IVANA ČIZMIĆ*
University Centre for Professional Studies
University of Split, Croatia

■ THE USE OF ARTICLES IN SPOKEN AND WRITTEN DISCOURSE

1. INTRODUCTION

Although it is well known that many second language (L2) learners of English have persistent trouble using articles properly until very late stages of acquisition, or do not ever reach native-like levels of performance, the primary causes of their difficulties still remain unclear. This study addresses the problem by examining the knowledge of the English article system that learners employ when selecting articles in a given situation.

Articles such as *a(n)* and *the* are the most commonly used words in English, and yet their usage is in fact surprisingly complex. Part of the complexity can be attributed to the fact that the English article system does not consist of one-to-one form and meaning relationships (Butler 2002). This complexity poses a number of challenges for L2 learners of English. This is true even for those learners who have studied English for a number of years. Teachers of English as a second language find it difficult to understand how and why their students choose to use articles in the ways that they do. Therefore, effectively teaching the article system to their students often remains an elusive goal.

Driven by my students' difficulties in acquiring the English article system, I decided to do a small-scale research on their use of articles in job interviews. I presume that students will make many mistakes using them. The reason for that might lie in the fact that there is no such category in the Croatian language which is students' native language (L1). Therefore, articles are expressed by using other categories such as demonstratives, numerals and pronouns.

This analysis will be corpus-based which is well suited to investigations of language use, which must be empirical, analyzing the actual distributional patterns in natural texts. According to Biber and Conrad (1999), corpora can also be used to compare grammatical structures by studying the ways in which seemingly similar structures occur in different contexts and serve different functions.

The primary aim of the present study is to investigate the differences in the use of the articles by non-native students of English when being interviewed for a job (spoken discourse) and native speakers of English (journalists) in daily newspaper articles (written discourse).

---

*   Kontakt podaci (Email): icizmic@oss.unist.hr

2. THEORETICAL FRAMEWORK

There has been a body of research on the article acquisition and use by learners of English as a foreign language (Huebner 1985; Master 1987; Thomas 1989; Chaudron/ Parker 1990; Young 1996; Robertson 2000; Butler 2002; White 2003). The difficulties that L2 learners have in understanding articles are surprising if compared with the experience of children who are acquiring English as their L1. Children seem to acquire the article system at a relatively early age in L1 acquisition (somewhere between 2;8 and 3;8 years old), and they typically exhibit a low frequency of overall errors (Butler 2002). Huebner (1983, 1985) found in a longitudinal study that his adult L2 learner initially overused the definite article *the* with almost all nouns in his speech. Flooding of *the* was also observed in learners with low English proficiency from other L1 backgrounds such as Japanese (Chaudron/Parker 1990) and Czech and Slovak (Young 1996). The same results were obtained in the present study. Master (1987) also found that *the* was overused. He also reported that among learners whose L1 did not have an article system (such as Croatian), the appropriate use of *a* was delayed when compared with *the*. Based on these results, both Huebner and Master suggested that L2 learners initially might associate *the* with the feature of hearer's knowledge in contrast to children learning the English article system as part of their L1.

Thomas (1989), however, analyzed L2 learners' article usage in their production and presented a different claim from those of Huebner and Master. Thomas's findings included the following three observations: (a) whereas L1 children show accurate use of *a* at an early stage, the accurate use of *a* by adult L2 learners was delayed; (b) the most common errors by L2 learners across proficiency levels were overgeneralized zero articles; and (c) both L1 child acquirers and L2 learners overgeneralized *the* in first-mention contexts.

Zdorenko and Paradis (2008) did a survey on the acquisition of articles in child second language English with a longitudinal corpus of narratives from English second language children with first languages that do not have definite / indefinite articles (Chinese, Korean and Japanese) and L1s that do have article systems (Spanish, Romanian and Arabic). It was hypothesized that L1 transfer most likely plays a role in L2 learners' acquisition of English articles. In the course of acquisition, L2 learners of English have been documented to omit articles in both definite and indefinite contexts, and to misuse them, that is substituting one in the context of another (Huebner 1985; Master 1987; Thomas 1989), which is also true for Croatian learners of English. Zdorenko and Paradis concluded that both adult and child learners' acquisition patterns were influenced by the greater complexity of the indefinite article in English, and omission errors were most commonly produced by learners whose L1s lacked articles.

As far as omission errors are concerned, they also have been documented in numerous studies of English article acquisition. For instance, Huebner (1985) and Robertson (2000) found a lot of use of the zero article in oral language. It has also been widely reported that L2 English learners are significantly more accurate in choosing the definite article in definite contexts than they are in choosing the indefinite article in indefinite contexts.

3. RESEARCH QUESTIONS

In this study I will try to answer the following research questions:

a. To what extent do students of English as a foreign language use the definite and indefinite articles in the spoken register, i.e. job interviews?
b. What could be the reasons for their incorrect use of articles in case they make mistakes?
c. How can the frequencies of occurrence of the indefinite and the definite article be compared across two corpora?
d. Are articles more often used in spoken or written register?

4. PARTICIPANTS

The participants used in the present study were 20-year-old second year full-time students of Information Technology from the University Centre for Professional Studies in Split, Croatia. Within this institution, English is an obligatory subject in the first as well as in the second year of study. Students have to complete a general English course being a part of the first year curriculum, whereas they learn English for Specific Purposes (ESP) in the second year. Although they have been learning English for about ten years, they are at pre-intermediate level of English due to the fact that for most of them English was only a marginal subject during their secondary education. Therefore, their knowledge of English is rather poor.

5. METHODOLOGY

The data gathered for this study were both qualitative and quantitative in nature. The present study is based on an analysis of two corpora containing data from a spoken and a written register. The spoken corpus which is rather small was compiled of 5 students' job interviews containing 3,065 words while the written corpus was compiled of 49 newspaper articles taken from *The San Diego Union-Tribune*, an American West coast daily newspaper containing 52,893 words. All articles within this corpus were issued from 1st January 2004 till 2nd April 2004 in the News section. The written corpus was drawn from the materials prepared by Dr. Enikõ Csomay for her PhD-level course *Discourse Analysis: A Corpus Linguistic Perspective* held at the University of Pecs in August 2007. Since this corpus was compiled for the classroom-use only, it must be emphasized that it does not meet the criteria stated by Biber (1993) in his article on the representativeness in corpus design. Therefore, the conclusions of this research must be discussed bearing in mind the limitations of the corpus used in the analysis. The quantitative analysis of articles was undertaken with corpus analysis toolkit *Antconc 3.2.1.* for Windows (2007), while its concordance feature was used to capture all of the instances of the indefinite and definite articles in both corpora.

As far as the spoken register is concerned, students had to role-play a job interview. One student was an employer and the other one a candidate. The candidate

had to submit his/her CV and a letter of application for one of the given vacancies that seemed the most interesting for him/her. The employer asked the candidate questions concerning his education and previous work experience. All interviews were tape-recorded and then transcribed for further usage based on pre-defined transcribing conventions.

6. FINDINGS

The first issue that was interesting to see was the actual use of indefinite articles across two corpora. I wanted to see whether those articles are more often used by students or journalists or, in other words, in spoken or written register.

	Job interviews	Newspaper articles
Number of interviews / articles	5	49
Number of words	3 065	52 893
Occurrences	393	1 275
Percentage	12.82	2.41

Table 1. Use of the indefinite article *a* in spoken and written corpus

Results indicate that students seem to use the indefinite article *a* more often than journalists. There is probably a relationship between the frequency of articles and of nouns, according to which students use more nouns in their role-plays than journalists in their newspaper articles. It can also be concluded that this indefinite article is more used in spoken than in written corpus. The relatively high distribution of the indefinite article *a* in students' interviews might be attributed to two factors. Firstly, there is a high frequency of singular nouns in relation to plural nouns in conversation resulting in more indefinite articles than otherwise expected. Secondly, pronouns are preferred over noun phrases as anaphoric expressions in conversation resulting in fewer definite articles.

The findings from this study differ from those presented by Biber *et al.* (1999) in *Longman Grammar of Spoken and Written English* (LGSWE, henceforth) that reports the results of corpus-based analyses based on approximately 20 million words from four registers: conversation, fiction, newspaper language and academic prose.

According to Biber *et al.*, "the distribution of the indefinite article is relatively similar across spoken and written registers" (1999: 267). One of the reasons for different results could be the discrepancy in the corpus size. Therefore, we cannot be sure whether the results from a smaller corpus are representative or not. Another reason could be the insufficient students' knowledge of the use of articles. This could be seen in the following examples in which students make mistakes concerning the use of *a*:

Where do you see yourself in a nearby future?
I have a bills to pay.
What about working time? – It's a flexible....

The second question is to what extent students and journalists use another indefinite article *an*, and whether the results will be the same as in LGSWE.

	Job interviews	Newspaper articles
Number of interviews / articles	5	49
Number of words	3 065	52 893
Occurrences	4	198
Percentage	0.13	3.35

Table 2. Use of the indefinite article *an* in spoken and written corpus

As the results show, *an* is more used in written than in spoken corpus which is almost the same in LGSWE. According to their findings, "the form *an* of the indefinite article is much less frequent than *a*, although it is slightly more common in news", meaning newspapers (Biber *et al.* 1999: 267). In addition, conversation sometimes makes use of the demonstrative determiners *this/these* in introducing new entities. The higher frequency of *an* in newspapers reflects the vocabulary of the register, where *an* is required more commonly with Latinate vocabulary (which has many words beginning with *ab-, ad-, ex-, in-, ob-*, etc.)

The definite article *the* will probably be widely used across the corpora because it occurs with countable as well as uncountable nouns.

	Job interviews	Newspaper articles
Number of interviews / articles	5	49
Number of words	3 065	52 893
Occurrences	51	3 072
Percentage	1.66	5.8

Table 3. Use of the definite article *the* in spoken and written corpus

The overall results show that articles are generally more common in written than in spoken corpus. The definite article is almost twice as common as the indefinite article in the written register because it has a greater range of uses. Firstly, it combines with both countable and uncountable nouns as well as both singular and plural nouns. In addition, the definite article is used commonly for subsequent mention, and when used

cataphorically, it can also introduce new referents. In contrast, indefinite articles are used primarily to introduce a new referent.

As for students' mistakes in using the definite article *the*, here are some examples:

I currently work at the computer company. (mentioned for the first time)
Are you ready to work under the pressure?

Taking into account students' examples of the use or omission of definite as well as indefinite articles, it can be concluded that there are quite a number of cases where they actually omit articles instead of using them. The reason for that may lie in their lack of knowledge of grammar rules on one hand, and in their insufficient reading in general on the other hand. Reading magazines, newspapers and books written in the target language, in this case in English, could really help students to better understand English syntax and to get an idea of the real use of articles in authentic materials.

Here are some examples of students' omission of articles where they had to use them:

That's good answer. (That's a good answer)
Where do you see yourself in future? (in the future)
Do you have driving license? (a driving license)
*Where do you see yourself in 10 years? – *Married with 2 children, good job, good paycheck, my own car and house. (a good job, a good paycheck and a house)*

7. DISCUSSION

As it could be seen from students' sentences taken from the concordance list, it is apparent that they make mistakes in using all articles. The question is why Croatian students, as well as other non-native learners of English, find using articles so difficult. Firstly, one of the reasons may lie in the lack of such category in the Croatian grammar in which demonstratives, pronouns and mostly numerals play the role of articles. Therefore, students have never got acquainted with the notion of articles in their own grammar, and accordingly find them difficult to use. This statement could be connected to Lado and his theoretical foundations for the Contrastive Analysis Hypothesis that were formulated in his *Linguistics Across Cultures* (Lado 1957, available at http://en.wikipedia. org/wiki/Contrastive_analysis). In this book Lado claimed that "those elements which are similar to (the learner's) native language will be simple for him, and those elements that are different will be difficult". In other words, positive transfer occurs when we are trying to learn the elements from a second/foreign language which are present in the mother tongue and we are familiar with them, whereas negative transfer occurs when learning those elements which are different from our L1. While this was not a novel suggestion, Lado was the first to provide a comprehensive theoretical treatment and to suggest a systematic set of technical procedures for the contrastive study of languages. This involved describing the languages (using structural linguistics), comparing them and predicting learning difficulties. The difficulty in mastering certain structures in a

second language depended on the difference between the learners' mother language (L1) and the language they were trying to learn.

In its strongest formulation, the Contrastive Analysis Hypothesis claimed that all the errors made in learning the L2 could be attributed to *interference* by the L1. However, this claim could not be sustained by empirical evidence that was accumulated in the mid- and late 1970s. It was soon pointed out that many errors predicted by Contrastive Analysis were simply not made in learners' language. Even more confusingly, some uniform errors were made by learners irrespective of their L1. It thus became clear that Contrastive Analysis could not predict learning difficulties, and was only useful in the retrospective explanation of errors. These developments, along with the decline of the behaviourist and structuralist paradigms considerably weakened the appeal of Contrastive Analysis.

Secondly, students' problems with articles may arise from inadequate and insufficient presentation of this grammatical category in textbooks and grammar books they are using in the English course. To illustrate bad presentation of articles in a textbook, I've chosen *Cutting Edge Pre-intermediate* (Cunningham/Moor 2001) that was used in the first year of study by low-level students within their general English course. In my opinion, the authors of this book gave only general examples of using articles that are very often decontextualized. In addition, in this book there are not too many exceptions to the rule that often occur in English.

Speaking of grammar books, it should be noted that authors and teachers, when engaged in material development for the purpose of language instruction and assessment, should make decisions about the language actually used in everyday situations and about linguistic features and specific words that should be included when illustrating a grammatical feature. Recently, in quite a number of grammar books authors prefer using authentic materials for presenting language from natural texts rather than made-up examples (Biber and Reppen 2002). They also confront some basic issues such as: which grammatical features should be included in a lesson or book, which should be excluded, how much space should be given to included features, what should be the order of grammatical topics. All these questions should be discussed in details before writing a grammar book in order to provide an accurate reflection of the language actually used by speakers and writers in natural situations.

When talking about grammar presentation in the classroom, I can say from my experience that most students prefer using pictures, images and spatial understanding, not only for teaching grammar but for teaching new vocabulary as well. Therefore, they like employing visual learning style. Since grammar rules that should be learnt could be sometimes boring for some students, it is up to teachers to present them in an interesting and amusing way. Materials from the Internet are widely used in classrooms, partly because most students are acquainted with the possibilities of the Internet, and partly because they find those materials more interesting than those in the text- or grammar books.

In discussing article acquisition by L2 learners, one has to keep in mind two separate issues. Firstly, the frequency of each error type differs depending on the task performed (Tarone/Parish 1988). In general, production tasks, such as interviews and essay writing, have produced lower error rates than objective tasks, such as

cloze tests. Researchers have attributed the lower error rates in production tasks to learners' avoidance of uncertain uses of articles in these tasks. Thus, it appears that collecting data from a variety of tasks, in addition to production tasks such as oral interviews and essay writing, is important for examining different types of article use by L2 learners.

8. CONCLUSION

I want to emphasize once more that all results drawn from this study should be subjected to careful verification and re-examination, as the corpora used for the analysis were compiled for classroom-use only and therefore they were not representative according to principles of corpus design (Biber 1993). The results of the present analysis suggest that articles are generally more common in written than in spoken corpus. The definite article is almost twice as common as the indefinite article in the written register because it has a greater range of uses. It can be combined with both countable and uncountable nouns, as well as both singular and plural nouns. Moreover, the definite article is used commonly for subsequent mention, and when used cataphorically, it can also introduce new referents. In contrast, indefinite articles are used primarily to introduce a new entity.

It should be noted that the present study has some limitations. Data gathered for the spoken corpus are not sufficient in order to be representative, and the comparison was made between the spoken and the written corpus although there have already been some similar examples as in Biber *et al.* (2004) where they contrasted two different registers – classroom teaching and textbooks.

Despite such limitations, it is hoped that the information gained from this study may provide teachers with a new means of identifying and understanding learners' problems, and indicate areas where language instruction can be made more effective. This study indicates that learners' errors stem from a number of sources. Unfortunately, articles do not appear to have been the focus of very much attention in many L2 classrooms to date. However, the results of this study indicate the importance of considering a more suitable way of teaching the article system in order to best help students become proficient in the use of articles.

REFERENCES

Biber, D. 1993. Representativeness in Corpus Design. *Literary and Linguistic Computing* 8(4), 243-257.

Biber, D. and S. Conrad. 1999. Lexical Bundles in Conversation and Academic Prose. In H. Hasselgard and S. Oksefjell (eds.) *Out of Corpora: Studies in Honour of Stig Johansson*. Amsterdam/Atlanta, GA: Rodopi.

Biber, D. *et al.* 1999. *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education Limited.

Biber, D. and R. Reppen. 2002. What does Frequency Have to Do with Grammar Teaching? *Studies in Second Language Acquisition* 24, 199-208.

Biber, D. *et al.* 2004. If you look at...: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics* 25/3, 371-405.

Butler, Y.G. 2002. Second Language Learners' Theories on the Use of English Articles: An Analysis of the Metalinguistic Knowledge Used by Japanese Students in Acquiring the English Article System. *Studies in Second Language Acquisition* 24 (3), 451-480.

Chaudron, C. and K. Parker. 1990. Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition* 12, 43-64.

Cunningham, S. and P. Moor. 2001. *Cutting Edge Pre-Intermediate, Students' Book*. Harlow, Essex: Pearson Education Limited.

Huebner, T. 1983. *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Huebner, T. 1985. System and variability in interlanguage syntax. *Language Learning* 35, 141-163.

Master, P. 1987. *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, University of California, Los Angeles.

Master, P. 1997. The English article system: Acquisition, function, and pedagogy. *System* 25, 215-232.

Robertson, D. 2000. Variability in the use of the English article system by Chinese learners of English. *Second Language Research* 16, 135-172.

Tarone, E. and B. Parrish. 1988. Task-related variation in interlanguage: The case of articles. *Language Learning* 35, 373-395.

Thomas, M. 1989. The acquisition of English articles by first- and second-language learners of English. *Applied Psycholinguistics* 10, 335-355.

White, L. 2003. Fossilization in steady state L2 grammars: implications of persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6, 128-141.

Young, R. 1996. Form-function relations in articles in English interlanguage. In R. Bayley and D. R.Preston (eds). *Second language acquisition and linguistic variation*. Amsterdam: Benjamins, 135-175.

Zdorenko, T. and J. Paradis. 2008. The acquisition of articles in child second language English: fluctuation, transfer or both? *Second Language Research* 24 (2), 227-250.

Wikipedia. Entry on Contrastive Analysis. http://en.wikipedia.org/wiki/Contrastive_analysis.

SUMMARY

THE USE OF ARTICLES IN SPOKEN AND WRITTEN DISCOURSE

This paper deals with a survey describing the students' use of articles in a spoken register that was compared to the use of articles in a written register. The results showed that articles are generally more common in written than in spoken corpus.

The definite article is almost twice as common as the indefinite article in the written register because it has a greater range of uses. Since in the current language teaching practice and theory the emphasis is put on the communicative language competence, it is important to decide how to best help students become proficient in the use of articles. Due to the small size of the corpus, the conclusions of the present study should be subjected to further re-examination and verification.

KEYWORDS: students' usage of definite and indefinite articles, spoken and written register, grammar presentation.