

# The acquisition of L3 Spanish articles: What can be learned from a simple linear regression analysis?

Martin Testa  
University of Warsaw / Poland

**Abstract** – Despite being extremely frequent in Spanish, articles pose one of the biggest challenges for Polish L1/English L2 students who learn it as a L3, even among the most advanced learners. However, little attention has been paid to the influence of the task at hand, especially since certain registers and writing styles tend to make use of a higher number of noun phrases, thus increasing both the amount of article syntactical environments and the probability of an error to be made. The main purpose of this article is to investigate whether there is a statistically significant relation between the number of article-related errors and an increasing number of article environments. In order to do this, a simple linear regression analysis was run for both B1 and B2 groups separately, and finally the whole level ( $n = 72$ ). The results of the  $t$  tests suggest a significant linear relationship between the number of article-related errors and article tokens, between article-related errors and noun tokens, and between article and noun tokens. The article provides some considerations on possible actions to be taken both for researching and teaching the Spanish article system to [–ART] L1 learners.

**Keywords** – SLA; TLA; Spanish L3; article acquisition; simple linear regression

## 1. INTRODUCTION<sup>1</sup>

Although it may be true that articles might not be absolutely necessary in immediate environments (e.g. when ordering food; cf. Pica 1983), a more in-depth training of the semantic features of the article system seems necessary. This is especially the case when it comes to developing written competence, since in writing definiteness can only be expressed verbally.

---

<sup>1</sup> The author would like to express his very great appreciation to Prof. Jadwiga Linde-Usiekniewicz from the University of Warsaw for drawing attention to the issue of stylometric differences and their potential influence on the number of article-related errors. The author is also thankful to two anonymous reviewers for their thoughtful and constructive comments on an earlier version of this paper.



As will be shown in Section 2, below, evidence calls for a systematic pedagogical approach to the instruction of the article system, in which case learning the ‘function’ (cf. Pienemann 1998) of grammatical distinctions seems to be essential (cf. Widdowson 1988; Master 1994), especially at ‘intermediate levels’. Some authors (cf. Master 1994; Ionin *et al.* 2004) argue that extensive exposure to comprehensible input – namely by means of the ‘natural approach’ (cf. Krashen 1982; Krashen and Terrell 1983) – seems not enough for L2 article acquisition. Master (1994) acknowledges, however, that this method is appropriate at the beginning levels of L2 instruction, at which point instruction of grammatical distinctions might be too detailed (cf. Allen 1983; Little 1994; Master 1997). More studies in terms of naturalistic acquisition are then required before definite conclusions can be drawn.

Some research has been carried out on the influence of pedagogical intervention in terms of English L2 article acquisition. For example, Master (1987: 116) suggests the following sequence for the assessment of article environments:  $[\pm\text{count}] > [\pm\text{definite}] > [\pm\text{generic}] > [\pm\text{postmod}] > [\pm\text{proper}] > [\pm\text{idiomatic}]$ . The facilitating effects of the application of this kind of systematic long-term instruction is supported by Ekiert and Han (2016), who have called for more conclusive research on the effectiveness of pedagogical intervention in L2 article instruction.

The situation is even less clear when it comes to other languages. For example, in the case of Spanish FL handbooks the article seems to be only treated at the initial stages and the understanding of its semantics is almost immediately taken for granted and subsequently abandoned (cf. Tarrés Chamorro 2002; Lin 2003; Hidalgo 2015).<sup>2</sup> This proves to be an inadequate approach given that even the formal and semantic differences between the article systems of two [+ART] languages are highly abstract and difficult to address, let alone for learners whose L1 is [–ART], i.e. Slavic languages.

In terms of linguistic typology, languages can be classified into ‘article + tense’ (i.e. Romance and Germanic languages) and ‘aspect + case’ languages (i.e. Slavic languages). Taking into account this classification, Polish and Spanish are typologically different (cf. Nowikow 2017). Polish differs from Spanish in that definiteness can be expressed by means of different strategies: namely, word order, aspect, grammatical case, lexical choice (i.e. the numeral ‘one’), and deictic elements such as demonstratives or possessives (cf. Tarrés Chamorro 2002; Lin 2003; Shen Jie 2012; Hidalgo 2015;

---

<sup>2</sup> Some authors claim the same holds true for English (cf. Ionin *et al.* 2004; Sabir 2015).

Fernández Jódar 2017).<sup>3</sup> However, it should also be borne in mind that none of these strategies are completely equivalent to the semantics of articles (cf. Fernández Jódar 2017) and, as Ekiert and Han (2016: 151) note, even the very concept of *definiteness* “is often left to be inferred in a variety of ways.” This is, for instance, the case in the Polish sentence *Potknąłem się o psa* (‘I tripped over dog’). In this example, the lack of article does not give any hints as to whether the speaker is referring to *a dog* or *the dog* (cf. Allan 1986). In order to be more precise, for example, the speaker may either use the very name of the dog (i.e. *Potknąłem się o Reksia*), a determiner (i.e. *Potknąłem się o tego psa*), or fronting (i.e. *O psa się potknąłem*).

Studies such as Fernández Jódar (2006) have typically analyzed the ratio of article-related errors per one hundred words. However, it is still not clear whether a different approach to calculating the ratio of such errors would be more accurate, since different types of texts can be characterized by different writing styles (e.g. ‘nominal’ argumentative texts vs. ‘verbal’ narrative texts). The main purpose of this article is therefore to determine whether there exist statistically significant relations between: (a) the number of article-related errors and the number of article tokens; (b) the number of article-related errors and the number of noun tokens; and (c) the number of article tokens and the number of noun tokens.

The paper is organised as follows. Section 2 offers a review of the literature on article acquisition and instruction, which is intended to provide the reader with a basic understanding of the complexity of learning the article system, especially by learners whose mother tongue does not make use of articles: i.e. [–ART] L1 speakers. Section 3 provides a description of the data gathering procedures. The results will be presented in Section 4 and will be discussed in Section 5. Finally, Section 6 offers some concluding remarks and points out future lines of research.

## 2. LITERATURE REVIEW

There is extensive literature on the acquisition of English articles (cf. Master 1986, 1987, 1994, 2002; Ionin *et al.* 2004; Jaensch 2008; Yoo 2009; Sabir 2015; Ekiert and Han 2016; Sun 2016; or Şekerci Arıbaş and Cele 2019). As regards Spanish, some

---

<sup>3</sup> For a contrastive analysis between Spanish and Polish see Konieczna-Twardzikowa (1992), Tarrés Chamorro (2002) and Fernández Jódar (2017). For Spanish and Chinese see Lin (2003), Shen Jie (2012), and Hidalgo (2015).

research has been carried out on the acquisition of the Spanish article system by Polish L1 students (cf. Tarrés Chamorro 2002; Fernández Jódar 2006; Testa 2019, in press) and by Chinese L1 students (cf. Lin 2003; Shen Jie 2012; Hidalgo 2015).

Although it may be true that, as Master (1994: 230) note, “a person may communicate effectively in spoken English even when article use is entirely erroneous,” article-related errors are especially visible in writing, where a weak command in the L2 article system might as well undermine the students’ academic performance (cf. Master 1987, 1990, 1997). Many authors have pointed out that even high-intermediate level students use articles in an intuitive way (cf. Master 1994; Tarrés Chamorro 2002; Hidalgo 2015; Lema 2016). However, the way L1 speakers of [–ART] languages use articles is not random but actually responds to universal semantic distinctions (cf. Ionin *et al.* 2004, see section 2.1. below).

Despite being extremely frequent (cf. Hewson 1972), articles still pose great problems for students whose L1 does not have such a category (cf. Ringbom 1987; Master 1990, 2002; Mizuno 1999; Lin 2003; Park 2006; Harb 2014). These speakers often perceive articles as redundant (cf. George 1972) and frequently avoid their use (cf. Ringbom 2011), especially during the first stages of language learning (cf. Ringbom 2016). Jiang *et al.* (2011: 959, cited in Ekiert and Han 2016: 149) explain that grammatical morpheme acquisition is particularly demanding since its “related meaning is not grammaticalized in the learner’s L1, which means that the related meaning is not part of the routinely activated meanings in the learner’s mind.” Master (1990) also points out that the article system is one of the latest aspects of L2 English to be fully acquired. In fact, as a result of lacking such a category in their native languages, [–ART] learners are one stage behind [+ART] learners when it comes to acquiring the L2 article system and therefore need more time (cf. Bailey *et al.* 1974; Hakuta 1976; Master 1997) to activate them in their mind.

Nonetheless, article-related errors are found even among the most advanced students (cf. Lin 2003; Ekiert and Han 2016) and also among students whose L1 features such a category (cf. Odlin 1989). Knowing an additional [+ART] L2 may indeed be of help, but only if the learners’ command of the L2 is highly proficient or native-like, will their knowledge be available for transfer to the L3 (cf. Tarrés Chamorro 2002; Jaensch 2008; or Testa 2019, in press). At lower levels the students’ mastery of the L2 article system may be far from perfect and, therefore, not very helpful. Moreover, not all

article systems are alike and both formal and conceptual differences between [+ART] languages add an extra layer of difficulty when it comes to acquiring the article system of an L3. Such a difficulty may even be further increased because of lexicalised and idiomatic uses in the article system of the different languages.

Harb (2014: 98–99) shows that students transfer the semantic properties of articles from the L1 to the L2. For example, this is the case between English [+ART] and Spanish [+ART] (cf. García Mayo 2008; Ionin *et al.* 2008; Isabelli-García and Slough 2012), and between English and Arabic [+ART] (cf. Kharma 1981; Schulz 2004). However, it remains unclear why [+ART] speakers are more able to transfer [+DEF]-usage than [–DEF]-usage into L2 English, given that languages such as Spanish and German also have indefinite articles (Master 1987). In this respect, Hidalgo (2015) draws attention to the fact that Mandarin Chinese may sometimes use the numeral *yī* (‘one’) in order to mark [–definite], but it may be omitted and for that reason positive transfer does not always take place.

Ionin *et al.* (2004) suggest two important hypotheses: ‘the Article Choice Parameter’ and the ‘fluctuation hypothesis’. Regarding the Article Choice Parameter hypothesis, they suggest that [+ART] languages distinguish articles on the basis of either definiteness (i.e. English) or specificity (i.e. Samoan).<sup>4</sup> For example, on the one hand, Samoan features a marker for specificity *le* that corresponds to both English definite/specific *the* and indefinite/specific *a*. On the other hand, the Samoan marker for non-specific *se* may correspond to both English definite/non-specific *the* and indefinite/non-specific *a*. In fact, as Lyons (1999: 167) illustrates, definites can be either specific, as in *Joan wants to present the prize to the winner — but he doesn’t want to receive it from her*, or non-specific, as in *Joan wants to present the prize to the winner — so she’ll have to wait around till the race finishes*. Indefinites can also be either specific, as in *Peter intends to marry a/this merchant banker — even though he doesn’t get on at all with her* or non-specific, as in *Peter intends to marry a/this merchant banker — though he hasn’t met one yet*, as shown in Lyons (1999: 176). As Ionin *et al.* (2004: 17) point out, the Fluctuation Hypothesis holds that “L2 learners have full UG access to the two settings of the Article Choice Parameter” and that they fluctuate

---

<sup>4</sup> According to Ionin *et al.* (2004: 9), specificity involves “speaker intent to refer to an individual who exists in the actual world.” This explains errors in [–definite, +specific] contexts, as illustrated in *I am visiting a/\*the friend from college — his name is Sam Brown, and he lives in Cambridge now* (taken from Ionin *et al.* 2004: 41).

between those two settings “until the input leads them to set this parameter to the appropriate value.”

Jenks (2018) provides a typology of definiteness marking. ‘Bipartite’ languages have one form of the definite article for unique contexts (DEF<sub>weak</sub>) and a different one for anaphoric contexts (DEF<sub>strong</sub>). An example is German, as provided in Schwarz (2009: 41): *In der Kabinettsitzung heute wird ein neuer Vorschlag vom Kanzler erwartet* (‘In today’s cabinet meeting, a new proposal by the chancellor is expected’) vs. *In der Kabinettsitzung heute wird ein neuer Vorschlag von dem Minister erwartet* (‘In today’s cabinet meeting, a new proposal by the minister is expected’). ‘Marked anaphoric’ languages, such as Mandarin, use bare nouns in unique contexts (Ø) but tend to use a demonstrative in anaphoric contexts (DEF<sub>strong</sub>). Languages such as English and Cantonese are marked, i.e. both for unique and anaphoric contexts. Finally, Jenks explains that it seems theoretically impossible for a language to mark unique definites and not to have an anaphoric definite article, i.e. DEF<sub>weak</sub> for unique contexts and Ø for anaphoric ones.

We will now provide an overview in terms of LX English and LX Spanish article instruction research.

### 2.1. LX English article instruction

Master (1987) studies the acquisition of the English article system by L1 speakers of Chinese (–ART), Japanese (–ART), Russian (–ART), Spanish (+ART) and German (+ART), in terms of four main categories of article use, namely generic (*a*, *the*, Ø), specific definite (*the*), specific indefinite (*a*, Ø) and ambiguous generic (*a*, Ø).<sup>5</sup> Master (1997) claims that both [–ART] and [+ART] groups show internal homogeneity, although it should be borne in mind that he had used only one subject per L1, per level, and at least some part of the results might be influenced by individual variation. For example, [–ART] groups tend to go from an overextended use of Ø at the lower levels (which is something expected, since their L1s have no articles), to an overextended use of *the* once they have become aware of the morpheme, and then to a slowly increasing use of the indefinite article *a*. On the other hand, [+ART] groups overuse *the* right from the start and use Ø correspondingly less than their [–ART] counterparts. The acquisition

---

<sup>5</sup> For examples in each category, see Master (1987: 23).

of the indefinite article seems to happen as an independent process (cf. Pica 1983, 1985; Master 1987). [+DEF]-flooding seems common at the intermediate levels, probably due to recognition of the fact that English noun phrases need a specifier, and once “they realize that Ø can be a specifier too, and they start to increase their Ø-usage” (cf. Master 1987: 88). [+DEF]-flooding at the intermediate levels has also been observed among Chinese L1 learners of Spanish (cf. Lin 2003; Shen Jie 2012; Hidalgo 2015). Jenks (2018: 501), for instance, points out that, in Mandarin, “unique definites are realized with a bare noun, [whereas] anaphoric definites are realized with a demonstrative, except in subject position” and that, while lacking a definite article, Mandarin seems to distinguish between unique and previously mentioned definites (i.e. anaphoric).

Master (1990: 461) suggests that the article system may be taught “as a binary division between classification (*a* and Ø) and identification (*the*)” and further argues that “[a]ll the other elements of article usage can be understood within this framework.” Master (1990: 465) goes on to claim that “determining the correct article in English requires the simultaneous consideration of four features: ‘definiteness’ [±definite], specificity [±specific], countability [+count], and number [±singular].” According to Master (1990: 465), “number subset really only applies to [+count] nouns and should therefore only be considered a feature of that subset.” This increases the students’ cognitive load dramatically, and it is likely to slow down fluency considerably (cf. Krashen 1982). That is why Master (1990: 466) suggests “collapsing the features” [±definite] and [±specific] into one single feature [±identified], thus reducing the number of concepts to be taught to two: identification vs. classification.<sup>6</sup> Nevertheless, Master (1990: 474–476) acknowledges the complexity of article-usage when it comes to proper nouns and idiomatic phrases, which remain “in the realm of things which must be learned and memorized and for which there is rarely a productive rule.” The binary framework, identification vs. classification, is tested in Master (1994: 245) who concludes that “the article system can indeed be learned and that it is perhaps the systematic presentation of the article system that makes the difference.”

Master (1997) elaborates on the abovementioned binary framework and explores the differences between the two types of Ø (the zero article Ø<sub>1</sub>, and the null article Ø<sub>2</sub>). While Ø<sub>1</sub>, (cf. *I ate the pizza* vs. *I ate a pizza* vs. *I ate pizza*), is “the most indefinite

---

<sup>6</sup> Master (1990: 469) notes that classification appeals to the awareness of the [±count] feature (*a* vs. Ø), as in *What’s this? = It’s a whiteboard/It’s chalk/These are books/This is paper*.

article in the article system” (cf. Harb 2014: 91),  $\emptyset_2$  (cf. *After dinner, I’m going home*) is the most definite one, because it is applied in contexts where “definiteness can still be internally attained without the addition of the article *the*” (cf. Harb 2014: 92). Moreover, Master (1997: 226) suggests that at beginning levels of L2 instruction only extensive exposure is advisable given that rules might be yet obscure to learners (cf. also Allen 1983; Little 1994). According to Master, at intermediate levels the article system should be taught in a systematic way by means of the binary framework (cf. Master 1990) or input processing (cf. VanPatten and Cadierno 1993). Finally, at more advanced levels a lexical rather than syntactic approach seems more appropriate, for instance, lexical minimal pairs  $\emptyset$  vs. *the* (cf. Master 1995).

Master (2002) finds that with the use of an information structure framework in one month of teaching the English article system to a group of students (mixed L1s, +Art and –Art), these made small but significant improvement when compared to a control group which had been taught by means of more traditional explanations, and a third group which had received no article instruction. Master (2002: 331) suggests that language teachers “present canonical information structure as a preliminary guess in determining the appropriate article for any noun, providing a further potential aid in learning the article system.”

Ionin *et al.* (2004: 4) analyze article use by L1 Russian and L1 Korean (both [–ART] languages) learners of L2 English and argue that [+DEF]-flooding with indefinites “is systematic, being tied to the occurrence of the feature [+specific],” and that the same holds true for [–DEF]-flooding with definites (i.e. absence of the feature [+specific]). Ionin *et al.* (2004: 42) also point out that “as proficiency increases L2 learners are able to set the Article Choice Parameter (although many advanced learners still show fluctuation).” Their findings therefore suggest that there may be direct access to universal semantic distinctions in L2 acquisition, given that L1 transfer cannot account for access to the feature [+specific].

Jaensch (2008: 87) studies the acquisition of L3 German articles by native [–ART] Japanese speakers. She finds a positive effect of L2 English proficiency, since L1 Japanese/L2 English/L3 German learners seem to be “more aware of the definiteness feature – perhaps due to having acquired an L2 which has articles that mark definiteness in the same manner” – than [–ART] L1 Russian/L2 English and [–ART] L1 Korean/L2 English learners in Ionin *et al.* (2004).



Sabir (2015) analyzes article use by L1 Saudi (Hejazi) Arabic learners of L2 English in three tasks: article elicitation, acceptability judgment and elicited written production. Sabir finds no clear relation between explicit article instruction – i.e. instruction in definiteness, specificity and genericity, as well as translation activities – and article accuracy. Her results are consistent with the Fluctuation Hypothesis postulated by Ionin *et al.* (2004).

In line with Master (1987), Sun (2016) finds that, in her study, the [+ART] group produces much more correct article-related sentences than the [–ART] group. Sun (2016: 5) also finds that the Ø article “is the last one to be acquired and is the most difficult one for L2 learners of all levels,” whereas the “indefinite article *a* is the first one and the easiest one for L2 learners of all levels to acquire.” In fact, Sun claims that there is correspondence between the proficiency level and the acquisition of the indefinite article, although not the same can be claimed for the definite and Ø article.

Ekiert and Han (2016) analyze the acquisition of English articles by 65 Slavic learners (Polish *n* = 42, Russian *n* = 11, and Ukrainian *n* = 12) by using video retelling, missing word activities, and a translation methodology. Ekiert and Han (2016: 166) point out to that “[i]n general, the L1-Polish participants considered speaker-oriented identifiability sufficient for reference tracking in English” (i.e. givenness), thus having a harder time at choosing the right form of the article in the case of first-mention referents. In [+ART] languages, the definite article implies that the referent of the noun phrase must be identifiable either within the discourse or “uniquely identifiable to the hearer” (Birner and Ward 1994: 1), that is, there should be no room for ambiguity in the ears of the hearer (cf. Hawkins 1991). Following Trenkic (2002), Ekiert and Han (2016: 165–166) explain that in Slavic languages, “the way objects are referred to mirrors the state of knowledge of the speaker only,” and this often results in interference at the level of discourse. Following Yoo (2009), Ekiert and Han (2016: 150) point out that, although most student textbooks put emphasis on the anaphoric discourse-oriented use of *the* (i.e. *I bought a book. On Tuesday, I finished the book*), it is the cataphoric and situational uses of the definite article which are much more common in conversational English (i.e. *I had an argument at the office*).

Finally, Şekerci Arıbaş and Cele (2019) compare the performance of L1 Turkish/L2 German learners of L3 English and L1 Turkish learners of L2 English, and

find that L3 English learners are significantly more accurate than L2 English learners in all article contexts. Their results are also consistent with the Fluctuation Hypothesis.

In this section, we have seen that many studies support the claim that knowledge of article semantics can be transferred from the L1 to the target language, as well as from a L2 [+ART] to an L3, especially if the level in the L2 is high. Proficiency in the target language has also been linked to a better command of article use. Moreover, [+DEF]-flooding with indefinites and [-DEF]-flooding with definites have been found to be systematic and their motivation can be linked to a readjustment of the Article Choice Parameter. Therefore, although there is some evidence on the benefits of systematic explicit instruction, there is also evidence in support of the Fluctuation Hypothesis.

## 2.2. *LX Spanish article instruction*

Because of the perceived differences between [-ART] and [+ART] languages, Spanish articles pose a problem for Polish L1 students, regardless of the task and the level (cf. Fernández Jódar 2006: 97). Fernández Jódar (2006) identified up to twenty types of article-related errors, the most frequent being the unnecessary use of the definite article (e.g. *come \* la carne*).<sup>7</sup> Furthermore, Fernández Jódar (2006: 106) concludes that articles cannot be said to be fully acquired by Polish L1 students (cf. Tarrés Chamorro 2002), and that the greatest problem lies in the unnecessary use of Spanish articles with non-referential noun phrases and the incorrect avoidance of the article, particularly the definite one. On the other hand, Fernández Jódar (2006: 107) notes progress in the use of the zero article with proper nouns and idioms, as well in the use of the neutral article *lo*, which is absent at early levels of instruction. However, it is not clear whether this represents a better command of the semantics of the zero article or it is rather correct avoidance of the article, that is, a covert error.

Tarrés Chamorro (2002: 79) also argues that a high level in a [+ART] L2 might be beneficial, although the evidence is sometimes contradictory. Moreover, he argues that L2 English does not seem to be of much help, probably because students have not yet fully acquired the English article system either. Tarrés Chamorro (2002: 74) reports the

---

<sup>7</sup> Pragmatics and semantics are crucial to determining the correctness of a sentence like *come (\*) la carne*. The definite article is used in imperative sentences as in *¡Come la carne!* ('eat your meat!'). However, in declarative clauses, the definite article is not used: that is, when referring to the act of eating meat regularly *Mi hijo no come carne*. ('My son doesn't eat meat.')

case of an L1 Polish student, who had been in permanent and intense contact with the target language, and seemed to have achieved native-like control of the L3 Spanish definite article, mostly because of her extraordinarily high motivation and engagement. This could show that higher levels of motivation and engagement might be linked to more exposure to the target language, as well as to more attention to its distinctive features.

Lin (2003) studies 80 writing assignments by Chinese L1 (Taiwanese) multilingual learners of Spanish from four different levels at the University of Alcalá (Spain), and analyzes article-related errors in terms of *signifier* and *signified*-related errors (i.e. article omission; unnecessary use of article; and wrong choice of article). As a general strategy, Lin detects initial article omission to subsequent [+DEF]-flooding, and also finds that the most common type of error is article omission (especially [+DEF]-omission). At lower levels, Lin (2003) demonstrates that INDEF usage seems to be aided by the transfer of the Mandarin Chinese numeral *yī*.<sup>8</sup> Lin argues that the use of the Chinese numeral should result in positive transfer and, for this reason, the number of Ø-for-INDEF errors should be low. However, the Chinese numeral can be often omitted and, as a consequence, interference is still possible and positive transfer does not often take place (see also Hidalgo 2015). On the other hand, it has been shown (cf. Lapesa 2000; Lin 2003; Hidalgo 2015) that Mandarin allows the numeral to be placed before ‘classifying’ predicate nouns while Spanish does not (especially after copular verbs), thus resulting in interference, as in *\*soy un estudiante* (‘I am a student’) or *esto es \*un asunto mío* (‘this is my issue’). Once again, pragmatics and semantics are crucial in determining the correctness of these utterances. For example, the zero article is used after copular *ser* (‘to be’) as is the case with professions, cf. *Marta es actriz* (‘Marta is an actress’), but the definite article can be used in emphatic contexts *Marta es la actriz más famosa* (‘Marta is the most famous actress’), and the indefinite article may be used for comparison *Marta es una actriz* (‘Marta {behaves ~ looks ~ talks} like an actress’).

Shen Jie (2012) analyzes 90 writing assignments from A1-B2 Chinese L1 students of Spanish at the ‘Official School of Languages’ in Barcelona, and finds that the major issue for the students seemed to be the [±definite] feature, as well as the use of the Ø article when it comes to bare nouns.

---

<sup>8</sup> This is also noted in the case of Polish L1 learners (cf. Fernández Jódar 2006). We argue that this is due to the fact that students seem unaware of the rule that, with the exception of topicalized unaccountable nouns, Spanish nouns in subject position require the use of the article.

Hidalgo (2015) studies 20 article-related tasks (11 fill-in-the-blank and 9 translation exercises) from Chinese L1 learners of Spanish at B1 level, and finds that the most frequent errors to be DEF-for-INDEF ([+DEF]-flooding) as well as INDEF-for-DEF, followed by Ø-for-DEF.<sup>9</sup>

Finally, Testa (in press) explores whether knowledge of a [+ART] L2 was of any help when learning the article system of L3 Spanish. In order to do this, he analyzes both the production and comprehension of 25 Polish L1 intermediate-level students (B1 *n* = 11, B2 *n* = 14) by means of a translation task, a grammaticality judgment task, and post-task interviews. When it comes to the translation task, Spanish generic definite (Polish Ø) and bare nouns after copular *ser* ('to be') turned out to be the most difficult sentences to translate, with less than 40% of target-like answers each in both groups. In the case of the grammaticality judgment task, both groups find it particularly hard to recognize wrong Ø use (cf. George 1972) and, in contrast, have very little difficulty to correct wrong DEF-for-Ø and INDEF-for-Ø. Finally, the post-task interviews reveal that in many cases students tend to choose between article forms intuitively. They may be aware of some of the semantic/pragmatic distinctions but often fail at mapping them onto the right article form. Furthermore, during the post-task interviews none of the students pointed to English L2 influence, except for one student with native-like command of English. On the other hand, other languages (i.e. German, Italian and Portuguese) are considered to have an influence on the students' choices, which frequently results in interference (cf. Testa 2019: 40–41). The results echo Tarrés Chamorro's (2002: 51) suggestion that positive transfer might happen only if the level in the L2 is very high, that is, proficient user or native-like.

In this section we have observed that [+DEF]-flooding is also common in L3 Spanish as a general strategy by [–ART] learners, following initial article omission. Moreover, high proficiency in an [+ART] L2 seems to help in transferring article-related knowledge into L3 Spanish.

To the best of our knowledge, there are currently no studies that examine whether a higher number of article-related errors can be expected with increasing article usage. In other words, the type of task may exert an influence on the total number of article-related errors since certain writing styles tend to make a more frequent use of noun phrases – i.e. 'nominal' argumentative texts, as opposed to 'verbal' narrative texts) –

---

<sup>9</sup> This is especially the case with the verb *tener* ('to have'), as shown in Testa (2019: 273, 284, 295).

thus increasing both the amount of article environments and the probability of error. When it comes to Spanish, Fernández Jódar (2006) has analyzed the number of article-related errors per one hundred words but, as stated above, we may expect to find more articles and nouns in argumentative texts and, therefore, more environments for article-related errors. We argue that such an analysis would be a better indicator of the real ratio of article-related errors than the number of errors per one hundred words, or else the mean average of such errors.

In light of these considerations, the present study will address the following research questions:

1. Is there a significant relationship between the number of article-related errors (ArE) and the number of article tokens (AT)?
2. Is there a significant relationship between the number of article-related errors (ArE) and the number of noun tokens (NT)?
3. Is there a significant relationship between the number of article tokens (AT) and the number of noun tokens (NT)?

### 3. THE STUDY

The current study uses sample data collected from two intermediate-level (B1  $n = 35$ ; B2  $n = 37$ ) Spanish language courses that took place at the Faculty of Modern Languages at the University of Warsaw, between December 2016 and June 2017. The sampled population was highly homogeneous, since the participants were all multilingual Polish L1 university students doing a degree in Linguistics, with ages ranging from 19 to 25. They had English as their second language and were studying Spanish either as L3 (in most cases) or L4. The students had been admitted for the B1 and B2 courses on the basis of having passed an A2 and a B1 course, respectively.

The corpus consists of 72 compositions in total (15,288 words) coming from three different writing assignments about topics that had been discussed orally in class, in order to reduce the possible amount of errors due to ignorance of vocabulary. The topics are: 1) *Mis mejores vacaciones* ('The best holidays I have ever had') (B1  $n = 13$ ; B2  $n = 15$ ); 2) an e-mail to a friend telling them about a personal anecdote (B1  $n = 10$ ; B2  $n = 10$ ); and 3) *Ventajas y desventajas del uso de Internet a nivel educativo* ('Advantages

and disadvantages of using Internet for educational purposes') (B1 n = 12; B2 n = 12). All assignments were written in class and students were given a total of forty-five minutes to do them. The first assignment took place at the end of the winter semester (January 2017), the second one was carried out before the Easter break (April 2017), and the third assignment was scheduled at the end of the summer semester (June 2017). At no time were the students told that the activities were aimed at analyzing their use of Spanish articles. Moreover, since the data come from two narrative texts and one argumentative text there might be stylometric differences with regard to the number of article environments. For example, in the narrative texts a higher number of verbs should be expected, whereas in the argument text more noun phrases could be used.

In order to answer our three research questions a simple linear regression analysis was carried out. A simple linear regression analysis is normally used in order to determine whether the resulting regression line fits the data better than the average mean line. In our study, we will try to determine whether there is a significant relationship between: (a) the number of article-related errors (ArE) and article tokens (AT); (b) the number of article-related errors (ArE) and noun tokens (NT); and (c) between the number of article tokens (AT) and noun tokens (NT). If the results are significant, we may conclude that those relations can predict the amount of article-related errors better than the mean average number of errors (our first two research questions), as well as predicting the amount of articles better the mean average number of article tokens (our third research question). The equations used in the analysis are shown in Table 1.

Regression line	Slope calculation	Intercept calculation	Coefficient of determination
$y_i = b_0 + b_1x_i$	$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$	$b_0 = \bar{y} - b_1\bar{x}$	$r^2 = \frac{SSR}{SST}$
SST	SSE	SSR	t test
$\sum(y_i - \bar{y})^2$	$\sum(y_i - \hat{y}_i)^2$	$SSR = SST - SSE$	$t = \frac{b_1}{s_{b_1}}$
Standard error	Confidence intervals	Standar deviation of y *	Prediction intervals
$\sigma = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}}$	$y^* \pm t_{\alpha/2} s_{y^*}$	$s_{y^*} = s \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$	$y^* \pm t_{\alpha/2} s_{pred}$

Table 1: Equations used in the simple linear regression analysis

## 4. RESULTS

### 4.1. Regression analysis of groups B1 and B2 (separately)

#### 4.1.1. Number of article-related errors (ArE) vs. Number of article tokens (AT)

As can be noticed in Table 2 below, the regression line for group B1 ( $y_i = 0.2242x + 3.7771$ ) suggests that for every (1) article environment ( $x$ ) we would expect the amount of article-related errors to increase by 0.2242. In Figure 1, it can be seen how the mean line relates to the observed number of errors on the scatter plot. The coefficient of determination ( $r^2$ ) reveals that almost one third of the total sum of squares can be explained by using the estimated regression equation to predict the amount of article-related errors. In other words, as much as 30.78% of the variance in the amount of article-related errors can be explained by increasing article use. The standard error value ( $\sigma = 3.5611$ ) was obtained from the square root of the value resulting from the division of the SSE (418.4808) by the degrees of freedom ( $n - 2 = 33$ ). The confidence interval for the slope was estimated at (0.1055, 0.3429) and it does not contain zero, so the null hypothesis – that is, the mean average number of article-related errors fits the data better – can be rejected.<sup>10</sup> Finally, the t-test indicates that there is a statistically significant relationship between the number of article-related errors and the number of article tokens ( $t = 3.8324 > t_{crit} = 2.03$ ).

<sup>10</sup> All calculations in this study use  $\alpha$  0.05.

	ArE vs. AT		ArE vs. NT		AT vs. NT	
	B1 (df = 33)	B2 (df = 35)	B1 (df = 33)	B2 (df = 35)	B1 (df = 33)	B2 (df = 35)
$y_i$	$0.2242x$ $+ 3.7771$	$0.193x$ $+ 5.1755$	$0.1612x$ $+ 0.4559$	$0.1971x$ $- 0.5097$	$0.4906x$ $- 3.5209$	$0.2967x$ $- 7.8495$
$\bar{x}$	20.7429	24.2973	49.4571	55.4595	49.4571	55.4595
$\bar{y}$	8.4286	9.8649	8.4286	9.8649	20.7429	24.2973
$\sum (x_i - \bar{x})^2$	37000.6871	4929.7297	8537.6871	19679.1892	8536.6871	19679.1892
$SSE$	418.4808	982.678	382.7325	477.6808	1645.9889	3197.2757
$SSR$	186.0906	183.6463	221.8389	688.6435	2054.6968	1732.4540
$SST$	604.5714	1166.3243	604.5714	1166.3243	3700.6857	4929.7297
$r^2$	0.3078	0.1575	0.3669	0.5904	0.5552	0.3154

Table 2: Regression lines for groups B1 and B2

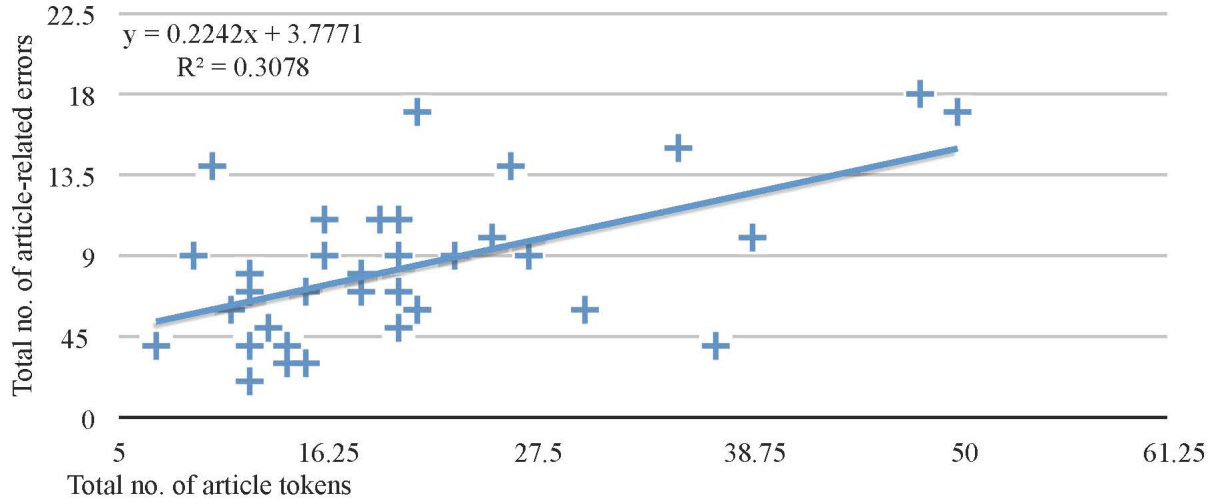


Figure 1: Article-related errors (ArE) vs. Article tokens (AT) (B1)

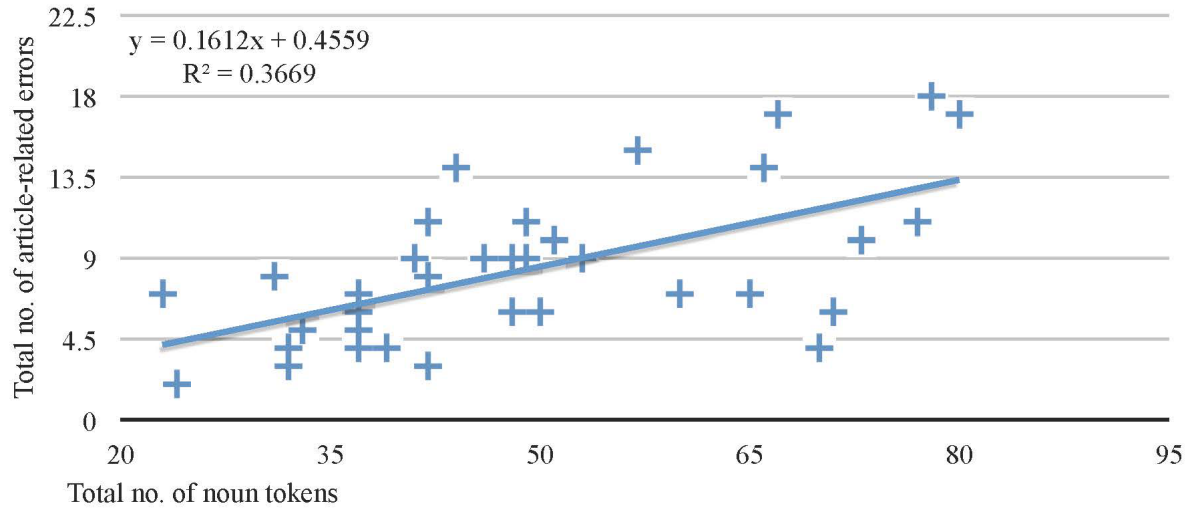


Figure 2: Article-related errors (ArE) vs. Noun tokens (NT) (B1)



With regard to group B2, the calculated regression line ( $y_i = 0.193x + 5.1755$ ) reveals a lower frequency of article-related errors per article token, as shown in Figure 2, above. The coefficient of determination ( $r^2$ ) suggests that the regression line is not a good fit, since only 15.75% of the variance in the amount of article-related errors can be explained by the use of articles itself. The standard error value ( $\sigma = 5.2987$ ) was obtained from the square root of the value resulting from the division of the SSE (982.678) by the degrees of freedom ( $n - 2 = 35$ ). There is confidence that the actual slope lies within the interval (0.0397, 0.3463) and it does not contain zero, so the null hypothesis can also be rejected. Finally, the  $t$ -test also reveals a significant relationship between the variables ( $t = 2.5563 > t_{\text{crit}} = 2.03$ ).

#### 4.1.2. Number of article-related errors (ArE) vs. Number of noun tokens (NT)

Because we are dealing with the same amount of article-related errors, the mean values for  $y$  and  $SST$  remain the same (cf. Table 2, above).

The regression line for group B1 was calculated as ( $y_i = 0.1612x + 0.4559$ ), which suggests that for every new noun ( $x$ ) we would expect the amount of article-related errors to increase by 0.1612 (cf. Figure 3). The coefficient of determination ( $r^2$ ) reveals that up to 36.69% of the variance in the amount of article-related errors can be explained by increasing noun use. The confidence interval for the slope was estimated at (0.0864, 0.236) and it does not contain zero, so the null hypothesis can again be rejected. Furthermore, the  $t$ -test suggests a significant linear relationship between the number of article-related errors and the number of noun tokens ( $t = 4.3686 > t_{\text{crit}} = 2.03$ ).

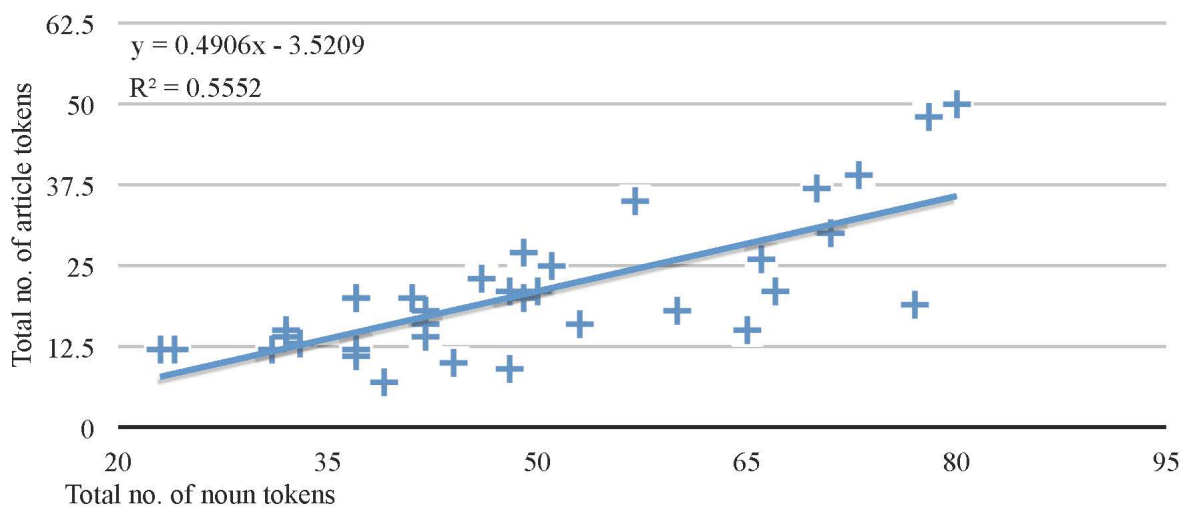


Figure 3: Article-related errors (ArE) vs. Noun tokens (NT) (B1)

With respect to group B2, the regression line ( $y_i = 0.1871x + 0.5097$ ) suggests a slightly higher frequency of article-related errors per noun token, as shown in Figure 4. The coefficient of determination as ( $r^2$ ) suggests a fairly good fit given that up to 59.04% of the variance in article-related errors can be explained by the number of nouns in the writing assignments. There is confidence that the slope lies within the interval (0.1336, 0.2407), and the null hypothesis can also be rejected as the interval does not contain zero. The  $t$ -test confirms a statistically significant linear relationship between the variables ( $t = 7.1141 > t_{\text{crit}} = 2.03$ ).

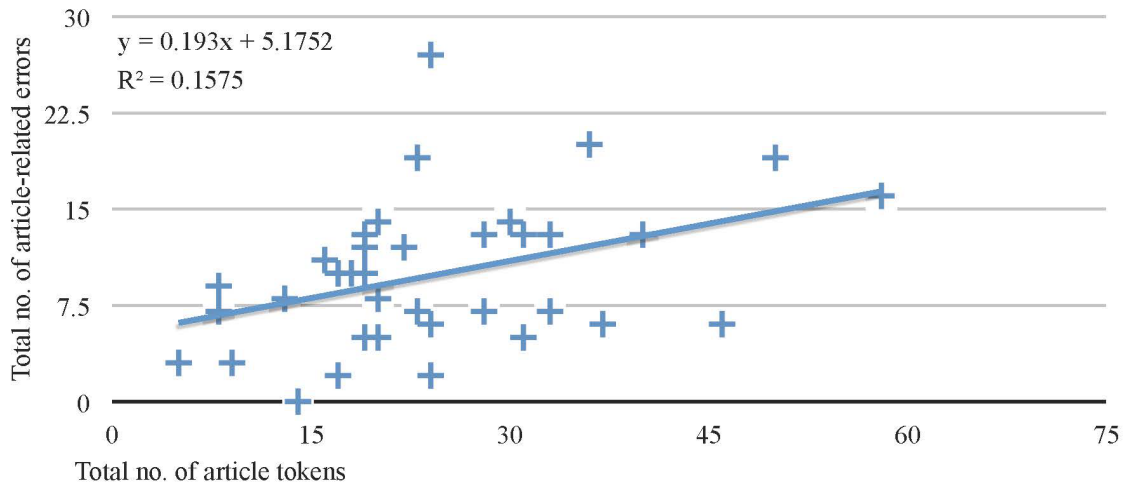


Figure 4: Article-related errors (ArE) vs. Article tokens (AT) (B2)

#### 4.1.3. Number of article tokens (AT) vs. Number of noun tokens (NT)

The regression line for group B1 ( $y_i = 0.4906x + 3.5309$ ) suggests for every new noun ( $x$ ) the amount of articles is expected to increase by 0.4906, as shown in Figure 5, below. The coefficient of determination ( $r^2$ ) indicates that up to 55.52% of the variance in the use of articles can be explained by the number of nouns used in each writing assignment. There is confidence in the interval (0.3354, 0.6458) containing the true slope of the regression line, and because the interval does not contain zero, the null hypothesis can be rejected. The  $t$ -test reveals a statistically significant relationship between the number of article tokens and noun tokens ( $t = 6.4215 > t_{\text{crit}} = 2.03$ ).

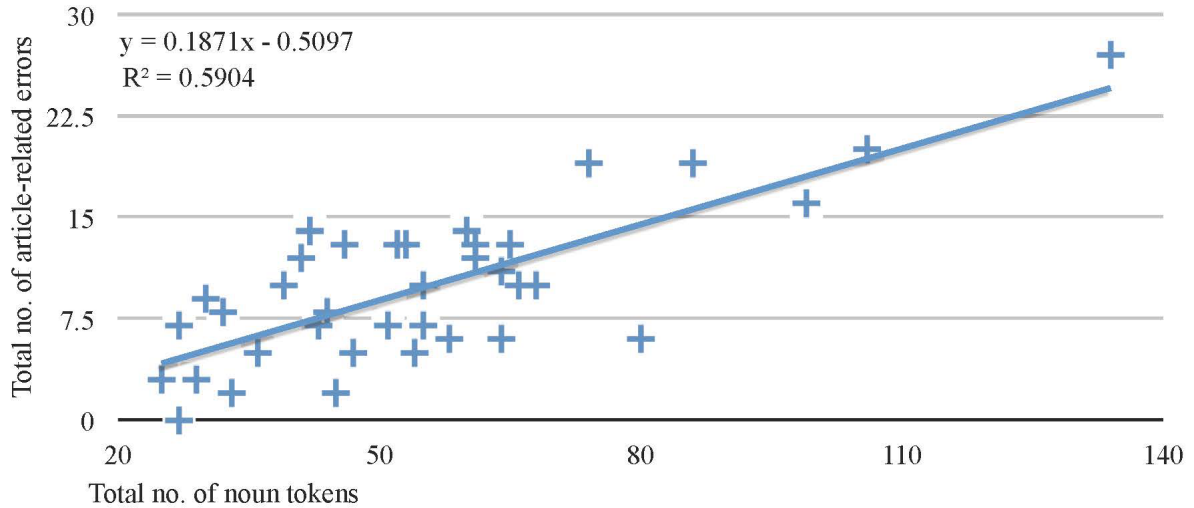


Figure 5: Article-related errors (ArE) vs. Noun tokens (NT) (B2)

As for group B2, the estimated regression line is ( $y_i = 0.2967x + 7.8425$ ) and this suggests a lower frequency of articles per noun token, as illustrated in Figure 6. However, the coefficient of determination ( $r^2$ ) does not suggest a good fit, since only 35.14% of the variance in article tokens can be explained by the amount of nouns. There is confidence that the slope lies within the interval (0.1584, 0.435), and since the interval does not contain zero, the null hypothesis can again be rejected. The result of the  $t$ -test confirms that the relationship between the variables is statistically significant ( $t = 4.3568 > t_{\text{crit}} = 2.03$ ).

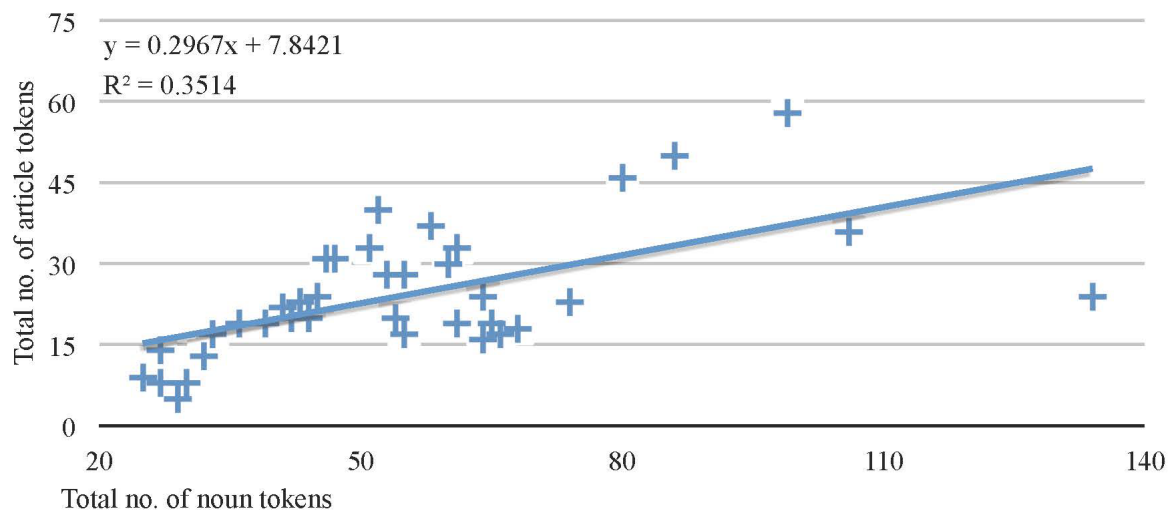


Figure 6: Article tokens (AT) vs. Noun tokens (NT) (B2)

#### 4.1.4. Confidence intervals for a specific value of article and noun tokens ( $x^*$ )

Finally, confidence intervals have been generated for the two groups separately in order to estimate: (a) a mean value of article-related errors every  $x^*$  article tokens; (b) a mean value of article-related errors every  $x^*$  noun tokens; and (c) a mean number of articles every  $x^*$  noun tokens. However, it should be stressed that confidence intervals were calculated for the rounded up value of  $\bar{x}$  in order to get predicted average mean value of  $y^*$  for a real value of  $x^*$ .<sup>11</sup>

##### A) Group B1

Firstly, by applying the regression equation is ( $y_i = 0.2242x + 3.7771$ ) an average mean of 8.4853 article-related errors is estimated for every 21 article tokens. The confidence intervals was then calculated by adding the margin of error ( $\pm t_{\alpha/2} s_{y^*}$ ), and it predicts that 95% of the sampled population in group B1 would produce a mean between 7.2624 and 9.7082 article-related errors for every 21 article tokens.

Secondly, the regression model ( $y_i = 0.1612x + 0.4559$ ) suggests an average mean of 8.3547 article-related errors for every 50 noun tokens has been estimated. According to the confidence interval, 95% of the sampled population are predicted to make a mean between 7.1849 and 9.5245 article-related errors for every 50 noun tokens.

Thirdly, according to the regression line ( $y_i = 0.4906x - 3.5209$ ), an average mean of 21.0091 articles is estimated for every 50 noun tokens. The generated confidence interval estimate that 95% of the sampled population will produce a mean between 18.5831 and 23.4351 articles for every 50 noun tokens. However, it should be borne in mind that this is an estimate of article use and does not represent correct article use.

---

<sup>11</sup> Although it would have been more precise to calculate it for the  $x$  value, it would not have yielded an accurate picture. This occurs because no student produces one noun/article and a half. Therefore, values for  $x^*$  have been rounded up.

## B) Group B2

Firstly, the regression equation ( $y_i = 0.193x + 5.1755$ ) predicts an average mean of 10.0005 article-related errors for every 25 article tokens. The confidence interval was then calculated by adding the error margin, and it predicts that 95% of the sampled population in group B2 would produce a mean between 8.8298 and 11.1712 article-related errors for every 25 article tokens.

Secondly, the regression model ( $y_i = 0.1871x - 0.5097$ ) predicts an average mean of 9.9679 article-related errors every 56 noun tokens. According to the confidence interval, 95% of the participants are expected to produce a mean between 8.7353 and 11.2005 article-related errors for every 56 noun tokens.

Finally, the regression line ( $y_i = 0.2967x + 7.8425$ ) predicts an average mean of 24.4577 articles for every 56 nouns. The generated confidence interval estimate that 95% of the sampled population will produce a mean between 21.2687 and 27.6467 articles for every 56 nouns in the B2 group. Once again, this represents an estimate of article use but does not represent correct article use.

## 4.2. Regression analysis of the intermediate Level Group (B1+B2)

### 4.2.1. Number of article-related errors (ArE) vs. Number of article tokens (AT)

The calculated regression line ( $y_i = 0.2115x + 4.3938$ ) suggests that for every new article environment ( $x$ ), we would expect the amount of error to increase by 0.2115, as illustrated in Figure 7, below. The coefficient of determination ( $r^2$ ) reveals that only 21.91% of the variance in article-related errors can be explained by using the number of article tokens. The confidence interval (0.1318, 0.2912) does not contain zero so the null hypothesis (the mean average number of article-related errors fits the data better) can be rejected. Furthermore, the  $t$ -test reveals a significant linear relationship between the number of article-related errors and the number of article tokens ( $t = 4.4340 > t_{\text{crit}} = 1.67$ ).

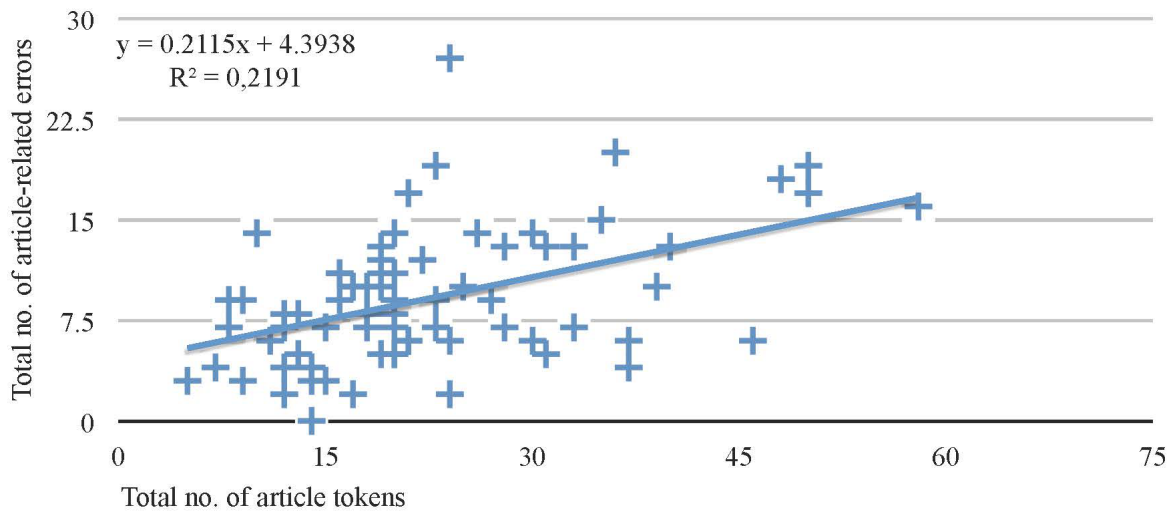


Figure 7: Article-related errors (ArE) vs. Article tokens (AT) (B1+B2)

#### 4.2.2. Number of article-related errors (ArE) vs. Number of noun tokens (NT)

Due to the fact that we are dealing with the same amount of ArE, the mean values for  $y$  remain the same, as illustrated in Table 3, below. In this case, the regression line ( $y_i = 0.1806x - 0.3218$ ) suggests that for every new noun ( $x$ ) the amount of article-related errors is expected to increase by 0.1806, as can be seen in Figure 8. The coefficient of determination ( $r^2$ ) indicates a fairly good fit, since up to 52.06% of the variance in article-related errors can be explained by the number of nouns in each composition. There is confidence that the slope lies within the interval (0.146, 0.2152), and since it does not contain zero, the null hypothesis can then be rejected. In addition, the  $t$ -test confirms a statistically significant linear relationship between the variables ( $t = 8.7346 > t_{\text{crit}} = 1.67$ ).

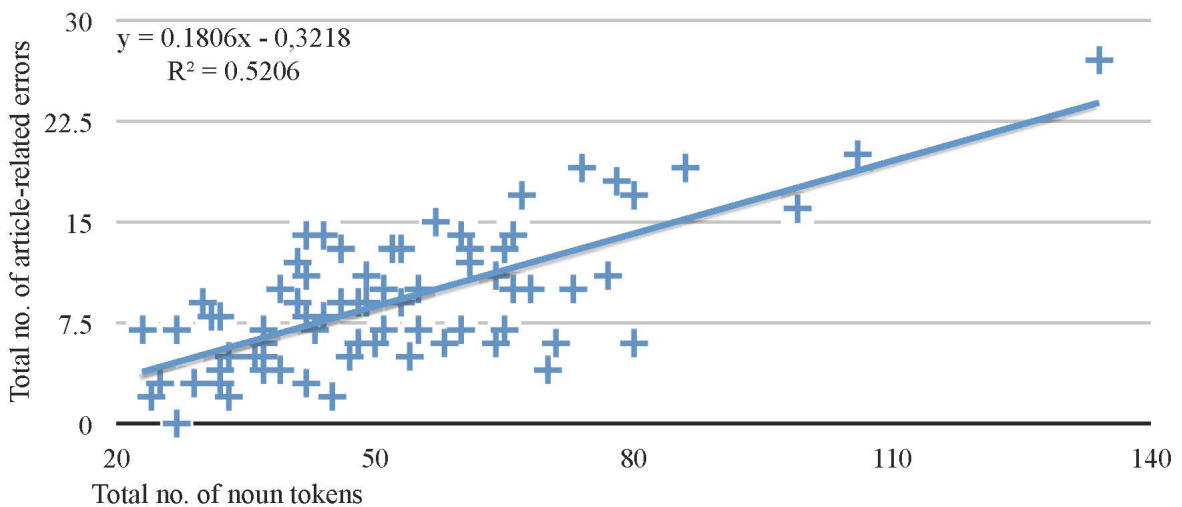


Figure 8: Article-related errors (ArE) vs. Noun tokens (NT) (B1+B2)

	ArE vs. AT	ArE vs. NT	AT vs. NT
$y_i$	$0.2115x + 4.3938$	$0.1806x - 0.3218$	$0.3607x + 3.6184$
$\bar{x}$	22.5694	52.5417	52.5417
$\bar{y}$	9.1667	9.1667	22.5694
$\sum (x_i - \bar{x})^2$	8857.6528	28863.8750	28863.8750
$SSE$	1411.8733	866.6796	51202.6274
$SSR$	396.1267	941.3204	37755.0253
$SST$	1808	1808	8857.6528
$r^2$	0.2191	0.5206	0.4239

Table 3: Regression lines for the Intermediate Level group (n = 72)

#### 4.2.3. Number of article tokens (AT) vs. Number of noun tokens (NT)

In this case, the resulting regression line ( $y_i = 0.3607x + 3.6184$ ) indicates that for every new noun ( $x$ ) the amount of article tokens is expected to increase by 0.3607 (cf. Figure 9). The coefficient of determination ( $r^2$ ) suggests that 42.39% of the variation in the number of article tokens can be explained by the number of nouns in each composition. There is confidence in the interval (0.2768, 0.4446) containing the true slope of the regression line. The null hypothesis can be rejected because the interval does not contain zero. Finally, the  $t$ -test reveals a significant relationship between the variables ( $t = 7.1853 > t_{\text{crit}} = 1.67$ ).

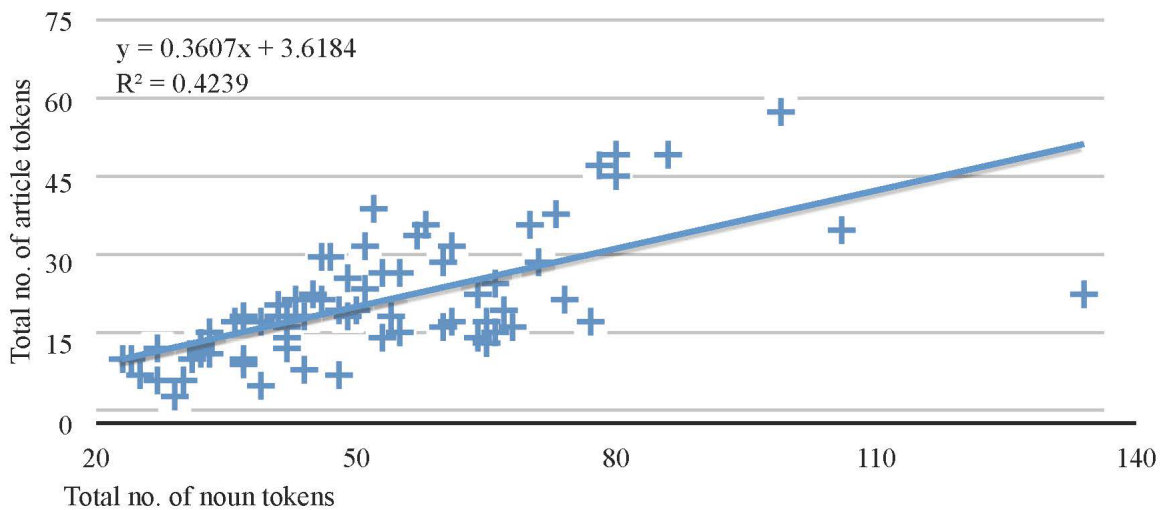


Figure 9: Article tokens (AT) vs. Noun tokens (NT) (B1+B2)

#### 4.2.4. Confidence intervals for a specific value of article and noun tokens ( $x^*$ )

Finally, confidence intervals have been generated in order to estimate: (a) a mean value of article-related errors every  $x^*$  article tokens; (b) a mean value of article-related errors every  $x^*$  noun tokens; and (c) a mean number of articles every  $x^*$  nouns. However, it should be stressed that confidence intervals were calculated for the rounded value (cf. note 11, above) of  $\bar{x}$  in order to get predicted average mean value of  $y^*$  for a real value of  $x^*$ .

Firstly, according to the regression model ( $y_i = 0.2115x + 4.3938$ ), an average mean of 9.2583 article-related errors is predicted for every 23 article tokens. After adding the margin of error ( $\pm t_{\alpha/2} s_{y^*}$ ), the confidence interval indicates that 95% of the sampled population in the Intermediate Level group was predicted to produce a mean between 8.3734 and 10.1432 article-related errors for every 23 article tokens.

Secondly, the regression model ( $y_i = 0.1806x - 0.3218$ ) predicts that the students will produce a mean value of 9.25 article-related errors for every 53 noun tokens. Moreover, the confidence interval suggests that 95% of the students will produce a mean between 8.557 and 9.943 article-related errors for every 53 nouns.

Thirdly, the regression line ( $y_i = 0.3607x + 3.6184$ ) predicts an estimate of 22.7355 articles to be produced every 53 noun tokens. If we add the margin of error, the resulting interval predicts that 95% of the sampled population will produce a mean between 21.0541 and 24.4169 articles for every 53 nouns.

## 5. DISCUSSION

Although the aim of this study is it is not to analyze the types of article-related errors, or their possible motivations, a link can be established between our results and Ionin *et al.*'s (2004) Fluctuation Hypothesis. As can be seen in Figures 1–9, there is a great variance in terms of article-related errors, and the regression models – though they do not represent perfect fits – relate better to the observed number of errors than the mean line. This shows that the amount of article-related errors is hard to predict even in participants with a high degree of homogeneity, as is the present research.

With regard to our first research question, all the *t*-tests confirm statistically significant relations between the number of article-related errors and the number of



article tokens. This is also supported by the fact that none of the confidence intervals contain zero. This result was to be expected since more article tokens represent more linguistic contexts in which articles may occur.

Significant relations have also been found in relation to our second research question, that is, between the number of article-related errors and the number of noun tokens. Moreover, in this case the coefficients of determination ( $r^2$ ) indicate good fits (over 50%). This seems to suggest that a piece of text which makes use of a higher number of noun phrases will require more articles by default and, as a consequence, the number of article-related errors can be expected to increase. That is why it is essential to analyze article-related errors in relation to the total number of noun phrases, i.e. article environments, and not just to the total amount of words per one hundred words. In order to deal with this issue, Testa (2019) suggests assigning an *error coefficient* (between 0 and 1) which results from dividing the total number of article-related by the total number of article environments. In this way, two separate learners who would make two errors in a total of four article environments, and fifteen errors in a total in thirty article environments would both share an error coefficient of 0.5. Conversely, a sentence like *Me gusta \*uno vodka* ('me.DAT likes INDEF.MSG vodka') would mean an error coefficient of 2 (INDEF-for-DEF; M-for-F).

Finally, as for research question number 3, a significant relationship has been attested between the number of article tokens and the number of noun tokens. Moreover, the coefficient of determination ( $r^2$ ) indicates a good fit in the B1 group. In other words, these are cases in which over 50% of the variance in article-related errors and article tokens can be explained by the total number of noun tokens. Nevertheless, the coefficient of determination with regard to the relationship between AT and NT should be taken with caution, because more frequent article usage can be rarely explained without resorting to a higher amount of linguistic contexts in which article take place, namely in noun phrases.

## 6. CONCLUDING REMARKS

The present study has drawn attention to the need to consider the number of article environments, i.e. article tokens in noun phrases, when analyzing article-related errors.

Statistically significant relations have been attested between these variables and they fit the data better than the mean average number of article-related errors an article tokens.

Moreover, the evidence with regard to LX English and LX Spanish article acquisition suggests that students lack knowledge as regards the semantic traits associated to article use. This calls for a deeper analysis of such traits, particularly at an intermediate level of learning given that the article system is one of those categories that does not seem to be acquirable by means of comprehensible input alone (cf. Pica 1985; Master 1994).

Finally, some key questions that have not been addressed in this study remain open. One aspect that calls for more in-depth treatment is the effectiveness of pedagogical interventions in terms of L2 article instruction (cf. Ekiert and Han 2016). For instance, it would be interesting to analyze whether a Spanish-oriented version of Master's binary framework is a helpful pedagogical tool when it comes to teaching the Spanish article system. Moreover, even though Master's (1997) distinction of two types of Ø article ('indefinite' Ø<sub>1</sub> vs. 'definite' Ø<sub>2</sub>) seems to apply to Spanish in most cases, further research is required to show whether this distinction may accelerate the acquisition of the Spanish article system by [–ART] L1 learners.

#### REFERENCES

- Allan, Keith. 1986. *Linguistic Meaning*. London: Routledge.
- Allen, Virginia F. 1983. *Techniques in Teaching Vocabulary*. Oxford: Oxford University Press.
- Bailey, Natalie, Carolyn Madden and Stephen Krashen. 1974. Is there a natural sequence in adult second language learning? *Language Learning* 24/2: 235–243.
- Birner, Betty and Gregory Ward. 1994. Uniqueness, familiarity, and the definite article in English. In Susanne Gahl, Andy Dolbey and Christopher Johnson eds. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, California: Berkeley Linguistics Society, 93–102.
- Ekiert, Monika and ZhaoHong Han. 2016. L1-fraught difficulty: The case of L2 acquisition of English articles by Slavic speakers. In Rosa Alonso ed. *Crosslinguistic Influence in Second Language Acquisition*. Bristol: Multilingual Matters, 147–172.
- Fernández Jódar, Raúl. 2006. *Análisis de Errores Léxicos, Morfosintácticos y Gráficos en la Lengua Escrita de los Aprendices Polacos de Español*. Poznań, Poland: Adam Mickiewicz University dissertation.
- Fernández Jódar, Raúl. 2017. El artículo. In Waczesław Nowikow ed. *Gramática Contrastiva Español-polaco*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 353–377.

- García Mayo, María del Pilar. 2008. The acquisition of four non-generic uses of the article the by Spanish EFL learners. *System* 36/4: 550–565.
- George, Herbert. 1972. *Common Errors in Language Learning*. Rowley: Newbury House.
- Hakuta, Kenji. 1976. A case study of a Japanese child learning English as a second language. *Language Learning* 26: 321–351.
- Harb, Mustafa A. 2014. A closer look at the English article system: Internal and external sources of difficulty revisited. *International Journal of Linguistics* 6/4: 87–101.
- Hawkins, John. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27/2: 405–442.
- Hewson, John. 1972. *Article and Noun in English*. The Hague: Mouton.
- Hidalgo, Andrea. 2015. *Estudio Contrastivo Español-chino: El Artículo Indefinido y su Tratamiento en los Manuales de Enseñanza de Español como Segunda Lengua*. Córdoba, Argentina: Universidad Nacional de Córdoba dissertation.
- Ionin, Tania, Heejeong Ko and Kenneth Wexler. 2004. Article semantics in L2-acquisition: The role of specificity. *Language Acquisition* 12/1: 3–69.
- Ionin, Tania, María L. Zubizarreta and Salvador Maldonado. 2008. Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua* 118/4: 554–576.
- Isabelli-Garcia, Christina and Rachel Slough. 2012. Acquisition of the non-generic definite article by Spanish learners of English as a foreign language. *OnOmázein* 25/1: 95–105.
- Jaensch, Carol. 2008. L3 acquisition of articles in German by native Japanese speakers. In Roumyana Slabakova, Jason Rothman, Paula Kempchinsky and Elena Gavruseva eds. *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference*. Somerville, Massachusetts: Cascadilla Proceedings Project, 81–89.
- Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49/3: 501–536.
- Jiang, Nan, Eugenia Novokshanova, Kyoko Masuda and Xin Wang. 2011. Morphological congruency and the acquisition of L2 morphemes. *Language Learning* 61/3: 940–967.
- Kharm, Nayef. 1981. Analysis of the errors committed by Arab university students in the use of the English definite/indefinite articles. *International Review of Applied Linguistics in Language Teaching* 19/4: 333–345.
- Konieczna-Twardzikowa, Jadwiga. 1992. Caso y definitud en la lengua española desde la perspectiva polaca. *Estudios hispánicos* 2: 171–175.
- Krashen, Stephen. 1982. *Principles and Practice in Second Language Acquisition*. London: Pergamon.
- Krashen, Stephen and Tracy Terrell. 1983. *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon Press.
- Lapesa, Rafael. 2000 [1974]. Un, una como artículo indefinido en español. In Rafael Lapesa ed. *Estudios de Morfosintaxis Histórica del Español*. Madrid: Gredos, 477–487.
- Lema, Rebeca. 2016. Las interferencias del español L2 en el estudio del gallego L2. *Itinerarios: Revista de Estudios Lingüísticos, Literarios, Históricos y Antropológicos* 23: 61–78.
- Lin, Tzu Ju. 2003. *La Adquisición y el Uso del Artículo por Alumnos Chinos*. Alcalá de Henares, Spain: Universidad de Alcalá dissertation.

- Little, David. 1994. Words and their properties: Arguments for a lexical approach to pedagogical grammar. In Terence Odlin ed. *Perspectives in Pedagogical Grammar*. Cambridge: Cambridge University Press, 99–122.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Master, Peter. 1986. *Measuring the Effect of Systematic Instruction in the English Article System*. Los Angeles, California: University of California (Unpublished Paper).
- Master, Peter. 1987. *A Cross-linguistic Interlanguage Analysis of the Acquisition of the English Article System*. Los Angeles, California: University of California dissertation.
- Master, Peter. 1990. Teaching the English articles as a binary system. *TESOL Quarterly* 24/2: 461–478.
- Master, Peter. 1994. The effect of systematic instruction on learning the English article system. In Terence Odlin ed. *Perspectives on Pedagogical Grammar*. Cambridge: Cambridge University Press, 229–252.
- Master, Peter. 1995. Consciousness raising and article pedagogy. In Diane Belcher and George Braine eds. *Academic Writing in a Second Language*. Norwood, NJ: Ablex, 183–204.
- Master, Peter. 1997. The English article system: Acquisition, function, and pedagogy. *System* 25/2: 215–232.
- Master, Peter. 2002. Information structure and English article pedagogy. *System* 30: 331–348.
- Mizuno, Mitsuharu. 1999. Interlanguage analysis of the English article system: Some cognitive constraints facing the Japanese adult learners. *International Review of Applied Linguistics in Language Teaching* 37/2: 127–153.
- Nowikow, Waczesław. 2017. Tiempos verbales. In Waczesław Nowikow ed. *Gramática Contrastiva Español-polaco*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 127–178.
- Odlin, Terence. 1989. *Language Transfer*. Cambridge: Cambridge University Press.
- Park, Sung B. 2006. *The Acquisition of Written English Articles by Korean Learners*. Carbondale, Illinois: Southern Illinois University dissertation.
- Pica, Teresa. 1983. The article in American English: What the textbooks don't tell us. In Nessa Wolfson and Elliot Judd eds. *Sociolinguistics and Language Acquisition*. Rowley, Massachusetts: Newbury House Publishers, 222–233.
- Pica, Teresa. 1985. The selective impact of classroom instruction on second language acquisition. *Applied Linguistics* 6/3: 214–222.
- Pienemann, Manfred. 1998. *Language Processing and Second Language Development: Processability Theory*. Amsterdam: John Benjamins.
- Ringbom, Håkan. 1987. *The Role of the First Language in Foreign Language Learning*. Clevedon: Multilingual Matters.
- Ringbom, Håkan. 2011. Perceived redundancy or crosslinguistic influence? What L3 learners' material can tell us about the causes of errors. In Gessica De Angelis and Jean-Marc Dewaele eds. *New Trends in Crosslinguistic Influence and Multilingualism Research*. Bristol: Multilingual Matters, 19–24.
- Ringbom, Håkan. 2016. Comprehension, learning and production of foreign languages: The role of transfer. In Rosa Alonso Alonso ed. *Crosslinguistic Influence in Second Language Acquisition*. Bristol: Multilingual Matters, 38–52.
- Sabir, Mona H. 2015. *Explicit Instruction and Translation: A Generative View of the Acquisition of English Articles*. Leeds, United Kingdom: University of Leeds dissertation.

- Schulz, Eckehard. 2004. *A Student Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
- Schwarz, Florian. 2009. *Two Types of Definites in Natural Language*. Amherst, MA: University of Massachusetts dissertation.
- Şekerci Arıbaş, Derya and Filiz Cele. 2019. 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*. Advanced online publication .
- Shen Jie. 2012. *El Artículo en la Enseñanza de ELE. Estudiantes de Origen Chino*. Barcelona, Spain: Universidad de Barcelona dissertation.
- Sun, Ganzhao. 2016. The acquisition of English articles by second language learners: The sequence, differences, and difficulties. *SAGE Open* 6/1: 1–8.
- Tarrés Chamorro, Iñaki. 2002. *El Uso del Artículo por Estudiantes Polacos de E/LE*. Barcelona, Spain: Universidad de Barcelona dissertation.
- Testa, Martín. 2019. *Análisis de Variables Psicolingüísticas en la Interlengua de Alumnos Polacos de Español L3*. Warsaw, Poland: University of Warsaw dissertation.
- Testa, Martín. In press. The acquisition of Spanish articles by L1 Polish students. *Neofilologia: Perspektywy Transdyscyplinarności*.
- Trenkic, Danijela. 2002. Establishing the definiteness status of referents in dialogue (in languages with and without articles). *University of Cambridge Working Papers in English and Applied Linguistics* 7: 107–131.
- VanPatten, Bill and Teresa Cadierno. 1993. Explicit instruction and input processing. *Studies in Second Language Acquisition* 15/2: 225–243.
- Widdowson, Henry. 1988. Grammar, nonsense, and learning. In William Rutherford and Michael Sharwood Smith eds. *Grammar and Second Language Teaching*. New York: Newbury House, 146–155.
- Yoo, Isaiah. 2009. The English definite article: What ESL/EFL grammars say and what corpus findings show. *Journal of English for Academic Purposes* 8/4: 267–278.

*Corresponding author*

Martin Testa  
 University of Warsaw  
 Institute of Iberian and Ibero-American Studies.  
 ul. Oboźna 8  
 00–332 Warsaw  
 Poland  
 e-mail: m.testa@uw.edu.pl

received: August 2019  
 accepted: October 2019