

ACQUISITION OF ARTICLES BY ESTONIAN LEARNERS OF SPANISH AS L3: A CORPUS STUDY

VIRGINIA RAPÚN MOMBIELA
University of Tartu
virginia.rapun@ut.ee
ORCID: 0000-0001-8643-3410

GUILLEM CASTAÑAR RUBIO
Tallinn University
guillem.castanar_rubio@tlu.ee
ORCID: 0000-0001-5969-0024

ANNA TUDELA ISANTA
The Open University
anna.tudela-isanta@open.ac.uk
ORCID: 0000-0002-4379-706X

ABSTRACT

This study delves into the use of Spanish articles by Estonian students learning Spanish as their third language (L3). This topic has received little attention to date, despite evidence that it represents one of the greatest challenges Estonian learners of Spanish face, and it appears that these difficulties persist even at higher levels of Spanish proficiency (Kruse 2018: 126). This article aims to provide a comprehensive overview of how Estonian students use Spanish articles. 345 written texts from a learner corpus have been analysed. Results show that the most common error among learners is omission of the article, in line with previous research, but also an unexpectedly high incidence of morphological and syntactic errors. Moreover, it has been detected that there is not a significant improvement in article acquisition between levels A1 and B1. By addressing these objectives, this study contributes to a deeper understanding of the specific difficulties faced by Estonian learners and the broader landscape of article acquisition in a third language context.

KEYWORDS: article acquisition, learner corpus research, third language acquisition, longitudinal approach, Spanish as a foreign language, error analysis.

ADQUISICIÓ D'ARTICLES PER PART D'APRENENTS ESTONIANS D'ESPANYOL COM A L3: UN ESTUDI DE CORPUS

RESUM

Aquest estudi aprofundeix en l'ús dels articles en espanyol per part d'estudiants estonians que aprenen espanyol com a tercera llengua (L3). Aquest tema ha rebut poca atenció fins ara, malgrat l'evidència que representa un dels majors reptes a què s'enfronten els aprenents estonians de l'espanyol, i sembla que aquestes dificultats persisteixen fins i tot en nivells superiors de competència en espanyol (Kruse 2018: 126). Aquest article té com a objectiu proporcionar una visió global de com els estudiants estonians utilitzen els articles en espanyol. S'han analitzat 345 textos escrits d'un corpus d'aprenents. Els resultats mostren que l'error més comú entre els aprenents és l'omissió de l'article, en línia amb investigacions anteriors, però també una

Data de recepció: 15/viii/2024
Data d'acceptació: 30/ix/2024
Data de publicació: desembre 2024

incidència inesperadament alta d'errors morfològics i sintàctics. A més, s'ha detectat que no hi ha una millora significativa en l'adquisició d'articles entre els nivells A1 i B1. En abordar aquests objectius, aquest estudi contribueix a una comprensió més profunda de les dificultats específiques dels aprenents estonians i del panorama més ampli de l'adquisició d'articles en un context de tercera llengua.

PARAULES CLAU: adquisició de l'article, recerca amb corpus d'aprenents, adquisició de terceres llengües, enfocament longitudinal, espanyol com a llengua estrangera, anàlisi d'errors.

1. INTRODUCTION

A substantial and expanding body of research has explored the acquisition of the L2 English article, yet the acquisition of articles in other L2 languages, such as Spanish, has received comparatively little attention. While there exist some studies dedicated to the acquisition of articles in Spanish L2 (AASL2), it is only in the last two decades that a handful of research projects have begun to concentrate on this area, but not much has been analysed about the acquisition of articles by learners who speak Spanish as their L3. This study delves into the use of Spanish articles by Estonian students learning Spanish as their third language, a topic that has received little attention to date, despite evidence that it represents one of the greatest challenges they face, and it appears that these difficulties persist even at higher levels of Spanish proficiency (Kruse 2018: 126).

This article aims to provide a comprehensive overview of how Estonian students use Spanish articles. To do so, we use the *Corpus Longitudinal de Español de Aprendientes Estonios - CLEAE* ('Estonian Learners of Spanish Longitudinal Corpus') (Rapún Mombiela 2023), a learner corpus which has a subsection of 245 written texts produced by 15 Estonian students of Spanish as L3. Our study underlines that errors found in the corpus show similar results to previous studies, with omission being the most common error, and no significant longitudinal improvement of the use of articles. Moreover, our results highlight the relevance of analysing morphosyntactic errors. The results of this study contribute to a deeper understanding of the specific difficulties faced by Estonian learners and the broader landscape of article acquisition in a third language context.

2. LITERATURE REVIEW

The efforts undertaken thus far in the field of L2 Spanish article acquisition stem predominantly from heterogenous initiatives and lack a uniformity of methodologies and theoretical frameworks. This diversity in approach has led to a rich tapestry of findings, yet it simultaneously poses challenges for synthesising results and drawing comprehensive conclusions about the learning and acquisition processes. To draw more generalisable conclusions about the acquisition of Spanish articles by L2 learners within structured learning

environments,¹ this literature review includes studies in which Spanish is taught as a foreign language. Additionally, the informants in these studies share homogeneous characteristics, primarily being adults who share the same first language (L1).

The exploration of the use and acquisition of Spanish articles has been conducted through two distinct methodological approaches: experimental methods (Barrera Rodríguez 2008a, 2008b; Cuza *et al.* 2013; Fernández López 1988; Furukawa & Okami 2016; Ionin *et al.* 2013; Kirova & Camacho 2021; Kobashi 2018; Montrul & Ionin 2012; Musto 2018; Nakagawa 2004; Prodanovic *et al.* 2023; Santiago Alonso 2017a, 2017b; Testa 2020; Xiuchuan 2015, 2017, 2018) and learner corpus research (Kruse & Rapún Mombiela 2023; Lee *et al.* 2018; Lin 2003; Rapún Mombiela & Tramallino 2023; Rubinjoni 2017; Said-Mohand 2007; Santiago Alonso 2009; Shen 2012; Testa 2019, 2020; Valverde & Ohtani 2014; Yonekawa & San Mateo Valdehita 2022).

In the context of Second Language Acquisition (SLA), experimental methods refer to a range of systematic and controlled techniques used to investigate how individuals learn a second language. These methods are designed to test specific hypotheses about the processes and factors that influence L2 learning. They often involve manipulating one or more variables to observe their effects on language learning outcomes.

In the domain of AASL2, various experimental methodologies have been employed: to shed light on this complex area. Grammaticality or acceptability judgement tasks, used to assess individuals' intuitive understanding of the correctness of sentences in a language, have been widely applied (Barrera Rodríguez 2008a, 2008b; Ionin *et al.* 2013; Kirova & Camacho 2021; Kobashi 2018; Montrul & Ionin 2012). Truth value judgement tests, employed to assess comprehension, particularly in the domain of semantics and pragmatics, were used by Ionin *et al.* (2013) and Montrul & Ionin (2012). Other methodologies employed a context (Cuza *et al.* 2013) or used felicitousness tasks (Ardura 2004), designed to assess a participant's ability to judge the appropriateness or felicity of a linguistic element within a given context.

The study of Ardura (2004) introduced a listening comprehension task and examined the effect of the mass/count noun distinction and verb type on intermediate L1-English L2-Spanish learners' interpretation of definite plurals, using both psychological and non-psychological verbs to neutralise syntactic

¹ It is imperative to emphasise that studies on article acquisition not based on learner output, presenting descriptive and contrastive approaches to interlanguage in a theoretical manner, are excluded from this study. For those interested in a more descriptive approach, the following sources are recommended: Bop (2023), Garachana (2008), Montero Gálvez (2011, 2014), Rothman *et al.* (2018), Santiago Alonso (2010, 2013, 2014, 2016, 2017b, 2022), Yang (2000). Additionally, other works focusing on the general analysis of the entire Spanish interlanguage have also been omitted. However, for those interested in this broader perspective, these are the recommended references: Alarcón Moreno (2013), Fernández Jódar (2006), Nergiz (2017), Kruse (2018).

differences. Furthermore, the remainder of research conducted using experimental methods has incorporated article-cloze tests or gap-fill exercises, which involve asking participants to fill in blanks with appropriate articles, thereby assessing their mastery of article usage in various contexts.

A corpus constitutes a collection of spoken or written texts, or both, generated in natural contexts, encapsulating a language or a linguistic variety in its entirety or specific aspects. These texts are digitally stored and encoded for rigorous scientific analysis (Rojo 2014: 371). The distinguishing feature of a learner corpus, as opposed to a general corpus, lies in its composition of electronic texts from second language learners, created in natural or near-natural settings (Granger 2008: 338). This differentiation pertains to the classroom's confined setting, where students are limited to performing a narrow range of functions (Gilquin 2015: 9).

In contrast to experimental methods, learner corpora offer a fundamentally different approach to understanding language acquisition. The most significant difference lies in the nature of the language samples being analysed, since learner corpora use samples from free language production. In these corpora, learners are not directed towards any particular aspect of the language. Instead, they produce text with a communicative purpose that extends beyond mere linguistic objectives. This approach allows for the observation of language use in a more authentic, unguided context, offering insights into how learners naturally use and process the language. Such corpora capture the spontaneous, real-world application of language, which can reveal patterns and nuances that controlled experiments might overlook. Consequently, the integration of learner corpora and experimental methodologies offers a holistic approach to SLA research, each complementing the other to provide a comprehensive understanding of language learning processes in diverse contexts.

Learner corpus studies frequently focus on examining accuracy, which refers to the extent of deviation from a specific language norm (Housen & Kuiken 2009: 3). Accuracy is operationalised through Error Analysis (EA), which relies on the systematic examination of learners' errors (Thewissen & Anishchanka 2022: 213). This approach enables researchers to not only identify and categorise errors, as well as but also to infer the underlying processes affecting language acquisition.

In AASL2, there is a substantial body of research particularly focusing on learners whose L1 lacks articles, such as Japanese (Barrera Rodríguez 2008a, 2008b; Fernández López 1988; Furukawa & Okami 2016; Kobashi 2018; Nakagawa 2004; Valverde & Ohtani 2014; Yonekawa & San Mateo Vadehita 2022), Chinese (Lin 2003; Shen 2012; Xiuchuan 2016, 2017, 2018) and Korean learners (i.e., Lee *et al.* 2018). Additionally, there have been studies involving Slavic languages users, such as Slovenian (Santiago Alonso 2009, 2017a), Polish (Tarrés Chamorro 2002; Testa 2019, 2020; Zimny 2017), as well as Serbian (Prodanovic *et al.* 2023, Rubinjoni 2017) and Russian (Denst 2006). Most recently,

the field has expanded to include research on Estonian learners,² a group that has only just begun to be explored (Kruse 2018, Nikitina 2019, Rapún Mombiola & Tramallino 2023, Kruse & Rapún Mombiola 2023).

Languages with articles have been less frequently studied. There have been investigations into Spanish article acquisition by L1 English users (Ardura 2014, Cuza *et al.* 2013, Ionin *et al.* 2013, Kirova & Camacho 2021, Montrul & Ionin 2012). However, studies focusing on users of Romance languages,³ which typically possess articles, are particularly rare.

Fernández López (1988) offers a contrast by examining the Spanish article acquisition among student groups with different L1 backgrounds, including Japanese, German, Arabic, and French. His results showed that Japanese students faced more pronounced challenges, continuing to make errors even at advanced levels of proficiency. Fernández López attributed these persistent difficulties primarily to language transfer, highlighting the significant impact of L1 structure, especially in languages without articles. However, literature underlines that L1 as well as L2 have an impact when acquiring the use of articles in Spanish. Musto (2018) found that students often selected options in line with the syntax of their L2, English, even though their L1, Italian, is typologically closer to Spanish. This indicates a significant influence of their L2 in the learning process.

In regard to the accuracy of Spanish article use, speakers of Slovenian (Santiago Alonso 2009), Serbian (Rubinjoni 2017), Polish (Tarrés Chamorro 2002; Zimny 2017; Testa 2019, 2020), English (Said-Mohand 2007), and Japanese (Yonekawa & San Mateo Valdehita 2022) have been explored. These studies point out that differences between L1 and Spanish influence the omission or incorrect use of definite and indefinite articles in Spanish. Moreover, it seems that as learners progress, there is a decrease in errors (Said-Mohand 2007; Testa 2019, 2020; Zimny 2017; Yonekawa & San Mateo Valdehita 2022), though this trend is not always linear and may fluctuate. The relationship between articles and other determiners is complex, with a tendency for learners to use demonstratives and possessives to express referentiality (Lin 2003; Rubinjoni 2017; Barrera Rodríguez 2008a, 2008b). Despite these challenges, there is compelling evidence indicating that learners progressively improve their command of Spanish articles as they achieve higher proficiency levels, though the pace and degree of improvement

² The Estonian language belongs to the non-Indo-European, Finno-Ugric branch of the Uralic language family.

³ Please note that this study does not include findings from a separate investigation into the acquisition and use of articles by students with Romanian as their first language (L1) and either Catalan or Spanish as their second/third languages. This exclusion is due to the fact that the participants in the referenced study are minors and are also in an immersion situation, being immigrants. For those interested in this specific area of research, the study by Popa (2016) is recommended for further reading.

notably diverge among different learner groups (Rubinjoni 2017, Montrul & Ionin 2012, Cuza *et al.* 2013).

However, errors, particularly omissions, are not only prevalent but also tend to persist even at advanced stages of language proficiency (Rubinjoni, 2017, Santiago Alonso 2016). Furthermore, the comprehension of the necessity of articles is generally low, with learners often relying on context to identify noun references. These errors tend to fossilise over time, showing minimal resolution despite prolonged study (Reymóndez Fernández & Santiago Alonso 2004).

In the case of Estonian learners of Spanish, Rapún Mombiela & Campillo (2023) and Kruse & Rapún Mombiela (2023) applied EA to assess the accuracy of articles and prepositions in temporal expressions in Spanish produced by Estonian students. Additionally, Kruse & Rapún Mombiela (2023) extended their research to include the study of fluency⁴ in written texts, utilising a keylogger for detailed analysis. The outcomes of both studies indicated a significant improvement in article use as the Estonian students progressed in their Spanish language proficiency. Furthermore, it was found that students with a higher level of proficiency in English tended to produce fewer deviant structures compared to those with lower English proficiency. The fluency analysis revealed a direct correlation between the level of linguistic competence in English and smoother transitions in Spanish discourse among the students. This suggests a heightened metalinguistic awareness and greater precision in the use of functional words.

In certain instances, Spanish has been approached as a second language/foreign language (Barrera Rodríguez 2008a, 2008b; Cuza *et al.* 2013; Fernández Jódar 2006; Fernández López 1988; Ionin *et al.* 2013; Kirova & Camacho 2021; Lee *et al.* 2018; Lin 2003; Montrul & Ionin 2012; Nakagawa 2004; Said-Mohand 2007; Shen 2012; Valverde & Ohtani 2014; Xiuchuan 2015, 2017, 2018; Yonekawa & San Mateo Valdehita 2022) while other studies have acknowledged that students of Spanish often have prior experience with learning another foreign language before beginning their Spanish studies, thus treating Spanish as a third language (Furukawa & Okami 2016; Kruse & Rapún Mombiela 2023; Musto 2018; Prodanovic *et al.* 2023; Rapún Mombiela & Tramallino 2023; Santiago Alonso 2009, 2017a, 2017b; Testa 2019, 2020).

This shift towards treating Spanish as an L3 underlines the progression in the evolution of Second Language Acquisition (SLA) theories embodied by Third Language Acquisition (TLA), which considers not only the learner's first language (L1) but also the entire linguistic spectrum that the student has been exposed to throughout their life. There exists a broad consensus among scholars that any prior linguistic knowledge can influence the learning of an additional language

⁴ Fluency pertains to the ability to produce discourse smoothly and continuously, without significant interruptions. Historically, fluency has been primarily associated with spoken language (Thewissen & Anishchanka 2022: 210), but over the past decade, its importance in writing processes has been increasingly acknowledged, recognizing its role in assessing competence in a L2.

(Aribaş and Cele 2021, Cenoz 2001, De Angelis 2007, Odlin 2003, Williams & Hammarberg 1998). Thus, the existing knowledge learners possess regarding language use plays a crucial role in understanding TLA. This is due to the fact that learners beginning L3 learning have already attained an understanding of complex concepts and ideas expressed in their first and second languages (McManus 2021: 8). Scholarly works on TLA have disclosed that, in the case of multilingual learners, a significant portion of language transfer (particularly at the foundational levels) appears to originate from the L2 (De Angelis 2007, Williams & Hammarberg 1998).

Addressing the findings of such divergent studies, each anchored in distinct perspectives and methodologies, poses a significant analytical challenge. However, it is possible to discern a coherent narrative amidst this diversity. Notably, the impact of the first language of learners, especially from languages devoid of articles, emerges as a pivotal element that consistently introduces prolonged obstacles in mastering article use in Spanish (Denst 2006, Fernández López 1988, Rubinjoni 2017).

In AASL2, research has predominantly used learner corpora with written samples. Exceptions to this norm include the studies by Said-Mohand (2007) and Tarrés Chamorro (2002).

Regarding the classification of errors,⁵ we can discern four types of taxonomy, contingent upon the research questions posed: 1) descriptive, when elucidating the nature of the error from a purely linguistic perspective (omission, addition, inappropriate selection, morphological errors); 2) syntactic, when it encompasses information about the syntactic function fulfilled by the analysed noun phrase in the sentence, such as subject, for example; 3) functional or usage-based, when it delineates the type of relationship established between the article and its referent: generic, anaphoric, deictic, associative anaphoric, non-anaphoric, or cataphoric; 4) etiological, in which the cause of the error is described, which can be either interlinguistic, involving some form of transfer from another language known to the learner to some degree, or intralinguistic, involving simplification, overgeneralization or other aspects of language rules.

Although these classifications are not mutually exclusive and are perfectly compatible, researchers tend to work either solely with the descriptive or with a combination of the descriptive and another. An example of the combination of descriptive and syntactic analysis can be found in Santiago Alonso (2009), and of the descriptive and functional analysis in Barrera Rodríguez (2008), Rubinjoni (2017), and Xiuchuan (2015, 2017, 2018).

Regarding the etiological analysis, the research landscape is diverse, with some investigations focusing on crosslinguistic influence (Cuza *et al.* 2013; Fernández López 1988; Furukawa & Okami 2016; Ionin *et al.* 2013; Kruse & Rapún

⁵ For a comprehensive taxonomy of errors in Spanish interlanguage, it is recommended to consult Alexopoulou (2006).

Mombiela 2023; Lee *et al.* 2018; Lin 2003; Montrul & Ionin 2012; Musto 2018; Prodanovic *et al.* 2023; Rapún Mombiela & Tramallino 2023; Rubinjoni 2017; Testa 2019, 2020; Xiuchuan 2015, 2017, 2018). However, intralinguistic factors behind errors have not been as thoroughly investigated, indicating a clear area for further study.

While many authors propose some form of error taxonomy, it is challenging to find studies that justify their taxonomy with a comprehensive report on the intercoder reliability process. In many cases, it is not clear whether the study in question has been conducted by more than one annotator to ensure reliability. This lack of clarity raises questions about the consistency and reproducibility of the findings. A notable exception to this trend is the work of Valverde & Ohtani (2014), who proposed an annotation scheme based on a rigorous process of design and evaluation. Researchers have primarily focused on the causes related to interlanguage errors, including the aforementioned crosslinguistic influence and intralinguistic factors (cases of reduction to a simplified system or overgeneralization).

A minority of AASL2 studies have incorporated a semantic dimension into their analysis, adding depth to the understanding of article use and acquisition. In particular, two semantic parameters have been identified as pivotal in understanding these linguistic phenomena: the Nominal Mapping Parameter (NMP) (Chierchia 1998) and the Article Choice Parameter (ACP) (Ionin *et al.* 2004). Cuza *et al.* (2013) examine how advanced English-speaking learners of L2 Spanish grasp the semantic properties of Spanish definite plural and bare plural subjects. They address the transfer from English to Spanish, which poses a learnability challenge due to differences in how the two languages use articles with plural noun phrases for generic or specific interpretations. The study suggests that learning difficulties might be mitigated by recasting the subset/superset relationships in terms of the NMP settings for [+argument] and [+predicate]. It proposes that advanced learners might need substantial exposure to the target language to overcome initial L1 specifications and achieve full NMP resetting.

Building on the Article Choice Parameter framework, Furukawa & Okami (2016) and Lee *et al.* (2018) predict that the use of Spanish articles by Japanese students (whose L1 does not have articles) will fluctuate between an overuse of definite articles in indefinite contexts and a use more closely aligned with the target language. Both studies agree that proficiency level in Spanish will allow for improvement in accuracy and the ability to produce language segments more closely resembling those of the target language. However, in the discussion section both studies refrain from attributing the excessive use of definite articles in indefinite contexts to a misalignment of a semantic parameter. Rather, they suggest that the influences of the learners' L1 and L2, along with a shortfall in specialised instruction on article use, are the primary factors of the errors. In contrast, Xiuchuan (2015), in her study of L1 Chinese learners, does attribute the initial unstable state in the use of articles to the fluctuation parameter, wherein

the concept of specificity plays a significant role. This is identified as the foremost cause in the emergence of deviant structures.

This review underscores the imperative for further research that delves into semantic frameworks and advocates for refined error annotation methodologies underpinned by rigorous inter-coder reliability. It also highlights the necessity for a deeper examination of both inter- and intralingual error causes and emphasises the importance of specialised instructional strategies explicitly designed to address the nuances of article use.

3. STUDY GOALS

This study aims to provide an initial overview of the use of articles by Estonian students of Spanish. After reviewing the existing literature in the field of Spanish article acquisition, we aim to verify whether, as other studies have shown, omission errors are also the most common in the interlanguage of Estonian students and whether these errors persist over time or are overcome as a more advanced language level is reached. To this end, a longitudinal learner corpus will be used, and the annotation scheme will be tested with a rigorous inter-rater validation analysis to qualify it for use in future studies. As this is a preliminary approach, we will not focus on a single feature but will instead aim to provide as detailed a description as possible of the use of the Spanish article over time by Estonian students.

4. DATA AND SAMPLE DESCRIPTION

The longitudinal corpus used for this study is a subsection of 345 written texts from the *Corpus Longitudinal de Español de Aprendientes Estonios - CLEAE* ('Estonian Learners' of Spanish Longitudinal Corpus') (Rapún Mombiela 2023). The set of samples used in this study contains a total of 58.729 words and 65.690 tokens. Of these, a total of 14.440 items have been analysed. These items consist of all the noun phrases in the corpus as well as those categories that would require an article, such as some adverbial or prepositional phrases.

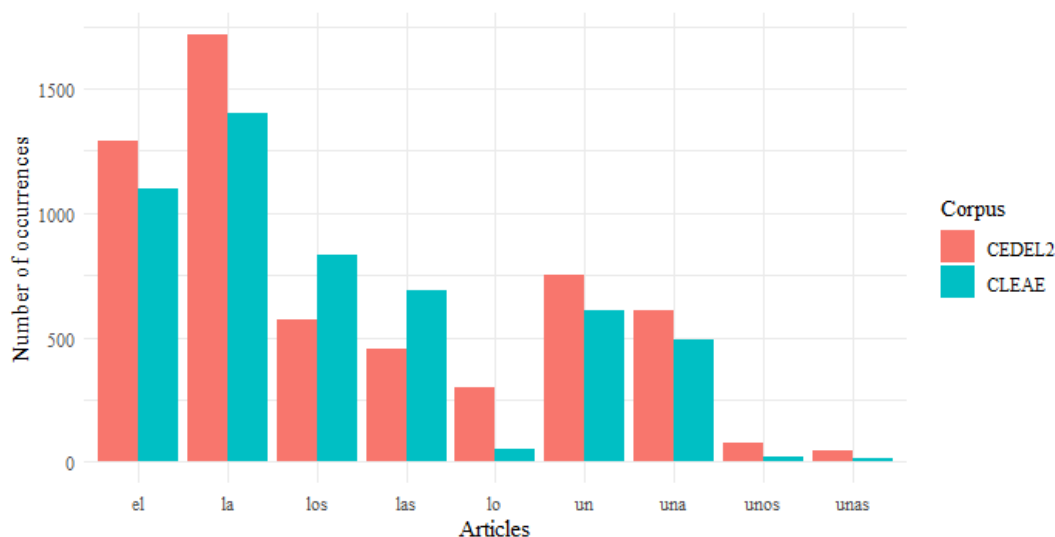
The productions come from 15 Estonian students (all women) of Spanish during their first two years of their degree in Romance Studies, specialising in Spanish. The participants' ages range from 19 to 38 years. The only L1 for all participants is Estonian. They all have high proficiency in English (6 participants have a B2 and 9 participants have a C1 level) and consider this language their dominant L2. They all have some experience studying other foreign languages (German, Russian, Finnish, Swedish, French, and/or Italian). At the beginning of their studies, 6 of the 15 participants had previous experience learning Spanish, and 9 did not. Students of Spanish Studies at the University of Tartu with prior knowledge of Spanish take a placement test before starting the first course. All

the participants in this study who took the test were placed at a false beginner level (i.e., they had some knowledge of Spanish but did not reach level A1), except for one student who demonstrated a slightly higher level and was assigned level A2.1. None of the participants had spent more than a month living in a Spanish-speaking country.

5. METHODOLOGY

In a study where the information comes from a learner corpus, the first step is to verify whether a certain feature has a different frequency of use in the students' language compared to L1 speakers' language. If this comparison shows that the feature is indeed less or more frequent in the interlanguage, it may indicate underuse or overuse due to transfer effects (Osborne 2015: 339) or intralinguistic causes. To do this, we compared our corpus with a selection of texts from the L1 Spanish corpus from CEDEL2 (Lozano 2022),⁶ a large database containing the language produced by learners of Spanish as a second/foreign language (L2). For comparative purposes, CEDEL2 also contains data from native speakers whose L1 is Spanish. Below, we present a comparison of article usage in the two corpora.⁷

Graph 1 shows the absolute number of occurrences of each grammatical form in both corpora. It is evident that the L1 Spanish corpus (CEDEL2) has a significantly higher number of occurrences for almost each form compared to the L3 Spanish corpus (CLEAE). The form *la* is the most frequent in both corpora, but its frequency is much higher in the L1 corpus.

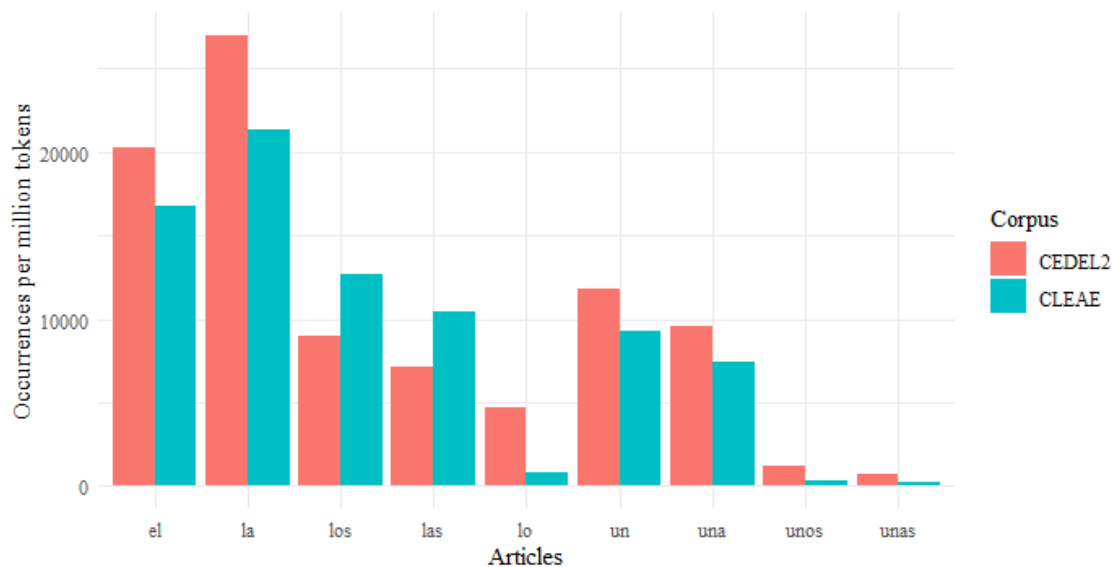


GRAPH 1. Number of occurrences of articles in CLEAE vs. CEDEL2.

⁶ For this study, we selected texts produced by women with Spanish as their L1, aged between 19 and 30 years, and only the descriptive texts that closely resembled those produced by our informants.

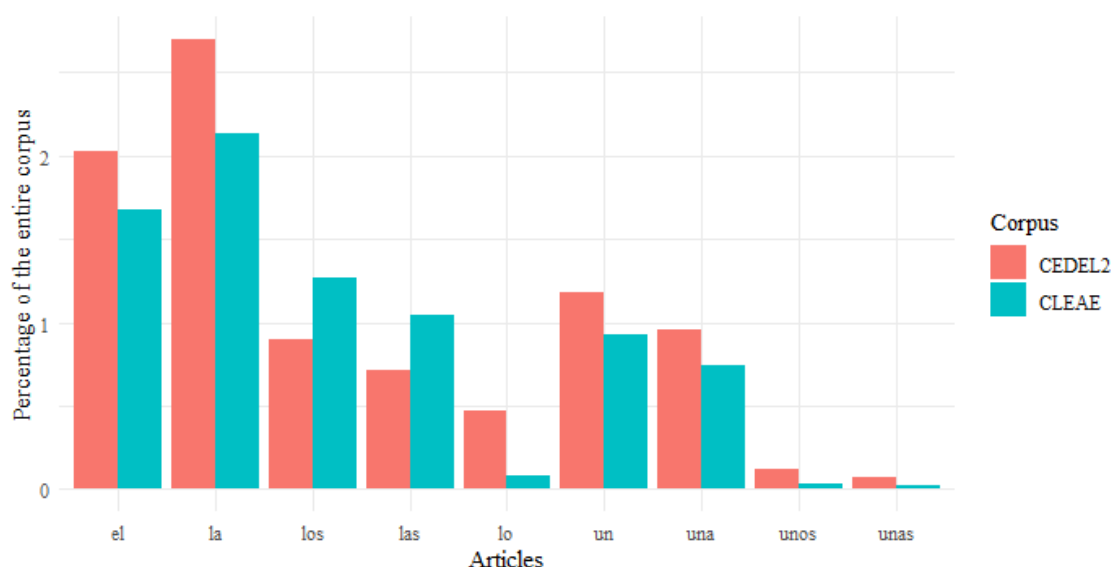
⁷ To carry out the comparison, we used the Sketch Engine tool (Kilgarriff *et al.* 2014).

Graph 2 presents the relative frequency per million tokens for each grammatical form. The forms *la* and *el* remain very frequent in both corpora. However, other forms like *lo*, *unos*, and *unas* have a significantly lower relative frequency in L3. The L3 corpus shows lower values compared to the L1 in all the forms except from *los* and *las*.



GRAPH 2. Occurrences per million tokens in CLEAE vs. CEDEL2.

The third graph represents the percentage each form constitutes of the entire corpus. The forms *la* and *el* make up a larger percentage of the L1 corpus (3,69 % and 2,02 %, respectively) than the L3 corpus (2,13 % and 1,67 %, respectively). This indicates that experienced users of Spanish use these forms more consistently in the natural flow of discourse.



GRAPH 3. Percentage of the entire corpus in CLEAE vs. CEDEL2.

After comparing the two corpora, an underuse of the article by Estonian informants can be confirmed. Interestingly, while the use of plural definite

articles was more frequent among these students —unexpectedly so— this also highlighted a noteworthy difference in this regard.

5.1. Annotation system

This section discusses an annotation scheme that would allow researchers to categorise the students' use of the article as comprehensively as possible. To do so, each item was annotated with a series of labels. Firstly, the category of the linguistic element that the noun accompanied was defined, followed by the syntactic function that the phrase containing the article had in the sentence. The type of article used was then tagged, along with an indication of whether it was correct or incorrect. If the article was incorrect, the cause (interlinguistic or intralinguistic) was specified, and if it was interlinguistic, the source of the transfer was assigned.

5.2. Inter-coder reliability

The reliability of the scheme was tested⁸ on different levels: modified by the article linguistic element categorization, error identification, error labelling, and error correction. This approach had multiple purposes, such as validating the annotation scheme, identifying ambiguities in the source texts, or adequately evaluating interpretations (Artstein 2017: 2).

First, the three coders conducted a preliminary study in which they coded 60 items. Subsequently, they met for a joint qualitative analysis, discussing each of the categories and the discrepancies found. In this initial analysis, it became quite evident that the divergences, in most cases, laid in understanding whether the errors had an interlinguistic or intralinguistic nature. Next, the three coders proceeded to analyse 436 items. The selection of these items considered examples of all article usages intended for study, in all possible syntactic contexts, as well as examples of all errors represented in the corpus. Language samples were selected from all tasks and all participants.

Once the coding was completed, a quantitative study was conducted. For this, each variable under study was analysed using cross tables, which allowed for comparing the values corresponding to the relative assignment according to the different possible response categories per column. From this comparison, along with Chi-Square tests and effect size measured through Cramer's *V*, it was determined whether there was statistical evidence to confirm that the measurements from different observers were in agreement. The results of the statistical analysis of the inter-observer validation are presented below.

First, the categorisation of the noun was analysed. This analysis yielded disparate results (the *p*-value was .001 with a χ^2 statistic of 32.00, indicating a

⁸ This inter-rater validation analysis was conducted as part of the first author's doctoral project, and a more detailed version of it can be seen in her doctoral thesis once it is published.

highly significant alternative result), unexpected to the coders, who thought there was strong consensus regarding the nature of nouns, as shown in table 1.

Noun	Coder		
	Cod 1	Cod 2	Cod 3
Uncountable	20,6 % (90)	11,1 % (48)	20,9 % (91)
Singular countable	49,8 % (217)	59,7 % (259)	48,2 % (210)

TABLE 1. Categorisation of nouns as uncountable or singular countable according to the three coders.

As can be seen in this table, Coder 2 did not categorise singular countable nouns and uncountable nouns in the same way as Coders 1 and 3. Therefore, a small investigation was needed to understand this discrepancy. *Nueva Gramática de la lengua española* (Real Academia Española 2016: 1544-1552) proposes various semantic classifications for uncountable nouns, but there is no consensus among them. This lack of agreement is due to the difficulty of establishing precise boundaries between subgroups and the fact that existing classifications do not cover all uncountable nouns in Spanish. Additionally, uncountable nouns can be both abstract and concrete, which adds another layer of complexity to their classification. In Spanish, some nouns can be used as both countable and uncountable in different syntactic contexts, which is a notable characteristic of the language. This linguistic flexibility allows the same word to adapt according to the communicative need, without strictly following the typical rules of countable or uncountable nouns in other languages. Referring to our corpus, we found nouns, as highlighted in the following example, that were categorised as uncountable by Coders 1 and 3, but Coder 2 considered them countable: *Es pasta con crema y con pollo.*

A noun like *crema* ('cream') can be difficult to categorise and may also be interpreted as a countable noun since there are contexts in which it can be used as such (Davies 2016). However, working with a methodology as EA, it is necessary for coders to always consider the context in which the words under study are found. Thus, in the example at hand, we would label the word *crema* as an uncountable noun because in the context it is used, it denotes a substance (Real Academia Española 2016: 1545). *Nueva gramática de la lengua española* (Real Academia Española 2016: 1529-1543) gives certain guidelines that can help discern the countable or uncountable nature of a noun. By following these guidelines, the coders re-coded all the problematic items and achieved consensus. The quantification of errors was a relatively straightforward variable to code, as identifying errors related to the use of articles at beginner and intermediate levels does not pose a problem for experienced users of Spanish. In fact, the statistical analysis ($\chi^2 = 5,81$, $p = .214$, $V = .047$) showed that there were no significant statistical differences between the three coders. However, when categorising the type of error, significant differences were found ($\chi^2 = 32,25$, $p = .004$). Although the effect size indicated that these differences were not large

($V = .177$), the coders decided to discuss the discrepancies. The main divergence was caused by errors involving the presence of an unnecessary definite article. The first coder had categorised this type of error on 44 occasions, while the second had assigned this error to 34 items, and the third to only 26.

After reviewing the items in context, the annotators concluded that they had not interpreted them in the same way and decided to follow a principle of minimal change as recommended by Valverde & Ohtani (2014: 239): sentences in which the student's choice is acceptable, even if not the best option, should not be marked as errors. That is, the annotator's goal should be to produce an acceptable result rather than a perfect one. If the coder cannot decide whether an item contains an error or not, it should be considered correct. With this adjustment in mind and after a new coding, an optimal result was obtained, and the inter-rater validation analysis ($\chi^2 = .713$, $p = 8.88$) showed that there were no significant statistical differences from the perspective of the coders.

The etiological study of errors was more complicated, and the coders could not reach an agreement. The main problem lies in the difficulty of discerning the error cause of a bare noun phrase. On the contrary, in cases where the student incorporates an unnecessary article into the phrase, it is easier to understand whether transfer has occurred or not. The inter-rater validation analysis, as we can see below in table 2, yielded unacceptable results.

Cause	Coder			
	Cod 1	Cod 2	Cod 3	
Intralinguistic	62,2 % (107)	31,5 % (53)	37,0 % (64)	43,7 % (224)
Interlinguistic	37,2 % (64)	59,5 % (100)	63,0 % (109)	53,2 % (273)
Unclear	0,6 % (1)	8,9 % (15)	-	3,1 % (16)
<i>total</i>	33,5 % (172)	32,7 % (168)	33,7 % (173)	(513)
	P-value	Est.	Result	
χ^2	.001	85,13	highly significant	
	value			
V	.396		moderate	

TABLE 2. Results of the inter-rater validation analysis.

Upon closer examination of the samples, we realised that it was difficult to discern those instances where the error was due to the transfer from the L1/L2 or to the simplification or lack of knowledge of L3 rules. Scientific investigation of transfer has found it to be a challenging phenomenon to precisely identify (Osborne 2015: 334). We decided to differentiate between cases where English differs from Spanish in that the former does not use an article while the latter does, and other cases where both English and Spanish do use the article. If the analysed item was in a context that had been provided during instruction, and where English does not use an article but Spanish does, we concluded that crosslinguistic influence must be very strong. This is because if a person speaks two languages that do not use articles in that context, it is not surprising that they

omit the article in their third language as well. Therefore, we marked these instances as interlinguistic errors. In contrast, for omission errors where English also uses an article, we labelled them as “unknown cause” because we cannot determine whether they are due to L1 influence or a lack of knowledge of the L2 and L3 rules.

5.3. Analysis

The tabulation, sorting, and cleaning of the database were performed using Microsoft Excel. The statistical analysis was conducted with the software application IBM SPSS Statistics version 28. The statistical techniques and tests employed were:

1. Descriptive analysis of qualitative variables using frequency and percentage tables for the cross-tabulation of two such variables.
2. For the cross-tabulation of categorical variables, contingency tables were used along with the Chi-square test of independence and the estimation of effect size using Cramer’s V coefficient.
3. To measure the time variable (i.e., the difference in error rates across semesters), we used adjusted rates and applied the Kruskal-Wallis test.

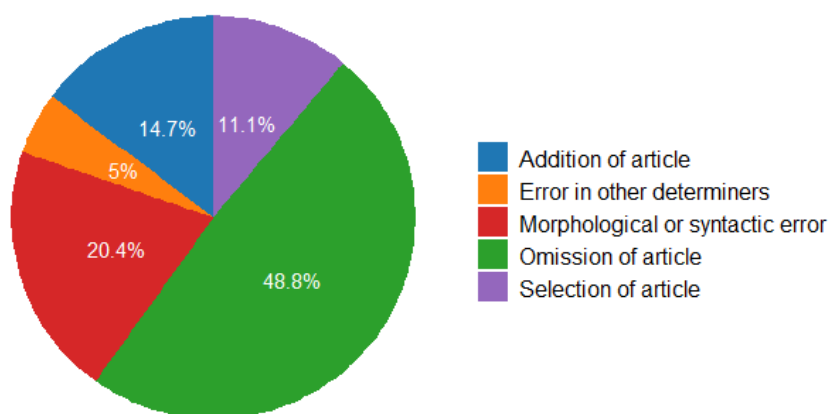
6. Results

We begin by offering a general descriptive analysis. Table 3 indicates that 20,8 % of the analysed structures contain some type of error that renders the construction unacceptable.

Category		Frequency
Type of determiner	Definite article	30,4 % (4.384)
	Indefinite article	7,8 % (1.124)
	Absence of article	43,1 % (6.230)
	Other determiner	18,4 % (2.658)
	Other word	0,3 % (44)
Type of linguistic item	Singular countable noun	36,8 % (5.314)
	Plural countable noun	30,40 % (4.390)
	Uncountable/abstract noun	16,71 % (2.413)
	Proper noun	10,31 % (1.489)
	Other category	5,78 % (834)
Acceptability	Acceptable	79,2 % (1.1439)
	Unacceptable	20,8 % (3.001)

TABLE 3. Categories of errors and frequency of appearance.

Of the 3,001 unacceptable instances, the most common error is omission, as shown in graph 4.



GRAPH 4. Distribution of error types in unacceptable instances.

The high incidence of omission errors was pointed out in previous research in the field of Spanish article acquisition (Rubinjoni 2017, Reymóndez Fernández & Santiago Alonso 2004, Santiago Alonso 2016). However, the morphological and syntactic errors are noteworthy, as such a high incidence was not expected.

Both Estonian and English lack gender marking, so gender errors could be attributed to interlinguistic influence. This hypothesis is further supported by the fact that these errors are the most frequent among the morphological and syntactic ones, as shown in table 4.

Morphosyntactic errors 100% (613)	Gender errors 79,77 % (489)	Definite article 35,23 % (216)
		Indefinite article 14,85 % (91)
		Other determiners 26,69 % (182)
	Number errors 10,6 % (65)	Definite article 6,68 % (41)
		Other determiners 3,92 % (24)
	Word order errors 9,62 % (59)	Definite article 1,63 % (10)
Other determiners 7,99 % (49)		

TABLE 4. Types and frequency of morphosyntactic errors.

However, when examining students' productions, it becomes clear that the issue does not always stem from transfer. Instead, it often arises from the overgeneralisation of certain rules, either by analogy or due to a lack of knowledge of grammatical norms, as seen in the following examples.

- (1) Por ejemplo recuerdo que {*el sál ~ la} era muy caro. ('For example, I remember that salt was expensive'.)
- (2) Y en {*los ~ las} manos de chica es un arma. ('And in the girl's hands there is a weapon'.)

- (3) Es {*un ~ una} celebración muy alegre y un poco triste porque es {*la ~ el} última día.
(‘It is a very joyful celebration and a bit sad because it is the last day’.)

Thus, it seems that difficulties related to gender are caused by a dual influence, both interlinguistic and intralinguistic, which explains the high number of gender errors in the corpus. While gender errors are numerous enough to warrant attention, it is important to emphasise that they do not necessarily stem from a lack of understanding in the use or interpretation of articles. The learner must realise that the language they are acquiring includes inherent gender characteristics and must learn the gender of each new noun. Although L2 learners whose L1 has grammatical gender might reach the conclusion more quickly that the L2 is a gendered language, determining the gender of each noun is not necessarily straightforward.

In addition to realising that nouns have gender, learners must also learn that gender agreement occurs in articles and adjectives, along with the appropriate morphological realisations of gender (Bruhn de Garavito 2002: 22). This requires a certain amount of time for assimilation, practice, and adaptation to this feature of the Spanish language, regardless of the learner’s L1.

As shown in table 5, the highest frequency of errors occurs in the prepositional phrase term category, where 41,3 % of all errors are concentrated. The direct object and subject categories also show significant error rates, accounting for 26,9 % and 18,1 % of the errors, respectively. Omission errors are the most prevalent, comprising 48,8 % of all errors, but morphological and syntactic errors also represent a great challenge making up 20,4 % of the total

The chi-square test result ($\chi^2 = 277,39$, $p = .001$) indicates that the differences in error frequencies across categories are highly significant. This means that the observed distribution of errors is unlikely to be due to chance and reflects a real pattern in learner behaviour. Although the Cramer’s V value of .152 suggests a small effect size, the pedagogical implications of these findings cannot be denied. The high incidence of omission errors and the concentration of errors in the prepositional phrase term category highlight specific areas where learners struggle the most.

It has been demonstrated that prepositions are also a particularly problematic aspect for Estonian students (Kruse 2018, Kruse & Rapún Mombiela 2023, Rapún Mombiela & Tramallino 2023) since Estonian has very few prepositions and instead favours the use of postpositions (Kruse 2018: 118). Studies in this area have shown that difficulties with prepositions do not diminish over time. Prepositional phrases therefore present a dual challenge, as they contain two elements absent in Estonian: articles and prepositions. Transfer from English can be a negative influence in the case of prepositions, as these often do not have a direct translation. This is particularly true for locative and directional prepositions, of which English has an abundance to

express different nuances. Spanish, however, is much vaguer in this regard (Butt *et al.* 2019: 481).

Function	Type of error					
	Addition	Omission	Selection	Morph-synt.	Other det.	
Subject	2,7 % (12)	23,8 % (348)	11,1 % (37)	16,3 % (100)	30,5 % (46)	18,1 % (543)
Direct object	39,1 % (173)	19,5 % (286)	32,2 % (107)	35,6 % (218)	15,9 % (24)	26,9 % (808)
Circumst.	1,8 % (8)	1,1 % (16)	4,2 % (14)	3,9 % (24)	1,3 % (2)	2,1 % (64)
Attribute	7,9 % (35)	7,3 % (107)	6,6 % (22)	6,5 % (40)	.7 % (1)	6,8 % (205)
Prep.ph. term	43,7 % (193)	43,0 % (629)	40,4 % (134)	34,6 % (212)	46,4 % (70)	41,3 % (1238)
Modifier	.2% (1)	.2% (3)	.3% (1)	.2% (1)	-	.2 % (6)
Complement	2,5 % (11)	5,0 % (73)	5,1 % (17)	2,6 % (16)	5,3 % (8)	4,2 % (125)
Other	2,0 % (9)	.1 % (1)	-	.3 % (2)	-	.4 % (12)
<i>Total</i>	14,7 % (442)	48,8 % (1.463)	11,1 % (332)	20,4 % (613)	5,0 % (151)	(3.001)
	p-value	Statistic		Result		
χ^2	.001	277,39 **		Highly significant		
	value	Interpretation				
V	.152	Small				

TABLE 5. Frequency of errors by function.

We observed a high number of errors (366) caused by the omission of the article in noun phrases functioning as the term of a prepositional phrase due to interlinguistic transfer. It is not difficult to assume that if Estonian speakers must produce prepositions and articles, two linguistic elements that are absent in their native language, they will find this task challenging and will likely seek assistance from their source language, which in this case is English, a language that has both elements. We found numerous English-influenced structures missing an article, as seen in the following sentence: [...] *era posible chatear con mis amigos gracias a {*\emptyset ~ las} redes sociales* ('it was possible to chat with my friends, thanks to social media').

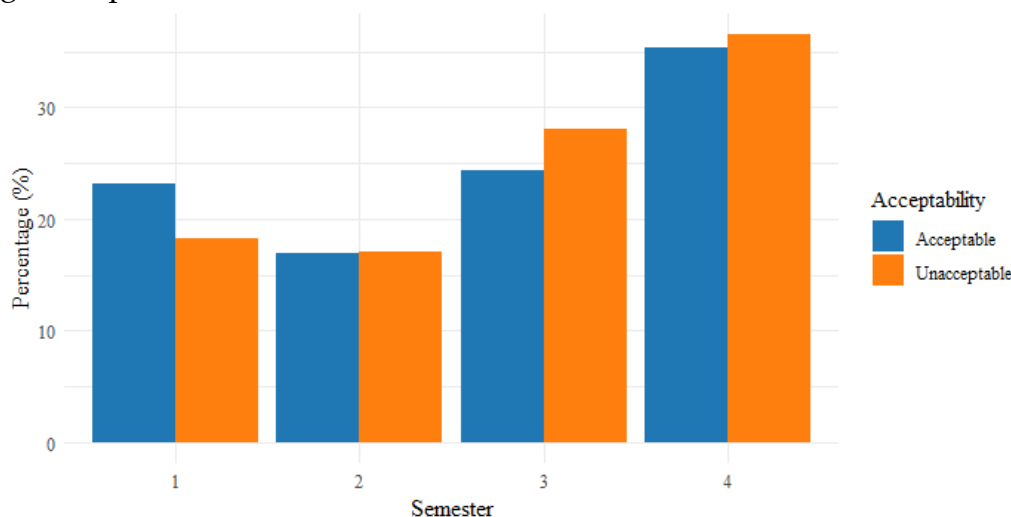
After identifying the most common types of errors, we proceeded to analyse the different semesters to determine whether there were any statistical differences that would allow asserting that changes occurred over time. Results are summarised in table 6.

Semester	Acceptability		
	Acceptable	Unacceptable	
1	23,2 % (2.656)	18,3 % (548)	22,2 % (3.204)
2	16,9 % (1.938)	17,1 % (514)	17,0 % (2.452)
3	24,4 % (2.793)	28,1 % (842)	25,2 % (3.635)
4	35,4 % (4.052)	36,6 % (1.097)	35,7 % (5.149)
<i>Total</i>	79,2 % (11.439)	20,8 % (3.001)	(14.440)
	P-value	Statistic	Result
χ^2	.001	39.76 **	Highly significant
	Value	Interpretation	
V	.052	insignificant	

TABLE 6. Acceptability of Spanish article structures by semester.

The chi-square test result ($\chi^2 = 39,76$, $p < .001$) indicates that the differences observed across semesters are highly significant. However, the Cramer's V value of .052 suggests that the effect size is insignificant, meaning that, while there is a statistically significant difference, the practical impact of these differences across semesters is minimal. This implies that the change over time in the acceptability of structures with articles is statistically noticeable but not strong enough to indicate a substantial shift in learner performance.

As seen in graph 5, the number of analysed expressions was higher in some semesters. This is due to two main reasons. Firstly, the nature of the tasks. In Semester 2, the tasks were descriptive, which encouraged the use of many noun phrases (Gutiérrez 2002: 80) but also led to the use of more stereotypical and simple structures. On the other hand, in Semesters 3 and 4, we observed a significant number of expository and argumentative tasks, which involved a greater presence of abstract nouns.



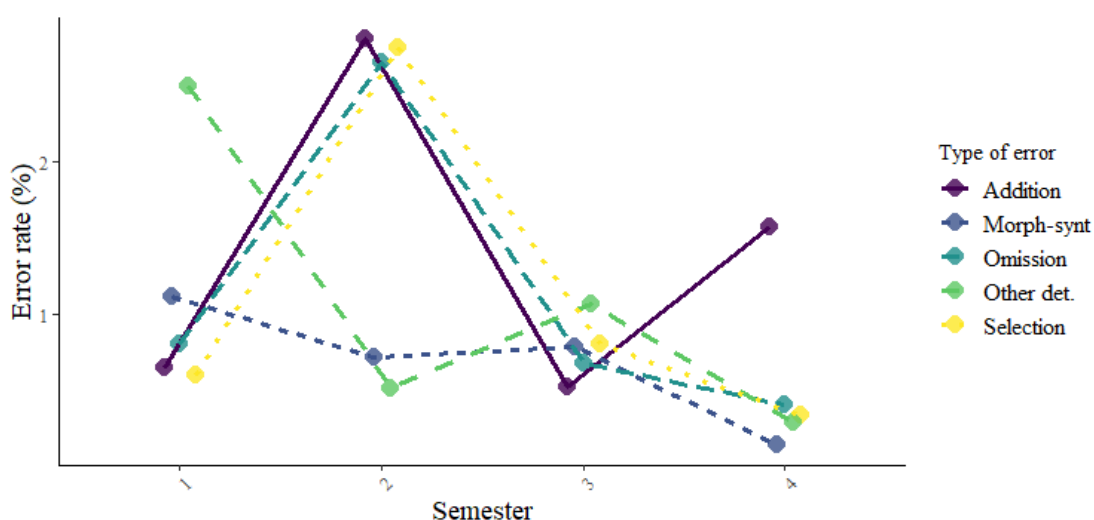
GRAPH 5. Acceptability of Spanish article structures by semester.

Secondly, as the students progressed in their language competence, they became increasingly capable of writing longer texts. This development meant that in the third and fourth semesters, we had a greater number of words and,

consequently, more structures to analyse. The growing length and complexity of students' writings reflect their advancing skills, which naturally led to an increase in the volume of data available, and in the number of errors because they were using more complex structures, in the later semesters.

Once the preliminary results from our contingency table were obtained, an analysis of adjusted error rates was applied to achieve a more accurate interpretation of the data. Since the number of words varies significantly between the different semesters, it was necessary to adjust the observed errors according to this factor to allow for fair comparisons between the groups. As graph 6 indicates, the error rates vary across the semesters, with certain fluctuations in some errors, such as addition or omission, but no clear or consistent trend indicating an increase or decrease in error rates as language competence progresses.

To compare the differences in adjusted error rates between the semesters, the Kruskal-Wallis test was applied. This non-parametric test is appropriate for comparing medians across multiple groups when the normality of the data cannot be assumed and when group sizes are unequal. The results of the Kruskal-Wallis test were as follows: $\chi^2 = 3$, $df = 3$, $p\text{-value} = 0.3916$. This indicates that there are no statistically significant differences in the adjusted rates between the semesters for each type of error. The p -value is higher than the common significance level of 0.05, suggesting that the null hypothesis cannot be rejected. The lack of statistical significance ($p > 0.05$) indicates that, after adjusting for the number of words, the error rates in the use of articles do not vary significantly across the four semesters. In practical terms, this suggests that Estonian students of Spanish are not making more or fewer article-related errors as their language competence advances.



GRAPH 6. Error rate by semester.

7. CONCLUSIONS

This study aims at conducting an in-depth study of the use of Spanish articles by Estonian students learning Spanish as their third language (L3). Although this topic poses major difficulties for Estonian learners of Spanish, it has received little attention to date. This article aims at achieving two main objectives. Firstly, it seeks to explain the methodologies and theoretical frameworks that have been employed in the field of Spanish article acquisition among L2/L3 learners. Secondly, it aims at providing a comprehensive overview of the usage of Spanish articles by Estonian students. By addressing these objectives, this study contributes a deeper understanding of the specific difficulties faced by Estonian learners and the broader landscape of article acquisition in a third language context.

Regarding the first goal, a thorough literature review is presented. It is claimed that the efforts undertaken thus far in the field of L2 Spanish article acquisition consist of heterogeneous initiatives that lack a methodological and theoretical uniformity. This overview reveals a need for further investigation in the field of semantic frameworks, and advocates for improved error annotation methodologies backed by rigorous inter-coder reliability. It also highlights the necessity of a deeper examination of both inter- and intralingual error causes and emphasises the importance of specialised instructional strategies designed to address the nuances of article use.

In order to attain the second goal, 345 written texts from the CLEAE corpus were analysed. This analysis shows that the most common error among learners is the omission of articles, in line with previous research. This finding is not exclusive to Estonian learners of Spanish, but a characteristic shared by many learners of Spanish with an L1 that does not use articles (Rubinjoni 2017; Santiago Alonso 2016; Tarrés Chamorro 2002; Zimny 2017; Testa 2019, 2020; Yonekawa & San Mateo Valdehita 2022) or uses them differently (Said-Mohand 2007). However, this study found an unexpected high incidence of morphological and syntactic errors. More precisely, the highest frequency of errors occurs in the prepositional phrase term category. This indicates that prepositions are another difficult aspect for Estonian learners of Spanish that deserves further investigation.

Moreover, it is observed that there is not a substantial improvement in the use of articles in the earlier stages of instruction (A1, A2 and B1). While previous studies indicate that errors in the use of articles decreases as learners progress (Said-Mohand 2007; Testa 2019, 2020; Zimny 2017; Yonekawa & San Mateo Valdehita 2022; Kruse & Rapún Mombiola 2023), this evolution is not always linear (Rubinjoni 2017, Montrul & Ionin 2012, Cuza *et al.* 2013). In fact, previous literature states that inaccuracies when using the article even persist even among advanced learners of Spanish (Rubinjoni 2017, Santiago Alonso 2016) and some of these errors might fossilise over time (Reymóndez Fernández & Santiago

Alonso 2004). This underlines the need for research to be carried out in further levels (B2 and C1) to observe how the acquisition of articles evolves among Estonian learners of Spanish.

BIBLIOGRAPHY

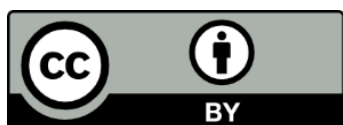
- Alarcón Moreno, Mauricia (2013), *Análisis de errores en la interlengua de aprendientes sirios de español como lengua extranjera*, Ph. D. dissertation, Madrid, Universidad Autónoma de Madrid.
- Alexopoulou, Angélica (2006), “Los criterios descriptivo y etiológico en la clasificación de los errores del hablante no nativo: una nueva perspectiva”, *Porta Linguarum: Revista Interuniversitaria de Didáctica de las Lenguas Extranjeras*, 5, 17-35.
- Ardura, Diego (2014), *The Definite Article System in L1-English L2-Spanish Learners*, master thesis, Amherst, University of Massachusetts.
- Aribaş, Derya Şekerci & Cele, Filiz (2021), “Acquisition of Articles in L2 and L3 English: The Influence of L2 Proficiency on Positive Transfer from L2 to L3”, *Journal of Multilingual and Multicultural Development*, 42(1), 19-36.
- Artstein, Ron (2017), “Inter-annotator Agreement”, in *Handbook of Linguistic Annotation*, Ide, Nancy & Pustejovsky, James (eds.), Dordrecht, Springer, pp. 297-313.
- Barrera Rodríguez, Francisco José (2008a), “Nongeneric Uses of Spanish Definite Article by Japanese Students”, *Polyglossia*, 14, 61-75.
- Barrera Rodríguez, Francisco José (2008b), “Notions of Definiteness, Specificity and Countability in the Grammaticality Judgments on Spanish Article System by Japanese Students”, *ポリグロシア*, 15, 45-63.
- Bop, Abdou Khadre (2023), “Contrastive Approach to the Determinative Articles in Wolof and Spanish: Didactical Implications in SFL”, *Revista SIGNOS ELE (Revista de español como lengua extranjera)*, 16(17), 1-24.
- Bruhn de Garavito, Joyce & White, Lydia (2002), “The Second Language Acquisition of Spanish DPs: The Status of Grammatical Features”, in *The Acquisition of Spanish Morphosyntax: The L1/L2 Connection*, Pérez-Leroux, Ana Teresa & Muñoz Licerias Juana (eds.), Dordrecht, Springer, pp. 153-178.
- Butt, John, Benjamin, Carmen & Moreira Rodríguez, Antonia (2019), *A New Reference Grammar of Modern Spanish (Routledge Reference Grammars)*, London/New York, Taylor & Francis.
- Cenoz, Jasone (2003), “The Additive Effect of Bilingualism on Third Language Acquisition: A Review”, *International Journal of Bilingualism*, 7(1), 71-87.
- Chierchia, Gennaro (1998), “Reference to kinds across languages”, *Natural Language Semantics*, 6, 339-405.
- Cuza, Alejandro, Guijarro-Fuentes, Pedro, Pires, Acrisio & Rothman, Jason (2013) “The Syntax-semantics of Bare and Definite Plural Subjects in the L2 Spanish of English Natives”, *International Journal of Bilingualism*, 17(5), 634-652.
- Davies, Mark (2016) = Now Corpus (News on the Web) [online]. Davies, M., 2016 [Accessed on: 07/10/2024]. Available at: <<https://www.english-corpora.org/now/>>.
- De Angelis, Gessica (2007), *Third or Additional Language Acquisition*, Bristol, Blue Ridge Summit: Multilingual Matters.

- Denst, Edyta (2006), *Uso del artículo en español para hablantes eslavos*, master thesis, Madrid, Universidad Complutense de Madrid.
- Fernández Jódar, Raúl (2006), *Análisis de errores léxicos, morfosintácticos, gráficos en la lengua escrita de los aprendices polacos de español*, Ph. D. dissertation, Poznań, Adam Mickiewicz University.
- Fernández López, M. Sonsoles (1988), "El uso del artículo en aprendices de español lengua extranjera", in *Actas del Primer Congreso Nacional de ASELE*, Fente Gómez, Rafael, Molina Redondo, José Andrés & Martínez González, Antonio (eds.), Málaga, Imagraf Impresores, pp. 109-118.
- Furukawa, Aya & Okami, Yurie (2016), "The Error Analysis of Spanish Articles: A View from Japanese", in *Proceedings of the 2nd International Conference on Linguistics and Language Studies*, Lee, F. K. C. et al. (eds.), Hong Kong, Chartered Institute of Linguistics Hong Kong Society, pp. 62-72.
- Garchana, Mar (2008), "Gramática y pragmática en el empleo del artículo", *MarcoELE*, 7, 1-19.
- Gilquin, Gaëtanelle (2015), "From design to collection of learner corpora", in *The Cambridge Handbook of Learner Corpus Research*, Granger, Sylviane, Gilquin, Gaëtanelle & Meunier, Fanny (eds.), Cambridge, Cambridge University Press, pp. 9-34.
- Granger, Sylviane (2008), "Learner corpora", in *Corpus Linguistics: An International Handbook*, vol. 1, Lüdeling, Anke & Kytö, Merja (eds.), Berlin, Walter de Gruyter, pp. 259-275.
- Gutiérrez Cansigno, Yvonne (2002), "El texto descriptivo en la enseñanza aprendizaje de una lengua extranjera", *Revista Fuentes Humanísticas*, 13(24), Article 24.
- Housen, Alex & Kuiken, Folkert (2009), "Complexity, Accuracy and Fluency in Second Language Acquisition", *Applied Linguistics*, 30(4), 461-473.
- Ionin, Tania, Ko, Heejeong & Wexler, Kenneth (2004), "Article Semantics in L2 Acquisition: The Role of Specificity", *Language Acquisition*, 12(1), 3-69.
- Ionin, Tania, Montrul, Silvina & Crivos, Mónica (2013), "A Bidirectional Study on the Acquisition of Plural Noun Phrase Interpretation in English and Spanish", *Applied Psycholinguistics*, 34(3), 483-518.
- James, Carl (1998), *Errors in Language Learning and Use*, London, Longman.
- Kilgarriff, Adam et al. (2014), "The Sketch Engine: Ten years on", *Lexicography*, 1(1), 7-36.
- Kirova, Alena & Camacho, José (2021), "Failed Gender Agreement in L1 English L2 Spanish: Syntactic or Lexical Problem?", *Glossa: A Journal of General Linguistics*, 6(1), 1-35.
- Kobashi, Saori (2018), *Enseñanza y aprendizaje del artículo en español por parte de estudiantes japoneses*, Ph. D. dissertation, Madrid, Universidad Complutense de Madrid.
- Kruse, Mari (2018), *La transferencia en personas plurilingües. Los falsos amigos como un obstáculo y una oportunidad en la enseñanza y aprendizaje de lenguas extranjeras*, Ph. D. dissertation, Tartu, University of Tartu.
- Kruse, Mari & Rapún Mombiela, Virginia (2023) "¿Influye el nivel de competencia en la producción de lenguas extranjeras? Estudio sobre la fluidez y la precisión en el uso de artículos y preposiciones en español como tercera lengua por parte de estudiantes estonios plurilingües", *e-Aesla*, 9, 17-32.

- Lee, Juyeon, Park, Soyoung & Heinz, Michael (2018), "Exploring Patterns of Article Use by Advanced Korean Learners of English and Spanish", *International Review of Applied Linguistics in Language Teaching*, 56(1), 79-101.
- Lin, Tzu-Lu (2003), *La adquisición y el uso del artículo por aprendientes chinos*, Ph. D. dissertation, Taichung City, Providence University.
- Lozano, Cristóbal (2022), "CEDEL2: Design, Compilation and Web Interface of an Online Corpus for L2 Spanish Acquisition Research", *Second Language Research*, 38(4), 965-983.
- McManus, Kevin (2021), *Crosslinguistic Influence and Second Language Learning*, New York, Routledge.
- Montrul, Silvina & Ionin, Tania (2012), "Dominant Language Transfer in Spanish Heritage Speakers and Second Language Learners in the Interpretation of Definite Articles", *The Modern Language Journal*, 96(1), 70-94.
- Montero Gálvez, Sonia (2011), "El artículo y otros fantasmas del nombre", *RedELE. Revista Electrónica de Didáctica del Español como Lengua Extranjera*, 8(21), 1-28.
- Montero Gálvez, Sonia (2014), "Los nombres escuetos y el artículo Ø", *Verba Hispanica*, 22, 37-58.
- Musto, Salvatore (2018), "A First Recognition of the Acquisition of Articles in Spanish L3 by Italian Students with English L2", *MarcoELE. Revista de Didáctica Español Lengua Extranjera*, 26.
- Nakagawa, Margarita (2004), "Empleo de los artículos por los estudiantes japoneses de español", *Cuadernos Canela*, 16, 161-178.
- Nergiz, Bengisu (2017), *Análisis de errores en textos escritos por estudiantes turcos del nivel B1 de español como lengua extranjera*, Ph. D. dissertation, Barcelona, Universitat Pompeu Fabra.
- Nikitina, Iuliia (2019), *Análisis de errores en el aprendizaje de los artículos en español por parte de aprendientes de habla estonia y rusa*, master thesis, Tartu, University of Tartu.
- Odlin, T. (2012), "Crosslinguistic Influence in Second Language Acquisition", in *The Encyclopedia of Applied Linguistics*, Chapelle, Carol A., Heboken, New Jersey, John Wiley & Sons.
- Osborne, John (2015), "Transfer and Learner Corpus Research", in *The Cambridge Handbook of Learner Corpus Research*, Granger, Sylviane, Gilquin, Gaëtanelle & Meunier, Fanny (eds.), Cambridge, Cambridge University Press, pp. 333-356.
- Pizarro Escabia, María (2012), *El tratamiento del artículo en manuales de ELE de nivel inicial (A1-A2)*, master thesis, Madrid, Universidad Nebrija.
- Popa, Simona (2016), *Language transfer in second language acquisition: Some effects of L1 instruction (Romanian) on L2/L3 learning (Catalan/Spanish)*, Ph. D. dissertation, Lleida, Universitat de Lleida.
- Prodanovic, Marijana, Corbolokovic, Saša & Gagić, Aleksandra (2023), "Correlation between Serbian Aspect in Adjectives and Articles in English and Spanish", *Revista Publicando*, 10, 82-90.
- Rapún Mombiela, Virginia (2023), "CLAE: Corpus Longitudinal de Español de Aprendientes Estonios. Diseño y estudio piloto", in *Creatividad, innovación y diversidad en la enseñanza del español como LE/L2*, Álvarez García, Esther & Villayandre Llamazares, Milka (eds.), León, Servicio de Publicaciones Universidad de León, pp. 351-362.

- Rapún Mombiola, Virginia & Tramallino, Carolina Paola (2023), "Acquisition of Spanish Temporal Adverbials by Multilingual Estonian Learners: A Proposal for Automatic Analysis", *Sustainable Multilingualism*, 23(1), 63-90.
- Real Academia de la Lengua Española (2016), *Nueva gramática de la lengua española*, Madrid, Planeta.
- Reymóndez Fernández, Iván & Santiago Alonso, Gemma María (2004), *Gramática mínima de errores. Análisis de errores españoles producidos por aprendices eslovenos*, Ljubljana, CIP.
- Rojo, Guillermo (2014), "Hispanic Corpus Linguistics", in *The Routledge Handbook of Hispanic Applied Linguistics*, Lacorte, Manel (ed.), London/New York, Routledge, pp. 371-378.
- Rothman, Jason, González Alonso, Jorge & Miller, David (2018), "The Acquisition of Second Language Spanish Morphosyntax", in *The Cambridge Handbook of Spanish Linguistics*, Geeslin Kimberly L. (ed.), Cambridge, Cambridge University Press, pp. 689-715.
- Rubinjon, Vlatka (2017), *Análisis de errores en el uso del artículo determinado en producciones escritas de aprendientes serbios de español como lengua extranjera*, Ph. D. dissertation, Alcalá de Henares, Universidad de Alcalá.
- Said-Mohand, Aixa (2007), "La adquisición del artículo definido: evidencia oral y escrita", *RedELE: Revista Electrónica de Didáctica ELE*, 10, 1-14.
- Santiago Alonso, Gemma Maria (2009), "La adquisición del artículo determinado por aprendices eslovenos de E/LE: análisis cuantitativo y cualitativo de errores", *Vestnik za tuje jezike*, 1(2), 35-51.
- Santiago Alonso, Gemma Maria (2010), "De la gramática cognitiva a la gramática pedagógica: hacia un modelo operativo del artículo español", *Verba Hispanica*, 18, 187-197.
- Santiago Alonso, Gemma Maria (2013), "La enseñanza significativa del artículo en español", *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 13, 991-995.
- Santiago Alonso, Gemma Maria (2014), "El análisis contrastivo como instrumento didáctico: los problemas de adquisición del artículo en español por aprendices eslovenos en los diferentes niveles del MCER", *Skupni evropski jezikovni okvir – navzkrižni pogledi*, 54, 397-410.
- Santiago Alonso, Gemma Maria (2016), "El artículo en español: problemas de conceptualización y valores esenciales", *Verba Hispanica*, 24, 125-145.
- Santiago Alonso, Gemma Maria (2017a), "Los efectos de la instrucción de procesamiento con input/output enriquecido para la adquisición del artículo en español", *Colindancias: Revista de la Red de Hispanistas de Europa Central*, 8, 221-249.
- Santiago Alonso, Gemma Maria (2017b), "Instrucción de procesamiento para la adquisición del artículo en español", *Verba Hispanica*, 25, 357-378.
- Santiago Alonso, Gemma Maria (2022), "La semántica del verbo en relación con la selección de predicados y distribución del artículo: hacia una aplicación didáctica para la enseñanza/aprendizaje del artículo", *Lingüística*, 62(1-2), 343-363.
- Shen, Jie (2012), *El artículo en la enseñanza de ELE: estudiantes de origen chino*, Ph. D. dissertation, Barcelona, Universitat de Barcelona.
- Tarrés Chamorro, Iñaki (2002), *El uso del artículo por estudiantes polacos de E/LE*, master thesis, Barcelona, Universitat de Barcelona.

- Testa, Martin (2019), "The Acquisition of Spanish L3 Articles: What Can be learned from a simple linear regression analysis?", *Research in Corpus Linguistics*, 7, 84-112.
- Testa, Martin (2020), "Exploring the Relation between Psychotypologies and Article Acquisition", *Psycholinguistics*, 27(2), 252-289.
- Thewissen, Jennifer & Anishchanka, Alena (2022), "Interaction between Grammatical Accuracy and Syntactic Complexity at Different Proficiency Levels: Insights from a Learner Corpus", in *Complexity, Accuracy and Fluency in Learner Corpus Research*, Leńko-Szymańska, Agnieszka & Götzt, Sandra (eds.), *Studies in Corpus Linguistics*, 104, Amsterdam, John Benjamins, pp. 209-240.
- Valverde, Pilar & Ohtani, Akira (2014), "Annotating Article Errors in Spanish Learner Texts: Design and Evaluation of an Annotation Scheme", in *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation*, Aroonmanakun, Wirote, Boonkwan, Prachya & Supnithi Thepchai (eds.), Phuket, Chulalongkorn University, pp. 234-243.
- Williams, Sarah & Hammarberg, Björn (1998), "Language Switches in L3 Production: Implications for a Polyglot Speaking Model", *Applied Linguistics*, 19(3), 295-333.
- Xiuchuan, Lu (2015), *La adquisición del artículo del español por los aprendientes chinos*, Ph. D. Dissertation, Madrid, Universidad Autónoma de Madrid.
- Xiuchuan, Lu (2016), "El estudio del estado inicial de la adquisición del artículo por los aprendientes chinos", in *Nuevos enfoques en investigación lingüística*, San Segunda, Pablo Rauno, Nieto Caballero, Guadalupe & Fernández de Molina Ortés, Elena (eds.), Badajoz, Universidad Extremadura, pp. 189-206.
- Xiuchuan, Lu (2017), "La secuencia de la adquisición de los diferentes valores del artículo español por los aprendientes chinos", *RedELE. Revista Electrónica de Didáctica ELE*, 29, 7-34.
- Xiuchuan, Lu (2018), "La adquisición del valor genérico del artículo español por los aprendientes chinos", in *Monográficos SINOELE: IX Congreso Internacional de la Asociación Asiática de Hispanistas*, Rangponsumrit, Nunghatai et al. (eds.) SINOELE, Bangkok, 447-459.
- Yang, Sung-Hye (2000), "La adquisición del artículo español por parte de los coreanos", in *Nuevas perspectivas en la enseñanza del español como lengua extranjera: Actas del X Congreso Internacional de ASELE*, Franco Figueroa, Mariano et al. (eds.), Cádiz, Universidad de Cádiz, pp. 749-755.
- Yonekawa, K. & San Mateo Valdehíta, A. (2022), "Análisis de errores en el uso del artículo en textos de aprendientes japoneses de español L2. Estudio de corpus", *Cuadernos CANELA*, 33, 71-89.
- Zimny, Ágata (2017), *Análisis de errores en la adquisición del artículo español por alumnos polacos de ELE*, master thesis, Madrid, Universidad Complutense de Madrid.



© Virginia Rapún Mombiela, Guillem Castañar Rubio & Anna Tudela Isanta, 2024.

Llevat que s'hi indiqui el contrari, els continguts d'aquesta revista estan subjectes a la [licència de Creative Commons: Reconeixement 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).