

Analysis of Errors by Beginner Students from Croatia: The Case of the Definite Article in Catalan, Galician and Spanish

Blanca León González, Martí Mestre Meliá and María Rosales Fernández
Department of Ibero-Romance Studies, University of Zadar

Abstract

Romance languages are characterised by the systematic use of definite articles, which is in contrast with Slavic languages which lack this grammatical category. The aim of this study is to analyse errors in Croatian speakers' use of definite articles in the context of the initial learning of Catalan, Galician and Spanish. The conclusions of this article were drawn from a corpus of essays written by Croatian students at A1 and A2 CEFR levels. This research allows us to observe and catalogue the different errors in the use of definite articles, as well as to confirm the hypothesis that motivated this work. The fact that these three languages – Catalan, Galician and Spanish – share both a common descriptive and prescriptive reality concerning the use of articles facilitated a cross-analysis in which interferences were minimal. In summary, the results demonstrate that the use of articles is one of the main grammatical difficulties in the transient interlanguage of Croatian speakers learning Romance languages, and highlight the need to create a method that prevents the fossilization of such errors.

Key words: *foreign language learning; Ibero-Romance languages; second language acquisition.*

Introduction

A high frequency of errors in the use of definite articles has been detected among Croatian students at the University of Zadar's Department of Ibero-Romance Studies. For this reason, this paper focuses on identifying and analysing the main problems for Croatian students regarding the use of definite articles in Catalan, Galician and Spanish, the three Iberian languages taught at this Department.

Due to historical and geographical reasons, German, Italian and French have traditionally been the most important foreign languages in Croatia. However, in the last decades there has also been a significant increase in the interest in learning Spanish (Elías Gutiérrez, 2006). Currently, in higher education, both Zagreb and Zadar universities offer Iberian studies as part of their academic programme. The latter also offers courses in Catalan and Galician, taught by native lecturers. Therefore, the University of Zadar constitutes an optimal framework for a study including more than one Iberian language.

Theoretical Framework

Our theoretical framework is based on the linguistic theory of “Errors Analysis” (EA), developed by Stephen Pit Corder in the 1960s. This method consists of the collection, detection, description and explanation of errors made by second-language learners in the target language. According to Corder (1974, pp. 120–125), “systematically analysing errors made by language learners makes it possible to determine areas that need reinforcement in teaching”, and also “provides us with a picture of the linguistic development of a learner and may give us indications as to the learning process.”

As previously explained, the aim of our work is to focus on the errors regarding definite articles. Since the Croatian language lacks this category, articles constitute a problematic area which needs more attention. Taking also into account the theory of Interlanguage (Selinker, 1972), it is our intention to prevent the fossilization of this type of errors. In this regard, interlanguage could be defined as a series of stages in which the second-language learner develops a particular linguistic system, which mixes elements from both first (L1) and second language (L2).

It is widely known that Corder’s approaches have received criticisms, since EA only presents learners’ failures; furthermore, identifying the origin of such errors sometimes turns out to be extremely difficult (Schachter & Celce-Murcia, 1977). However, it is also true, as Santos Gargallo (1993) points out, that EA can still be useful, especially in the research of languages whose relations have not been studied for a long time, as is the case with the present research.

As Iberian studies are relatively new to the Croatian education system there are not many contrastive analyses of Ibero-Romance languages and Croatian. However, there are at least two articles exploring errors made by Croatian speakers when learning Spanish. This is the basic literature related to the subject which has helped us in the preparation of this work:

- De la Cava Fernández-Coronado, P., & Sesmiło Pina, A. (2006). “Errores específicos en estudiantes eslavos (polacos y croatas) en las destrezas orales”

Although this first article focuses more on errors made by Polish students, it provides the basis for a catalogue of general errors between Romance and Slavic languages. The authors point to the use of articles as one of the main and more

persistent errors among both Polish and Croatian students. This paper is relevant to our research as it suggests a basic categorisation of problems related to the use of articles that has partially been used in this paper: overuse, underuse and misuse. In the present papers, the errors are classified as addition, omission and false choice. The latter refers to cases in which a student uses an indefinite article instead of a definite one, and vice versa, contrary to respective grammar rules.

- Kočman, A. (2011). “Errores gramaticales comunes en alumnos croatas de español como segunda lengua”

This article identifies and examines the most common mistakes in Spanish based on an extensive corpus of texts written by high school Croatian students. It constitutes a turning point as it is the first study to put the focus only on Croatian speakers. The author reports three trends among students: a general avoidance of articles, a false choice in the type of article and doubts regarding the use of definite articles, especially in cases in which the article does not define or restrain the referent. Kočman also points out a tendency towards overuse in cases in which grammar rules require an absence of definite articles. As in the previous case, this paper is especially valuable as it hints at the main problems that can be encountered when analysing Croatian learners’ use of articles.

Once again, it should be noted that both papers, by Kočman and De la Cava Fernández-Coronado and Sesmilo Pina, are relatively brief and discuss errors made by students in several areas (not only the use of articles). For these reasons, there are no detailed or quantitative results that can be compared to ours.

Forms and Uses of Definite Articles

All Romance languages, although not exclusively, include the grammatical category called article. Of course, in each of them the uses and values of articles can vary; however, they all share the notion of article. In contrast, the vast majority of Slavic languages lack this concept.

According to the Royal Spanish Academy dictionary, an article is a lexical category that determines if the noun or reference is known by the receiver or not (Real Academia Española, 2001).

In Catalan, Galician and Spanish, there are two main types of articles: definite and indefinite articles. In this paper, we will focus on the definite article due to the importance of the reference in the communicative context.

Table 1
Forms of definite articles in the three languages covered by this study

	Singular		Plural	
	Masculine	Feminine	Masculine	Feminine
Catalan	el, l'	la, l'	els	les
Galician	o	a	os	as
Spanish	el, lo (neutral)*	la	los	las

In order to understand our analysis better, we consider that it is essential to also know the uses of this category. To do so, we have consulted the main grammars in the three languages: *Nueva gramática de la lengua española* (2010), *Gramàtica del català contemporani* (2002), *Normas morfolòxicas e sintàcticas do galego* (2003). The basic uses could be summarized as follows:

- General use: all individuals or items from a group.
e.g.: *Las frutas tienen muchas vitaminas.*
As froitas teñen moitas vitaminas.
Les fruites tenen moltes vitamines.
*(The fruits have many vitamins).
- Uniqueness: there is only one referent.
e.g.: *El Sol sale por el Este.*
O Sol sae polo Leste.
El Sol surt per l'Est.
(The sun rises in the east).
- Deictic use: the information is present, current or obvious in the context.
e.g.: *Cierra la ventana, porfavor, que hace mucho frío.*
Pecha a fiestra, porfavor, que vaimoito frío.
Tanca la finestra, si us plau, que fa molt de fred.
(Close the window, please, it is very cold).
- Anaphoric use: the information has previously appeared in the discourse.
e.g.: *Me he comprado un coche (...). El coche es azul y consume mucha gasolina.*
Merquei un coche (...). O coche é azul e consume moita gasolina.
M'he comprat r un cotxe (...). El cotxe és blau i consume ix molta gasolina.
(I have bought a car [...]. The car is blue and consumes a lot of gas).
- Endophoric use: there is a modifier or complement that determines and restrains the information.
e.g.: *Le dio la carta que había estado escribiendo toda la mañana.*
Deulle a carta que estivera escribindo durante toda a mañá.
Li vadonar la carta que havia estate scrivint tot el matí.
(He gave him the letter he had been writing all morning).
- Possessive use: the referent of the nominal group involves inalienable elements such as parts of the body.
e.g.: *Tiene los ojos azules.*
Ten osollos azuis.
Té els ulls blaus.
*(She has the blue eyes).

- Other uses: superlative, collocations for common actions...

e.g.: *Marta es la chica más alta de la clase.*

Marta é a rapaza máis alta da clase.

Marta és la noia més alta de la classe.

(Marta is the tallest girl in the class).

e.g.: *Tomo el autobús al trabajo todas las mañanas.*

Collo o autobús ao traballo todas as mañás.

Agafó l'autobús a la feina cada matí.

(I take the bus to work every morning).

It is important to remember that rules and uses regarding articles depend on many factors: the intrinsic features of the verb, the type of noun, its number and function, etc. and are ultimately intimately bound to the context.

Definite articles in both the Catalan and Galician language have other uses, such as introducing anthroponyms (*La María, A Laura*) or marking possessives (*la meua* – mine in Catalan, *a súa* –his/her in Galician). However, we will omit these errors from the present study because of the differences between these languages and Spanish, and the subsequent inequality that might bias the results of this research. Please also note that for the purposes of understanding the given examples they have been literally transcribed into English (the translations may therefore include errors).

Aims and Hypotheses

Aims

The general aim of this study is to analyse the use of definite articles in the Catalan, Galician and Spanish language by Croatian beginner students (levels A1 to A2). In addition to this main goal, the following specific aims were targeted:

- to identify the most common error in general, and the variation among the three analysed languages;
- to describe the evolution of article usage with regard to the different levels;
- to compare the frequency of the different types of errors according to modality, written or oral expression.

Hypotheses

As it has previously been explained, the starting point of this study was the observation of a large number of errors in the use of articles among Croatian students. This observation was then compared with previous analyses identifying specific errors made by Slavic students, and as a result the following hypotheses were formulated:

- As Catalan, Galician and Spanish share the same morphosyntactic system, productions in the three languages show a similar frequency of errors.
- The use of the definite article presents a great problem for Croatian students.
- First-level students (A1) make more omission errors due to their ignorance of grammar rules.

- Second-level students (A2) show a tendency towards article overuse.
- False choice errors are uncommon.
- Oral speech tends to contain more errors in the use of articles than written production, due to the spontaneous character of spoken interventions.

Errors in the use of articles do not disappear at the second level of learning; sometimes it is even possible for this type of error to become more frequent over time.

Methods

In order to carry out this study, a corpus of texts in Catalan, Galician and Spanish produced by beginner students (levels A1 and A2) was gathered and analysed. Table 2 shows the distribution of modalities and languages within the corpus.

Table 2
Distribution of corpus texts

	Written expression		Oral expression		Total
	A1	A2	A1	A2	
Catalan	10	10	5	5	30
Galician	10	10	5	5	30
Spanish	10	10	5	5	30
					90

Table 2 shows the distribution of corpus texts. There is a total of 90 texts, 60 of which are written compositions of approximately 100–150 words, and 30 are oral productions of a duration ranging from 1 to 5 minutes. As for the languages and levels, 10 written compositions and 5 oral presentations have been analysed for each language (Catalan, Galician and Spanish) and level (A1 and A2). Half of the corpus, that is, 30 written texts and 15 oral productions, has been specifically designed for this study. The task consisted of a description of a family portrait in which the use of articles is essential, since students have a reference to write/speak about. The other half consists of randomly chosen texts: from routine descriptions to personal presentations.

Table 3
Example of analysis

No.	Linguistic context	Type of error	Language	Modality	Level	Observations
1.01	<i>No li agradava portar les botes.</i> (She did not like wearing <i>the</i> boots).	Addition	Catalan	W.E.	A1	The referent (<i>boots</i>) should not be specific: general use. It does not require any type of article.
2.01	<i>A avoa está sentada ao borde de (o) sofá.</i> (The grandmother is sitting on the edge of [<i>the</i>] sofa).	Omission	Galician	W.E.	A1	Deictic use: both speaker and listener share the referent (<i>sofa</i>) in the picture.
3.01	<i>Ella tiene un pelo rizado.</i> (She has a curly hair).	False choice	Spanish	W.E.	A1	Parts of the body (<i>hair</i>) should be expressed via possessive use of the definite article, as they are inalienable parts of a person.

Table 3 is an example of a table created to introduce the obtained data in our research. The table is divided into seven fields:

- Number of identification: 1 for Catalan, 2 for Galician, 3 for Spanish.
- Linguistic context of the error.
- Type of error: this study focuses on three kinds of errors based on the classification proposed by Roberto Maldonado Martínez de la Casa (2013, pp. 33–43), which consists of addition, omission and false choice.
- Language: Catalan, Galician or Spanish.
- Modality: written or oral expression.
- Level: first (A1) and second level (A2), according to the Common European Framework of Reference for Languages (CEFR).
- Observations: we included some remarks about mistakes, such as article use or some other general rule.

Results and Discussion

Comparison of Frequency and Type of Error in General by Language

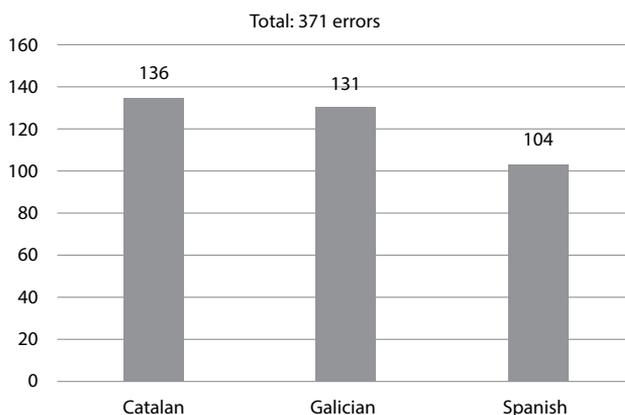


Figure 1. Frequency of error by language

Figure 1 shows that the number of errors was similar in the three analysed languages (136 in Catalan, 131 in Galician and 104 in Spanish). The slightly lower number of errors in Spanish can be explained by the fact that students have more contact and aplomb in the use of Spanish than in Catalan and Galician, which are elective courses and hence have fewer learning hours.

Figure 2 shows that omission was by far the most common mistake (225 cases). The lack of the grammatical notion of article in the learners' mother tongue makes the mistake more frequent. Figure 2 also shows that the second most common mistake was addition (122 cases), while the least common one was false choice (24 cases). This demonstrates that the problem surrounding articles is mainly related to over- and underuse. Finally, Figure 3 shows that the frequency and type of error were also very similar for each individual language.

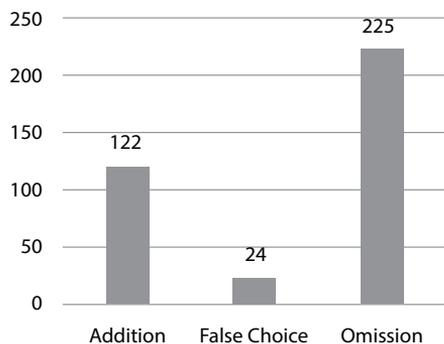


Figure 2. General overview by type of error

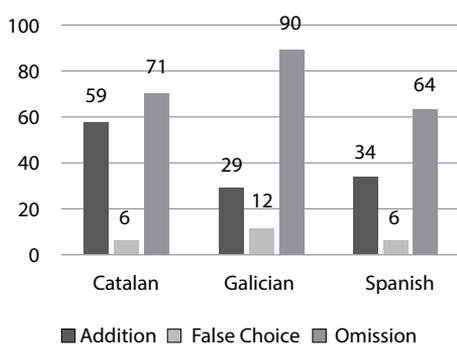


Figure 3. Frequency and type of error by language

Comparison of Frequency and Type of Error at Different Levels

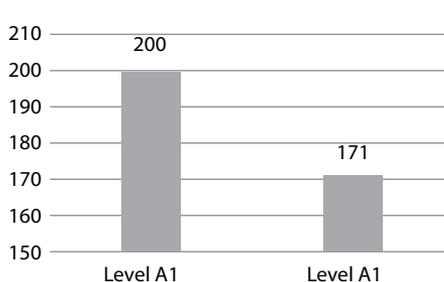


Figure 4. Frequency of error by level

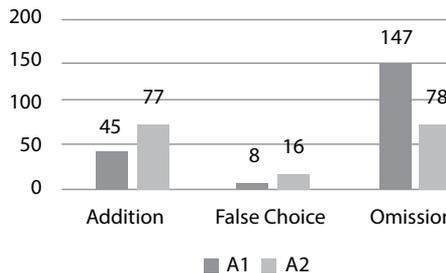


Figure 5. Type of error by level

Figure 4 shows that the frequency of error was higher in the A1-level group of students. However, there was not a dramatic decrease in the number of errors among A2-level learners (200 cases among A1 students vs. 171 cases among A2 students). Meanwhile, if the focus of attention is shifted to the evolution of the type of error over time, Figure 5 indicates that the most common error by far among A1 students is omission; as for A2 level students, the most common error is also omission, although it is almost as common as addition. This shows that false choice is not a common mistake in either of the levels.

These results can be explained by the fact that, due to the phenomenon of L1 transfer, first-level students tend to copy the system of their mother tongue, which in this case lacks the notion of article. In this way, students avoid using articles as they do not see the necessity for them. At higher levels, students are more aware of grammar rules and try to compensate, which leads to an overuse of articles. In any case, Figures 4 and 5 show that errors in the use of definite articles are common at both the A1 and A2 level. However, what changes over time is the type of error learners tend to make: from omission (A1 level) to omission and addition (A2 level), with false choice being quite uncommon.

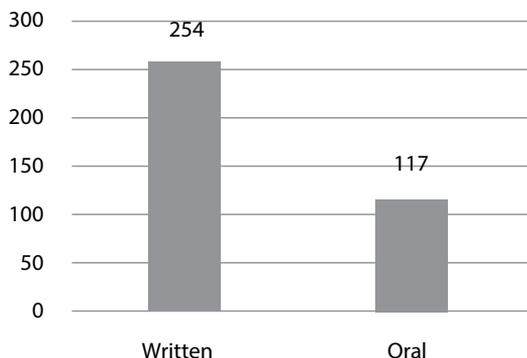


Figure 6. Frequency of errors by modality

Comparison of Frequency of Errors by Modality

For this analysis, a corpus of 60 written texts and 30 oral productions was gathered. The number of total errors was 371. Figure 6 shows the distribution of errors by modality (i.e. written or oral expression). However, it should be noted that there were twice as many written texts as there were oral ones. This means that the average number of errors in a written text is in fact 4.23 errors per written text, while the average number of errors per oral text is 3.9. This shows that there is indeed very little variation in the frequency of errors with regards to modality. These results could suggest that the root of Croatian speakers' errors in article use is indeed found in the interlanguage stage; therefore, the errors do not vary depending on the modality, but students' language learning level. However, further analysis is required.

Conclusions

After collecting and analysing a corpus of 90 texts (both written and oral) by Croatian learners, it can be suggested that definite articles constitute a problematic area in the learning of Ibero-Romance languages. The most common errors regarding articles in these languages could be divided into three categories: omission, addition and false choice in the use of articles.

The analysed data show a high and similar frequency of errors in the three Iberian languages analysed in this paper: Catalan, Galician and Spanish. By far the most common error regarding definite articles is the omission of articles, followed by the addition of unnecessary articles; false choice is the least common type of error. A1-level learners show a tendency towards underuse of definite articles, while A2-level students mainly make both omission and addition errors. At the same time, there is no significant difference in the frequency of errors between oral and written expression.

In summary, studying learners' errors can be helpful for identifying problematic areas in foreign language teaching, such as definite articles, and describing, categorising and preventing these errors.

References

- Corder, S. P. (1974). *Error Analysis*. Oxford: Oxford University Press.
- De la Cava Fernández-Coronado, P., & Sesmiolo Pina, A. (2006). *Errores específicos en estudiantes eslavos (polacos y croatas) en las destrezas orales*. Retrieved from http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/17/17_0415.pdf
- Elías Gutiérrez, C. (2006). *El español en Croacia*. En *Enciclopedia del español en el mundo*. Madrid: Instituto Cervantes, pp. 295-297. Retrieved from http://cvc.cervantes.es/lengua/anuario/anuario_06-07/pdf/paises_63.pdf
- Kočman, A. (2011). *Errores gramaticales comunes en alumnos croatas de español como segunda lengua*. Retrieved from http://marcoele.com/descargas/13/kocman-errores_croatas.pdf
- Maldonado Martínez de la Casa, R. (2013). *Análisis de errores en la adquisición/aprendizaje del español en estudiantes letones*. Universidad de Alcalá. Retrieved from <http://www.mecd.gob.es/dctm/redele/Material-RedEle/Biblioteca/2014bv15/2014bv1510robertomaldonado.pdf?documentId=0901e72b81b2db07>
- Real Academia Española. (2001). *Artículo*. En *Diccionario de la lengua española* (22nd edition). Retrieved from <http://www.rae.es/rae.html>
- Real Academia Galega (2003). *Normas ortográficas morfológicas do idioma galego*. Retrieved from <http://www.xunta.es/linguagalega/arquivos/normasrag.pdf>
- Real Academia Española (2010). *Nueva gramática de la lengua española*. Madrid: Espasa libros.
- Santos Gargallo, I. (1993). *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*. Madrid: Síntesis.
- Schachter, J., & Celce-Murcia, M. (1977). Some reservations concerning error analysis. *TESOL Quarterly*, 11, 441–451. <https://doi.org/10.2307/3585740>
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209–241. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Solà, J., Lloret, M., Mascaró, J., & Pérez Saldanya, M. (2002). *Gramàtica del català contemporani*. Barcelona: Empuries.

Blanca León González

Department of Ibero-Romance Studies, University of Zadar
Obala kralja Petra Krešimira IV 2, 23000, Zadar, Croatia
blanca.leon@eurekamadrid.es

Martí Mestre Meliá

Department of Ibero-Romance Studies, University of Zadar
Obala kralja Petra Krešimira IV 2, 23000, Zadar, Croatia
mestremelia@gmail.com

María Rosales Fernández

Department of Ibero-Romance Studies, University of Zadar
Obala kralja Petra Krešimira IV 2, 23000, Zadar, Croatia
mrosales@unizd.hr

Analiza pogrešaka kod hrvatskih studenata početnika: slučaj određenoga člana u katalonskome, galješkome i kastiljskome jeziku

Sažetak

Za romanske je jezike specifična sustavna upotreba određenih članova, koje slavenski jezici nemaju. U ovom se radu nastoje analizirati pogreške vezane uz upotrebu određenoga člana u kontekstu početnoga učenja katalonskoga, galješkoga i kastiljskoga jezika među izvornim govornicima hrvatskoga jezika. Zaključci su dobiveni na temelju korpusa koji čine eseji hrvatskih studenata čije znanje jezika odgovara razinama A1 i A2 prema ZEROJ-u. Ovo istraživanje omogućilo nam je da promatramo i katalogiziramo različite pogreške u vezi s upotrebom određenoga člana, kao i da potvrdimo hipoteze koje su potaknule njegovu izradu. Činjenica da je tim trima jezicima zajednička deskriptivna i preskriptivna stvarnost u vezi s upotrebom članova, olakšala je analizu u kojoj su interferencije minimalne. Zaključujemo da naši rezultati pokazuju kako je upotreba člana jedan od osnovnih gramatičkih problema u prolaznome međujeziku hrvatskih studenata koji uče romanske jezike, a posebna se pozornost posvećuje stvaranju metode koja bi spriječila fosilizaciju takvih pogrešaka.

Ključne riječi: *analiza pogrešaka; iberijski romanski jezici; određeni član; učenje stranoga jezika.*