



Acquisition of the non-generic definite article by Spanish learners of English as a foreign language

Christina Isabelli-García
Illinois Wesleyan University
Estados Unidos

Rachel Slough
University of Wisconsin - La Crosse
Estados Unidos

Abstract

This study explains the acquisition of the non-generic uses of the English definite article *the* by L2 learners. Chilean university students completed a questionnaire containing deleted obligatory uses of *the*. In all, four identified categories showed to have different accuracy rates. Of the four categories, the one most similar to the L1 was the least difficult to acquire while the most different resulted as the most difficult. This concurs with other research and can be explained by L1 rules that are transferred to the L2. The other two categories did not show the same acquisition rate as compared to previous studies and cannot be explained by L1 transfer given that both languages share the same rules. Time spent studying the English increases accuracy rates across all categories. Knowing that difficulty with the definite article can be attributed to L1 interference is useful when implementing pedagogical recommendations to teach the English article system.

Keywords: definite article acquisition; EFL; L1 transfer; non-generic uses; Spanish.

Afiliación: Christina Isabelli-García: Illinois Wesleyan University, USA. – Rachel Slough: University of Wisconsin-La Crosse, USA.

Correos electrónicos: cisabell@iwu.edu; slough.rach@uwlax.edu

Dirección postal: Christina Isabelli-García: Illinois Wesleyan University, PO Box 2900, Bloomington, IL 61702, USA. – Rachel Slough: University of Wisconsin-La Crosse, 1631 Pine Street, La Crosse, WI 54601, USA.

Fecha de recepción: abril de 2011

Fecha de aceptación: marzo de 2012

1. Introduction

The acquisition of the English definite article (*the*) is a difficult process for L2 learners (García Mayo, 2008; Ionin, *et al.*, 2008; Master, 1995, 1994; Dulay, *et al.*, 1982; Grannis, 1972). Researchers over the years have studied the source of this difficulty (Trenkic, 2007; Hawkins, *et al.*, 2006; Ionin, *et al.*, 2004; Master, 2002) and according to Liu and Gleason (2002) these studies examined the use of *the* as a whole but results from Master (1995) implied that “certain uses of the definite article might be more difficult than others” (Liu & Gleason, 2002: 5). In addition, several of these studies used participants from multiple language backgrounds whose L1 ranged in difference/similarity to the English use of the definite article. Attempting to identify different categories of the definite article and subsequently examining them separately as well as focusing on one language group may contribute reliable data on the issues with the acquisition of *the*. Knowing that a difficulty with the definite article can be attributed to L1 interference or not will be useful when implementing pedagogical recommendations to teach the English article system.

Huebner’s (1985) research, although focused on explaining form-function relationships over time, used the definite article to help classify the semantic uses of noun phrase reference. In doing so, he set forward a taxonomy of the English article system based on Bickerton’s (1981) semantic definitions of the noun phrase. We single out here the types that can be marked in English with the definite article *the*: generic nouns and referential definites. An adaptation of Huebner’s taxonomy is found in Díez-Bedmar and Papp (2008: 152) where the authors added another type that may be marked with the definite article: idioms and other conventional uses. Of these types of article categorizations, this study focuses specifically on the non-generic use of the English definite article that includes referential definites, idioms and other conventional uses.

1.1. Literature review

Identification of the types of non-generic use varies greatly from three –as outlined above– and four types (Liu & Gleason, 2002) to eight types (Hawkins, 1978). Invariably, the researchers that have identified fewer types have combined categories with

those that include more types. For the purposes of this study, which aims to replicate Liu and Gleason (2002) and García Mayo (2008), we identify the non-generic uses similar to the authors' four categories: cultural, situation, structural, and textual.

The participants in Liu and Gleason's (2002) study represented 18 native languages, three-fourths of which were from non Indo-European languages (such as Chinese, Korean and Japanese), languages that have different rules that restrict the use of *the*. The authors stated that since the group size of the Indo-European language speakers (languages that are more similar to English with regards to the use of *the* but may maintain some difference) was much smaller their "findings may not be reliable" (2002: 18). The present study used participants from only one Indo-European language group, Spanish, similar to García Mayo (2008). This study, however, used participants that spoke Southern Cone Spanish whereas García Mayo's participants were Peninsular Spanish (Spain). According to Ionin *et al.* (2008) Spanish-speakers learning English transferred the semantics of Spanish (L1) on to English (L2) articles. For those English definite articles that share Spanish semantic rules, L1-transfer may have allowed the learners to perform with a high accuracy rate.

As illustrated in Table 1, the three categories that share the same rules regarding non-generic use of the definite article are situation, structural, and textual uses.

The cultural uses sometimes differ in Spanish and English and within the same language as well. For example, with regards to some diseases, like "cancer", the definite article is omitted in English and Spanish: "She died of cancer", *Ella murió de cáncer*. While with the noun "plague" the article is not omitted in English while it is omitted in Spanish. For example the sentence "She died of the plague" was seen written with the feminine definite article *la* "... *murió de la plaga...*" ("*La peste negra*", n.d.). A difference between Spanish and English exists with some geographical names. In English the article is omitted before "lake" in "Lake Erie is in the United States" but not in Spanish *El lago Erie está en los Estados Unidos*.

TABLE 1
Categorizations of the non-generic definite article

Category	Use
Cultural	Used with a noun that is a unique and well-known referent in a speech community. The Atlantic Ocean is very blue. The flu is a common phenomenon in wintertime.
Situation	Used when the referent of a first-mention noun can be sensed directly or indirectly by the interlocutors. The teacher was lecturing in a hot room and asked his student, "Could you open the door?"
Structural	Used with a first-mention noun that has a modifier. The black dog that was eating a hot dog was mine.
Textual	Used with a noun that has been previously referred to or is related to a previously mentioned noun. <i>Arturo read a book yesterday. Today, he sold the book.</i>

1.2. Research questions

The research questions for this present study are: (i) Which of the four uses of the non-generic definite article presents more/less difficulty for L2 learners of English; (ii) Where do the differences in acquisition lie among the four groups and what can explain the differences; and, (iii) Do different rates of acquisition exist for each use as compared to each other? We hypothesize that because the four types vary considerably in context and rule complexity, they will not be equally difficult for ESL students and hence not acquired at the same time.

2. Methodology

2.1. Participants

The participant population was recruited from the adult university student population of the English Pedagogy program at a private university in Chile. The seventy participants were enrolled in one of the second through fifth levels of English (excluding the first level) with a reported native language of Spanish. All participants reported having taken English courses

at least through their four years of high school with some that began English study in elementary school. In total there were: 28 second year students; 15 third year students; 11 fourth year students, and 16 fifth year students.

2.2. Procedures

The data-collection tool was a ninety-one-sentence questionnaire containing sixty deleted obligatory uses of *the* and, as control items, forty zero articles where the definite article was not permitted. Each of the sixty deleted obligatory uses of the definite article fell into one of the four categories: cultural, situation, structural or textual. The participants were asked to analyze each sentence and insert *the* where they deemed appropriate. After the questionnaires were collected a number was assigned to each instance where the student should have inserted the article. Each instance was then categorized into one of the four categories and a spreadsheet was created with this structure. The data from the questionnaires was coded as “0” for a correct insertion and “1” for no insertion and therefore incorrect. An instance of no insertion was labeled as missed article use in the remainder of the document. Instances of incorrect insertion (or overuse) were tallied but not included in this study since most overuse dealt with the generic use of the definite article, which was not included in the scope of this study.

Several calculations were carried out. A one-way ANOVA measured the effect that a specific category had on the rate of missed article use, fixing the English proficiency level. A one-way mixed ANOVA measured the effect of English proficiency level, fixing the category of non-generic definite articles. We used a two-way mixed ANOVA to calculate whether the category and English proficiency level significantly affected the mean rate. We calculated the mean rate of missed use of the definite article in each category, fixing the English proficiency level, and vice versa. We then used the Tukey multiple comparisons of means to compare the differences between the four categories by English proficiency level.

3. Results and discussion

Looking at the mean rate of missed use of the definite article across categories, the cultural category had the highest rate of errors (.445) with textual at .27 and the situation and structural categories having the lowest rate of errors, .1325 and .1575 respectively. This is illustrated in Figure 1 and Table 2. All vary significantly from each other ($p < .0001$) except between the categories of situation and structural.

FIGURE 1
Comparison of means of missed obligatory use of the definite article

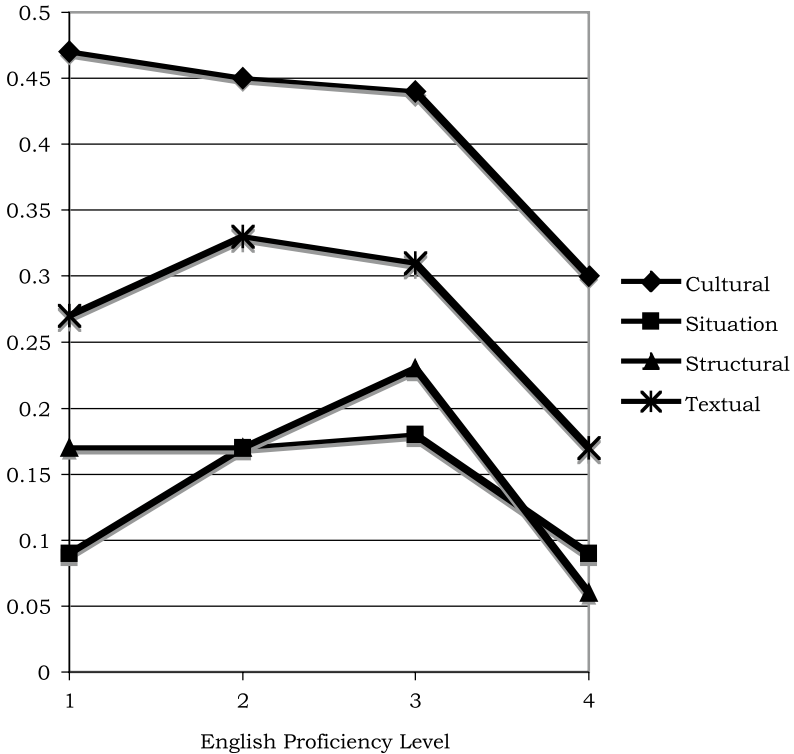


TABLE 2
Mean rate of missed obligatory use of the definite article

English Proficiency Level	Cultural	Situation	Structural	Textual	Mean, fixed proficiency level
2	.47	.09	.17	.27	.249
3	.45	.17	.17	.33	.282
4	.44	.18	.23	.31	.291
5	.30	.09	.06	.17	.155
Mean, fixed category level	.445	.1325	.1575	.27	

The mean rate of the missed use of the definite article across the lower English proficiency levels 2, 3, and 4, varies slightly. Level 2 has a mean of .249 and it lowers to .1555 at the top proficiency level 5 showing a significant overall increase in correct use of the obligatory definite article ($p < .001$). Levels 3 and 4 have a slightly more increased error rate than level 2.

All categories in proficiency level 2 evidence significant differences between categories with a p-value of 0 suggesting a strong category effect. The difference between levels 2 and 5 (.47 and .30, $p = .046$) suggests that by level 5 students make significant acquisitional progress making fewer errors in the cultural category. Slight improvement, but not significant, is seen in levels 3 and 4 suggesting that improvement is minor but by the time the student is at level 5, the data show that the overall improvement is significant.

In the situation category, students at proficiency levels 2 and 5 had the same score (.09) whereas many more errors were made at levels 3 and 4 (.17 and .18, respectively). We see here more evidence of destabilization or restructuring, which takes the form of U-shaped progress with significant backtracking (see Figure 1). The improvement seen in the situation category is not a linear, as seen in the cultural category, rather it is an upside down U-shaped progression. The U-shaped growth concept focuses on the fact that some behaviors appear, disappear, and then apparently reappear over time. U-shaped learning deals with a learner first having the correct hypothesis, then changing it to an incorrect hypothesis and then relearning the correct one.

In the structural category, students at proficiency levels 2 and 3 performed the same with a mean error rate of .17. Level 4 students saw an increase in error rate to a mean of .23 and then by level 5, a significant drop in error rate was seen to .06 ($p = .014$). Overall, structural uses of the definite article improved with time on language with backtracking occurring at proficiency level 4.

In the textual category students at the lower proficiency level had a mean error rate of .27, which increased to .33 at level 3, lowered to .31 at level 4, and lowered once more to .17 at level 5. A significant improvement occurred from level 3 to 5 ($p = .038$). For the textual category a similar kind of U-shape learning took place with a significant improvement seen from level 3 to 5.

To explain this seemingly counterintuitive direction of improvement (increasing error rate at the higher proficiency levels), we must look to discussions that have focused on scaffolding and restructuring in other L2 acquisition research. According to Donato (1994) when language learners are in a context where interaction occurs with a more advanced speaker, they may notice new or correct structures in the advanced speaker's language or feedback. This attention to "noticing" (Gass & Varonis, 1994) –a mismatch or discrepancy between what they know about the language and what native speakers produce– leads to the notion of scaffolding to describe the process by which learners develop their interlanguage through interaction (Donato, 1994; Lantolf & Appel, 1994). According to Lightbown (1985) restructuring also occurs through a process of destabilization in which an increase in error rate in one area may reflect an increase in complexity or accuracy in another, followed by overgeneralization of a newly acquired structure. When additional syntactic patterns become available to learners, restructuring or destabilization occurs. This destabilization is at the base of language change and explains the temporary increase in error rate.

Conclusions from Carlucci *et al.* (2006) led us to ask if backtracking, as seen in U-shaped learning, is necessary to achieve full power of algorithmic learning. The authors suggested that returning to wrong conjectures complements the paradigm of U-shaped learning when a learner returns to old correct conjectures. In the cases that they analyzed, the authors showed that returning to wrong conjectures is necessary to achieve

full learning power. We view backtracking and restructuring as serving the same functions in language learning.

4. Conclusions

The four non-generic uses of the English definite article have different accuracy rates in the performance of L2 language learners, with the situational category being the least difficult and the cultural category being the most difficult, concurring with conclusions from García Mayo (2008) and Liu and Gleason (2002). Our conclusions for the two other categories, textual and structural, showed a different accuracy rate, differing from García Mayo and Liu and Gleason. A result we find noteworthy is in regards to L1 transfer to L2. Ionin et al. (2008: 560) suggested “L2 learners whose L1 has articles transfer article semantics from their L1 to their L2”. On the one hand, the data support L1 transfer to L2 with the situational and cultural categories, but the two other categories, textual and structural, cannot be explained by L1 transfer.

The performance in the categories that caused the least and most difficulty (situational and cultural, respectively) can be explained by the L2 learners’ rules that are transferred from the L1. The most difficult category (cultural) does not share similar rules to English structure and the L2 learners restructure their interlanguage violating grammatical principals. The least difficult category (situational) does share similar rules to English structure and therefore a transfer of L1 rules to L2 grammar occurred. The data from these two categories support the notion that article semantics are transferred from the L1 to the L2. However, the remaining categories, textual and structural, cannot be explained by L1 transfer since both languages share the same rules in these categories. An example is “We went to a basketball game on Saturday. The players at the game were all very tall”. *Players* is not a(n) abstract concept, class or group in general and therefore it is obligatory to insert *the* before *players* in both the English and Spanish versions. The fact this rule exists in both the L1 and L2 and results in the second most difficult to acquire cannot be explained by transfer from L1, rather by restructuring in the interlanguage. The data from this study support a dual access hypothesis where the L2 learners have access to Universal Grammar but this may be partly blocked by the use of general thinking strategies.

García Mayo (2008) and Liu and Gleason (2002) concluded the structural use of *the* was acquired more readily than textual whereas our conclusions show the reverse. We suggest two possible factors. One may be due to the English language variety that is taught *in situ*, British English or American English. There are situations in which the definite article is used differently in several contexts in these two varieties. For example several non-generic nouns that refer to places take no definite article in both varieties when a role is implied: *at sea* (as a sailor), *in prison* (as a convict), and *at/in college* (for students). Among this group, British English has *in hospital* (as a patient) and *at university* (as a student), where American English requires *in **the** hospital* and *at **the** university* (though American English does allow *at college* and *in school*). When the implied roles of patient or student do not apply, the definite article is used in both dialects. A second possible factor may be due to the Spanish variety that is spoken. Southern Cone and Peninsular Spanish have slightly different rules that may explain the difference found in our study and that of García Mayo. Further investigation will have to be carried out to posit an additional factor to account for this performance

Overall, time spent studying the English non-generic definite article does increase accuracy rates across all categories, concurring with the results found by García Mayo (2008) and Ionin *et al.* (2008).

5. Bibliographic references

- BICKERTON, Derek, 1981: *Roots of language*. Michigan: Karoma Press.
- CARLUCCI, Lorenzo *et al.*, 2006: "Variations on U-shaped learning", *Information and Computation* 204 (8), 1264-1294.
- DIEZ-BEDMAR, María Belén & Szilvia PAPP, 2008: "The use of the English article system by Chinese and Spanish learners", in Gaëtanelle GILQUIN, Szilvia PAPP & María Belén DIEZ-BEDMAR (eds.): *Linking up contrastive and learner corpus research*, New York: Editions Rodopi: 147-175.
- DONATO, Richard, 1994: "Collective scaffolding in second language learning", in James LANTOLF & Gabriela APPEL (eds.): *Vygotskian approaches to second language research*, New Jersey: Ablex: 33-56.
- DULAY, Heidi, Marina BURT & Stephen KRASHEN, 1982: *Language two*. New York: Oxford University Press.
- GARCÍA MAYO, María del Pilar, 2008: "The acquisition of four nongeneric uses of the article *the* by Spanish EFL learners", *System* 36, 550-565.

- GASS, Susan & Evangeline VARONIS, 1994: "Input, interaction, and second language production", *Studies in Second Language Acquisition* 16, 283-302.
- GRANNIS, Oliver C., 1972: "The definite article conspiracy in English", *Language Learning* 22 (2), 275-289.
- HAWKINS, John A., 1978: *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. New Jersey: Croom Helm.
- HAWKINS, Roger *et al.*, 2006: "Accounting for English article interpretation by L2 speakers", in Susan H. FOSTER-COHEN, Marta MEDVED KRAJNOVIC & Jelena MIHALJEVIC DJIGUNOVIC (eds.): *EUROSLA yearbook*, Holland: John Benjamins: 7-25.
- HUEBNER, Thom, 1985: "System and variability in interlanguage syntax", *Language Learning* 25, 141-163.
- IONIN, Tania, Maria Luisa ZUBIZARRETA & Salvador Bautista MALDONADO, 2008: "Sources of linguistic knowledge in the second language acquisition of English articles", *Lingua* 118, 554-576.
- IONIN, Tania, Heejeong Ko & Ken WEXLER, 2004: "Article semantics in L2 acquisition: The role of specificity", *Language Acquisition* 12, 3-69.
- , "La peste negra" (n. d.): *Planeta Sedna*. [<http://www.portalplanetasedna.com.ar/malas01.htm>, date of retrieval: April 26th, 2011].
- LANTOLF, James P. & Gabriela APPEL, 1994: "Theoretical framework: An introduction to Vygotskian perspectives on second language research", in James LANTOLF & Gabriela APPEL (eds.): *Vygotskian approaches to second language research*, New Jersey: Ablex: 1-32.
- LIGHTBOWN, Patsy, 1985: "Great expectations: Second language acquisition research and classroom teaching", *Applied Linguistics* 6, 173-189.
- LIU, Dilin & Johanna L. GLEASON, 2002: "Acquisition of the article the by nonnative speakers of English: An analysis of four nongeneric uses", *Studies in Second Language Acquisition* 25, 1-25.
- MASTER, Peter, 1994: "Effects of instruction on learning the English article system", in Terence ODLIN (ed.): *Perspectives on pedagogical grammar*, New York: Cambridge University Press: 229-252.
- , 1995: "Consciousness raising and article pedagogy", in Diane D. BELCHER & George BRAINE (eds.): *Academic writing in second language: Essays on research and pedagogy*, New Jersey: Ablex: 183-204.
- , 2002: "Information structure and article pedagogy", *System* 30 (3), 331-348.
- TRENNIC, Danijela, 2007: "Variability in second language article production: Beyond the representational deficit vs. processing constraint debate", *Second Language Research* 23 (3), 289-329.

