a means of domination.	<i>Esfrutare</i> (it. sfruttare; pt. desfrutar, aproveitar)
2	The <i>exploitation/esfrutamento</i> of the South Atlantic...	<i>Esfrutamento</i> (it. sfruttamento; pt. exploração, aproveitamento)
3	The exploitation of the Atlantic brought a lot of <i>suffering/soferença</i> and destruction in Africa.	<i>Soferença</i> (it. sofferenza; pt. sofrimento)
4	...in the sense of tolerating the <i>presencia</i> /presence of the other without really accepting or including them. ...no one wants to accept the <i>presencia</i> /presence of racial hatred.	<i>Presencia</i> (it. presenza, pt. presença)
5	There is a <i>diferencial</i> /difference in war literature between reflection and consequence texts...	<i>Diferencia</i> (it. differenza, pt. diferença)
6	Using the rosary in their <i>pregueiras</i> /prayers would contribute to them having a place in paradise.	<i>Pregueiras</i> (it. preghiere; pt. orações)
7	The alternative representation of former subordinates pushes the contemporary <i>emarginados</i> /marginalized ...	<i>Emarginados</i> (it. emarginati; pt. marginalizados)
8	The literature on one side preserves the individual traumatic <i>testemunhança</i> /testimony of the events of the war.	<i>Testemunhança</i> (it. testimonianza; pt. testemunho)
9	Society is mainly androcentric, <i>machilista</i> /chauvinist that subjects women from the beginning.	<i>Machilista</i> (it. maschilista; pt. machista)
10	...is placed in the <i>amplia</i> /broader perspective of Atlantic Studies.	<i>Amplia</i> (it. ampia, pt. ampla)

Regarding the first two neologisms, *esfrutar* and *esfrutamento*, it should be noted that among the prefixes in Portuguese with meanings close to the Italian *s-* are *es-*, *des-*, *in-*, and *a-*, considering that the Italian *s-* has a negative, deprivation value, as well as a reverse action depending on the context (Grossmann & Rainer, 2004:137). Assuming the learner knew the prefixes in the TL given their level of Portuguese knowledge, the most likely hypothesis for using *es-*, instead of the other prefixes

mentioned above, is that they selected the affix based on a graphophonological adaptation principle to the TL system. The same reasoning seems suitable for neologism 2, whose accommodation to the TL, compared to the Italian word “sfruttamento”, affects the spelling of the prefix and the root.

Neologisms 3 to 5 exhibit remarkably similar construction patterns, as in all cases, the observed modification is the substitution of the Italian suffix *-nz(a)* with its Portuguese counterparts. In *soferença*, it can be noted that it was constructed by combining the Italian lexeme *soffere-* (from which words like “sofferente, insofferente, sofferto” are derived) with the Portuguese suffix *-nç(a)* instead of the Italian *-nz(a)*. The next two neologisms, 4 and 5, *presencia* and *diferencia*, were formed by adapting the formation pattern of the Italian word *conseguere+nz(a)*, and the deviation was caused by the inadequate suffix selection, as the Italian constituent *-nz(a)* can be replaced in Portuguese by either *-nça* or *-ncia*. These are examples of substituting the Italian suffix *-nz(a)* with the Portuguese affix *-ncia* when the appropriate choice would have been *-nç(a)*.

Considering neologism 6, *pregueiras*, the same pattern of adaptation to the TL's morphological system is observed: the lexical bases *pregu-* (from Italian) are combined with the suffix *-eira*, substituting the Italian correlate *-iere*.

In *emarginados*, there is an adaptation of the Italian participle suffix *-to* (*a, e, i*) to the corresponding Portuguese suffix *-do*, followed by the plural suffix *-s*. It should also be noted that the learner uses the prefixal scheme of L1 (*e+margin+are*) in a context where the TL uses suffixes (*pt. margin+al+izar*).

Neologisms 8 and 9, *testemunhança* and *machilista*, have peculiarities compared to the others: *testemunhança* lacks coincidence between the morphological structure of the word in the TL and the created neologism, while *machilista* presents a divergent lexical base, although the suffix is common between the two languages (L1 and TL). In the first case, even though there is no formal coincidence with L1 in any of the morphological constituents (base and suffix), there is interference from the word formation scheme that characterizes the Italian correlate, “testemonianza”. The learner constructs a noun in the TL with a similar morphological structure to the word in L1, transposing the base *testimonia-* and the suffix *-nza* to the Portuguese correlates “testemunha”- and *-nça*, respectively. The fact that word formation schemes do not coincide in Portuguese and Italian, however, is the cause of the deviation. While in Portuguese, the noun “testemunho” is formed by non-affixal derivation (also known as conversion), in Italian, it is formed by adding the suffix *-nz(a)* to the verb stem “testimoniare”. Conversely, the adjective *machilista* shares the same suffix *-ista* as the corresponding Portuguese word “machista”. Thus, the differences at the lexical base level (*maschil-* in Italian, *mach-* in Portuguese) allow us to classify it as a neologism.

The last of the listed neologisms, *amplia*, represents the group of neologisms constructed by adapting Italian words to Portuguese phonological/orthographic and morphological patterns. In *amplia*, for example, it is noticed that the corresponding Italian word “ampia” had the grapheme <l> added.

Before concluding this section, it is worth noting the formal adaptation of some previously presented neologisms to the spelling and morphology of the TL. In items 1 and 2, for instance, the learner, aware that Portuguese spelling does not include words beginning with the consonantal sequence *sf-*, added the initial vowel *e* to the prefix *s-*. As for *pregueiras* (“preghiere”/prayers), the suffix *-ier(e)* (nonexistent in Portuguese morphological system) was replaced by the suffix *-eir(a)*, with the addition of the plural suffix *-s*, a modification consistent with the Portuguese morphological system. Additionally, a modification in the lexical base spelling is observed by the substitution of the grapheme <gh> (nonexistent in Portuguese) for <gu>. One last observation related to the adaptation of the L1 spelling to the TL spelling is the suppression of double consonants in the neologisms *esfrutare*, *esfrutamento*, and *soferença*. Although this difference between Portuguese and Italian is taught to learners early on, it is not always practiced, mainly because it is more of a phonological than a graphemic issue.

The two interlinguistic neologisms formed from the Portuguese and Spanish language systems are described in Table 6.

Table 6. Interlinguistic Neologisms LE-TL.

No.	Words in Context	Neologisms
1	<i>Empeçou</i> também um processo... It <i>began</i> a process as well...	<i>Empeçou</i> (sp. empezò; pt. iniciou)
2	...o elemento que <i>garantizou</i> aos senhores brancos The element that <i>guaranteed</i> to the white man....	<i>Garantizou</i> (sp. garantizò; pt. garantiu)

In these two deviations, *garantizou* and *empeçou*, we have verb forms created based on words from Spanish (“garantizò, empezò”), adapted to the inflectional system of Portuguese verbs. Additionally, the learner, in the radical of *empeçou*, replaces the grapheme <z> with <ç>, which may have a phonological motivation.

4.3. Semantic Neologisms

Four semantic neologisms were identified, as shown in Table 7.

Table 7. Semantic Neologisms.

No.	Context	Semantic Neologisms
1	...alternative to the <i>conto</i> / <i>version</i> that has always being told	<i>Conto</i> (it. conto; pt. explicação, versão)
2	He feels no patriotic pride in being forced to <i>suportar</i> / support Portugal in the colonial war...	<i>Suportar</i> (it. supportare; pt. apoiar)
3	Based on the fear and violence that previously <i>subiam</i> / <i>suffered</i> the slaves.	<i>Subiam</i> (it. subire; pt. sofrer)
4	...the consequent problematic resolution of the post- abolition period represents today a heavy <i>herdade</i> / <i>heritage</i> that continues to show the effects...	<i>Herdade</i> (it. eredità; pt. herança)

The word “conto”, for example, has a semantic value coinciding with one of its meanings in L1, which attributes to the term, in the specific context of the learner’s production, the meaning of ‘narration, story’. It should be noted that the same sense of ‘narration’ and ‘story’ is also attested in Portuguese, both in the first entry of the Priberam dictionary, which defines it as “fictional story”, and in the first two entries of the *Dicionário da Língua Portuguesa* (2010), which consist of “1. short and fictional narrative in which the action generally focuses on a single theme or episode; 2. tale, fable.” However, “conto” is a polysemous term and the base of dozens of idiomatic expressions in Italian, far exceeding the more well-known meanings in Portuguese, restricted to the literary tale and to lie/deception. The relevant entry is “explanation”, and the idiomatic expression is “fare i conti con qualcuno”, whose approximate translation in Portuguese would be “to obtain an explanation, reparation”. Hence the strangeness of the occurrence of the word in the sentence in which it was used, since the most appropriate word for a native Portuguese speaker would be “explicação, versão”. Thus, this specific case could be classified as a semantic-pragmatic neologism, as the new meaning attributed to the word partly retains the original meaning of the word, while partially conveying the idea of the Italian term.

Returning now to the verb “suportar”, we classify it as a semantic neologism because we found only one entry in the four dictionaries consulted that corresponded to the learner’s intended meaning, which was “to support”. This meaning does not correspond to the usage in the Priberam Dictionary of the Portuguese Language, which are “to have on oneself; to endure; to permit, to tolerate; to suffer; to withstand”, and in the 2003 and 2010 editions of *O Dicionário da Língua Portuguesa*, “to have on oneself; to be the base or support of; to sustain the weight of; to endure; to suffer; to tolerate; to admit; to bear with”. However, in the 2014 edition of the *Dicionário Global da Língua Portuguesa*, we found a new entry that is “to support, to help”, which has been more recently adopted by Portuguese speakers, influenced by the verb “to support” in English.

The corresponding Italian term “supportare” is defined in the Treccani Encyclopedia as a “neologism derived from French *supporter* and English *support*”, whose figurative senses correspond to “help, support”. As seen, the same derivation process (from English) is also occurring in Portuguese, but it is still in an initial phase, causing, in many speakers of the language, either a greater cognitive effort to understand the statement or even misunderstanding of it. Due to this status of the word in the Portuguese lexicon, it seemed more appropriate to classify it as a semantic neologism.

The lexical item “herdade”, in the context of the statement “[...] the consequent problematic resolution of the post-abolition period represents today a heavy *herdade* that continues to show the effects”, is characterized as a semantic neologism because the attested meaning does not match the meaning that the word normally expresses in the TL. With significantly greater frequency in European Portuguese than in Brazilian Portuguese, the word “herdade” is used to designate a “large rustic property usually composed of sowing lands” and being synonymous with “quinta”. However, in the statement produced by the learner, the word is used in the sense of “inheritance”.

The last deviation classified as a semantic neologism is the use of “subiam” in the statement “[...] based on the fear and Violence that previously *subiam* the slaves”. Although “subir” in Portuguese and “subire” in Italian have an (almost) coinciding form, the issue is that the Italian term is a false cognate. According to the Zingarelli 2018 dictionary, the two main senses of “subire” are: (1) “to be forced to endure something harmful or unpleasant; (2) to submit to something”. Conversely, the entry for the verb “subir” in the online Portuguese dictionary Priberam is subdivided between intransitive uses (“to go up, to climb, to elevate, to increase, to become more expensive”) and transitive uses (“to climb, to traverse or pull up, to exalt, to magnify”). The learner’s statement leaves no doubt that the appropriate word would be “suffer” or even “to be subjected to”, but not any of the meanings of “subir” in Portuguese.

5. Summary of Data Analysis

In a corpus composed of 90 texts and approximately 49,000 words, a total of 28 deviations were found, including borrowings, interlinguistic neologisms from L1 or FL borrowings, and semantic neologisms. As described in Table 3, the lexical deviations were divided as follows: three deviations correspond to borrowings, and 25 to the creation of neologisms; among the three deviations classified as borrowings, all were from the informants’ L1; out of the total neologisms, 21 were classified as interlinguistic; among the interlinguistic neologisms, nineteen were created based on the L1 system, and two based on the Spanish system as an FL; no intralinguistic neologisms were verified; finally, four semantic neologisms were identified.

The higher number of neologisms compared to borrowings is related to the informants’ proficiency level and aligns with the findings of Garcia-Lecumberri and Gallardo (2003) and Llach (2010). The first researchers found that as the learner became more proficient in the TL, the cases of L1 lexical borrowings decreased, while neologisms formed by coinage and calque increased. That occurs, according to Llach (2010:6), because more advanced learners can establish semantic connections between L1 and TL and apply the TL’s graphophonological and morphological rules necessary for creating these types of neologisms.

It is essential to revisit a previously mentioned topic here, as it has a direct impact on our study, which is the lack of a standard categorization for lexical deviations. The fact that there is no unanimity in defining what constitutes a neologism in the field of foreign language acquisition means that the same data treated here as neologisms could be classified differently under other methodologies. As an illustration, it suffices to recall the lexical borrowing categories defined by Haugen (1950), summarized in the first part of this work: what the author classifies as loanblends and loanshifts are, in recent literature, neologisms. The most obvious consequence of this situation is the need, before adhering to the specific name assigned to the lexical deviation in the literature, to keep in mind the concept that underpins it. Although it is a basic principle of any investigation that desires a minimum of method, the linguistic proximity between the L1, the most commonly used FL as a borrowing

source, and the TL added an extra layer of difficulty to the task of categorizing deviations in our study.

L1 was the only language used in the three borrowing occurrences observed in our corpus – *paragone*, *declino e prosperoso*. As the data did not align with Kellerman's (1977) psycholinguistic hypothesis, alternative explanations were found in Leiria (2006). According to the researcher, there are three possible motivations for using L1 instead of a FL considered typologically closer to the TL by learners: unfamiliarity with the corresponding lexical item in the FL/FLs, the belief that, in fact, the lexical item in Portuguese matched in form and meaning with Italian, or a momentary lack of access to an already acquired lexical item. The fact that the methodology adopted in this study did not allow us to identify which of the motivations was selected by the informants limited us to accepting all three as possible.

Neologisms, in line with borrowings, were predominantly created based on the L1. Out of the total of twenty-one interlinguistic neologisms, nineteen had Italian as their support, and the other two were based on Spanish. The study concludes that informants preferred using L1 vocabulary over additionally spoken FLs due to the greater extent and activity of their L1 mental lexicon.

This result complemented the fact that we did not find intralinguistic neologisms, which may seem, at first glance, contradictory given the learners' TL proficiency. However, a slightly more careful analysis may reveal that exactly the learners' prominent level of Portuguese knowledge is one of the possible explanations for this fact. This is because, as the TL mental lexicon expands, the learner becomes more adept at recognizing the morphological (and morphophonological) combinations allowed by the linguistic system in acquisition, which would function as a kind of barrier to creating deviant lexical items constructed solely with the target language system. This hypothesis is further supported by the fact that, having more linguistic resources than a beginner learner, for example, the ability to assess the potential inadequacy ("strangeness") of a lexical item is greater, as is the possibility of replacing it with a synonym or an alternative expression that conveys the same concept, which would be much less likely to happen with a learner with a limited TL lexicon. Thus, it could be argued that when lexical deviation occurs, it happens either by borrowing or by creating neologisms based on borrowings preferably from the learner's L1 or even from a previously acquired foreign language, simply because, being more aware that the word they need is not available in the TL, they seek as a secondary strategy either the L1 or any FL they consider typologically closer to the TL.

Final Considerations

The motivation for developing this study resided in the possibility of identifying specific characteristics of deviations stemming from borrowings and the creation of neologisms due to two characteristics of the informants: the fact that they all shared the same L1, Italian, and their elevated level of proficiency in the TL, Portuguese. The results specifically related to the learners' proficiency level in the TL were threefold: (i) the higher occurrence of neologisms compared to borrowings, (ii) a greater use of L1 in both borrowings and the creation of neologisms; and (iii) the absence of intralinguistic neologisms, constructed entirely from the TL system.

Regarding neologisms, the study expanded the categorization proposed by Leiria (2006) to include the use of words that, although existing in the Portuguese lexicon, did not possess the meanings attributed to them by the informants. The words within this category were called semantic neologisms, and in our analysis, the semantic extension attributed to them was always derived from the informants' native language.

As seen throughout this text, the data largely corroborated the findings of previous studies with similar objectives, although with methodological differences concerning the classification of lexical deviations and the informants' L1. The most relevant finding of the study is the non-verification of the psychotypological hypothesis of languages in the process of Portuguese acquisition by Italian learners, both for lexical borrowing deviations and neologism creation deviations. In both cases, as seen throughout this study, L1 was the most employed language.

The hypotheses raised to explain such results, although they seem adequate to us, were sought in research with methodological differences compared to this study. In this sense, we understand that studies aiming to investigate these same questions, also considering a corpus composed of basic-level Portuguese informants, or even advanced-level informants but from a more extensive corpus, will certainly contribute new knowledge to the area of lexical acquisition.

In terms of this study's contribution and resuming Booij's (2005:264) statement that word creation processes are fundamental as they represent a window into learners' minds, we assess that the data treatment conducted here provided a more refined understanding of the mental processes learners engage in during lexical production in a foreign language.

It should be noted, however, that identifying these processes was permeated with doubts. Additionally, multiple hypotheses were proposed to explain certain neologisms when they appeared plausible.

We understand that many of our doubts, however, rather than revealing a theoretical-methodological lack of understanding of word formation processes, reflected the complexity of the classification task, as classifying presupposes understanding the mental processes involved in lexical deviation, and these occur largely unconsciously. Generally, not only are learners unaware of the deviations, but they would also likely be unable to explain the motivations that lead them to make them.

Considering this, another possibility for new studies could arise from research using introspective techniques to understand learners' neologism creation mechanisms, such as the use of verbal protocols. Although many of the hypotheses available in the literature to explain them have been considered coherent considering our data, in many cases, it was not possible to verify which one was exactly the most appropriate in a specific deviation occurrence because we did not have access to the informants' mental processes while – or shortly after – involved in producing their texts. We are aware of how laborious such studies are. However, in our view, they are also indispensable. The more accurately we understand the mechanisms that lead FL learners, regardless of proficiency level, to incur in lexical deviations, the better equipped we will be, as language professionals, to assist them in their bilingual acquisition process.

We hope this study will help non-native language teachers understand the mechanisms used by learners to effectively express themselves in a second language. And, naturally, also encourage them to integrate in their teaching practices didactic activities aimed at transforming their students' lexical deviations into successful word selection and formation strategies. These strategies can consist, for example, of simple exercises recognizing similar words between L1, closely related languages, and the TL, recognizing the most productive affixes, along with their meanings, in the TL, in contrast to L1, or more complex tasks, such as producing different words derived from a single lexeme in appropriate language contexts.

Conflict of Interest: The author declares no conflict of interest.

References

- Abutalebi, J & D. W. Green (2016). Neuroimaging of Language Control in Bilinguals: Neural Adaptation and Reserve. *Bilingualism: Language and Cognition* 19 (4): 689–698.
- Antoniou, M., & S. M. Wright (2017). Uncovering the Mechanisms Responsible for Why Language Learning may Promote Healthy Cognitive Aging. *Frontiers in Psychology* 8: 2217–2217.
- Bennati, R. (June 2022). Italian as L3 of Romanian university students: transfer from English and psychotypology effect. *Quaestiones Romanicae*, IX (2):116-130.
- Booij, G. (2005). *The Grammar of Words – an introduction to linguistic morphology*. Oxford University Press.
- Dicionário Priberam da Língua Portuguesa. <https://dicionario.priberam.org/>
- Dicionário da Língua Portuguesa (2010). Porto Editora.
- Dicionário de la Lengua Española. Real Academia Española. <https://dle.rae.es/>.
- Dicionário Global da Língua Portuguesa (2014). LIDEL.

- Dizionario de Italiano do Jornal A Repubblica. <https://dizionari.repubblica.it/italiano.html>
- Enciclopedia Treccani. <https://www.treccani.it/vocabolario/>
- Estrela, A. & Antunes, S. (2017). A sufixação num corpus de aquisição de PLE/L2. *Pelos Mares da Língua Portuguesa*, 3, 905-924.
- Ethnologue. Languages of the World. <https://www.ethnologue.com/>.
- Ferraz, A. P. (2006). A inovação lexical e a dimensão social da língua. In M. C. T. C. Seabra (ed.). *O léxico em estudo*, (pp. 218-234). UFMG.
- Garcia-Lecumberri, M. L. & Gallardo, F. F. (2003). English FL sounds in school - Learners of different ages. In M. P. Garcia-Mayo & M. L. Garcia-Lecumberri (eds.). *Age and the acquisition of English as a foreign language* (pp. 115-135). *Multilingual Matters*.
- Gass, S. M. & Selinker, L. (2008). *Second Language Acquisition – an introductory course* (3rd ed.). Routledge.
- Ghazi-Saidi, L., and A. I. Ansaldo. 2017. "The Neural Correlates of Semantic and Phonological Transfer Effects: Language Distance Matters." *Bilingualism: Language and Cognition* 20 (5): 1080–1094.
- Grossman, M & Rainer, F. (2004). *La formazione delle parole in italiano*. Niemeyer.
- Haugen, E. (1950). Analysis of linguistic borrowing. *Language* 26, 210-231.
- Herwig, A. (2001). Plurilingual Lexical organization: Evidence from lexical processing in L1-L2-L3-L4 translation. In J. Cenoz, B. Hufeisen & U. Jessner (eds.). *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 115-137). *Multilingual Matters*.
- Hilchey, M. D. & R. M. Klein (2011). "Are There Bilingual Advantages on Nonlinguistic Interference Tasks? Implications for the Plasticity of Executive Control Processes." *Psychonomic Bulletin & Review* 18: 625–658.
- Hopp, H. (2020). Morphosyntactic Adaptation in Adult L2 Processing: Exposure and the Processing of Case and Tense Violations. *Applied Psycholinguistics* 41 (3): 627–656.
- Kellerman, E. (1977). Toward a characterization of the strategy of transfer in second language Learning. *Interlanguage Studies Bulletin*, 2, 58-145.
- Llach, M.P.A. (2010). An overview of variables affecting lexical transfer in writing: a review study. *International Journal of Linguistics*, 2, 1-17.
- Lee, Y. Y. (2022). A conceptual analysis of typological distance and its potential consequences on the bilingual brain. *International Journal of Bilingual Education and Bilingualism*, 25(9), 3333–3346.
- Leiria, I. (2006). *Léxico, aquisição e ensino*. Fundação Calouste Gulbenkian/ Fundação para a Ciência e Tecnologia.
- Lowie, W. (1998). The acquisition of interlanguage morphology: a study into the role of morphology in the L2 learner's mental lexicon. [Doctoral thesis. University of Groningen]
- Naves, T., Mirapeis, I. & Celaya, M.L. (2005). Who Transfer More ... and What? Cross-linguistics Influence in Relation to School Grade and Language Dominance in ELE. *International Journal of Multilingualism*, 2, 113-134.
- Ng, T. Y., & Min, B. (2024). Factors affecting L3 acquisition: effects of perceived language distance and metalinguistic awareness on L3 Korean among Hong Kong learners. *International Journal of Bilingual Education and Bilingualism*, 1–14.
- Ò Laoire, M. & Singleton, D. (2009). The role of prior knowledge in L3 Learning and use: Further evidence of psychotypological dimensions. In L. Aronin & B. Hufeisen (eds.). *The exploration of multilingualism: Development of research on L3, multilingualism and Multiple language acquisition* (pp. 79-102). John Benjamins.
- Olsen, S. (1999). Errors and compensatory strategies: a study of grammar and vocabulary in texts written by Norwegian Learners of English. *System*, 27, 191-205.
- Padovani, R., G. Calandra-Buonaura, C. Cacciari, F. Benuzzi, and P. Nichelli (2005). Grammatical Gender in the Brain: Evidence from an fMRI Study on Italian. *Brain Research Bulletin* 65 (4): 301–308.
- Perfetti, C. A., Y. Liu, J. Fiez, J. Nelson, D. J. Bolger, and L. Tan. (2007). Reading in Two Writing Systems: Accommodation and Assimilation of the Brain's Reading Network. *Bilingualism* 10 (2): 131–146.
- Ringbom, H. (2007). *The Importance of Cross-linguistic Similarity in Foreign Language Learning: Comprehension, Learning and Production*. *Multilingual Matters*.
- Shook, A. & V. Marian (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition* 16 (2): 304–324.

Salaberry, M. Rafael (2020). The conceptualization of knowledge about aspect. In *Third Language Acquisition: Age, Proficiency and Multilingualism*. Edited by Camilla Bardel and Laura Sánchez. Eurosla Studies 3. Berlin: Language Science Press, pp. 43–65.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.

Zingarelli, N. (2017). Lo ZINGARELLI 2018. Vocabolario della lingua italiana. Zanichelli.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.