THE USE OF ARTICLES IN ACADEMIC WRITING

Renata Šimůnková

Technical University of Liberec
Faculty of Sciences, Humanities and Education
Studentská 2, 461 17, Liberec 1, Czech Republic
renata.simunkova@tul.cz

Abstract

Writing academic texts in English is an inevitable part of work of all academics and researchers. It is, however, often the case that scientifically excellent contributions are not published because of their lack of accuracy. Moreover, mistakes seem to be universal, articles being among the most frequent ones. This paper pursues the following aims: 1) systematically describe and categorize the use of articles in academic texts, 2) determine the most common mistakes in the use of articles made by L1 Czech writers, and 3) delimitate the problem areas and possible remedial steps. Four thousand and forty eight articles have been excerpted from six scientific texts in American English, classified and analyzed. Mistakes have been studied on scientific articles written by Ph.D. students with the conclusion that they can generally be classified into three areas. These areas should then logically represent the focus of instruction.

Introduction

Academic writing is a distinct register which can be distinguished from other registers by a number of typical features. One such feature is a high number of noun phrases. According to Biber and Gray, "academic writing is structurally "compressed", with phrasal (non-clausal) modifiers embedded in noun phrases" (Biber and Gray 2010: 12). In a different article they claim that "in particular, the distinctive communicative characteristics of academic writing (informational prose) have led to the development of a discourse style that relies heavily on nominal structures, with extensive phrasal modification and a relative absence of verbs" (Biber and Gray 2009: 128). Nominalizations are useful in academic writing because they convey an objective, impersonal tone. Nominalizations can also make the text more concise because they can pack a great deal of information in a few words. Noun phrases refer to the linguistic or situational context. The kind of reference of a particular noun phrase is related to context and expressed by means of determiners which occur in front of a head noun or its pre-modifiers. The most frequent central determiners are articles.

The high frequency of noun phrases logically results in a high number of articles in academic texts, and thus a demand on writers to be able to use them. The correct use of articles however represents one of the biggest challenges for learners of English as a second language. It is even more significant in the case of L1 Czech learners since Czech does not have articles, and learners therefore cannot draw on their experience and beneficial aspect of L1 transfer when acquiring the use of articles.

The present study targets at the following three aims:

- to study and classify the use of articles in English academic texts;
- to study and describe the most common mistakes Czech writers of academic texts make;
- to suggest simple and straightforward strategies that could help Czech academics and scientists to use articles correctly.

1 Articles in academic texts

In order to be able to use articles correctly, one first needs to know what the standard is. In other words, it is crucial first to see in what way native academic and scientific writers use articles. Therefore a deep insight into this matter has been conducted.

1.1 Methodology

In phase one of the study, a quantitative analysis has been carried out. For this purpose six different scientific articles written by native English authors have been selected and four thousand and forty eight articles (including zero article) have been excerpted. It proved rather complicated to find texts clearly written by native speakers since in the world of science the majority of research is team work, and moreover the teams are often multinational. Therefore the background of each author has been carefully checked and also authors having lived in the US for more than ten years have been considered native speakers. Since it has been assumed that there were going to be substantial differences in the use of articles with regard to the part of a text, only texts with a structure typical for scientific writing, which means texts consisting of an abstract, introduction, experimental part and discussion of results/conclusion, were included.

Once excerpted, the articles have been classified into categories according to reference they expressed. Frequencies with which the individual articles and the individual reference categories appeared in the texts as a whole and in their respective parts have been counted, expressed in percentages, and illustrated by means of graphs and tables.

1.2 Classification of reference

For the purpose of this article the classification of reference presented by Quirk(1985: 265 – 287) was used as a starting point which has been adapted and extended in order to better reflect the goals of the present study. The excerpted articles were thus classified into the following categories:

The definite article the:

- anaphoric reference
- cataphoric reference – “of phrase”
- cataphoric reference – other cases
- theory
- quantity
- other terminology (apart form quantity)
- situational
- generic

The indefinite article a/an:

- generic
- specific
- non-referring
- theory
- quantity
- other terminology (apart form quantity)

The zero article:

- generic
- specific
- non-referring
- theory
- quantity
- other terminology (apart form quantity)
- definite

Some of the categories above are going to be described in bigger detail, the rest strictly follows from Quirk (1985: 265 – 287). Generic reference represents reference to all members/representatives of a certain noun or to a typical representative of a noun. Definite cataphoric reference has been split into two categories, one including only cases of the postmodification by an "of phrase", and the second category all the other cases. Non-referring function of articles is according to Quirk (1985: 273) connected with "noun phrases in copular relationship. Here it has a descriptive role, rather than a referring role". The category called "theory" refers to the names of theories, such as the Theory of Relativity. In a similar way the category called "quantity" refers to the names of quantities, such as electric current, force, velocity and the category titled "other terminology" to all other words which could be considered specific for a certain field and therefore considered terms, such as quark, hyperbola. The last three categories have been added to Quirk's treatment of reference since they are almost exclusively scientific/academic writing specific and they are expected to be rather frequent in the studied texts.

1.3 Presentation and discussion of the results

Although there were some differences in the use of articles between the individual texts, we believe that it is possible to observe certain regularities in article distribution throughout a scientific article. Five out of the six pieces of text examined showed comparable results and the conclusions below are based on average values gained from the five articles. The sixth text provided values substantially different from the other ones and was not included in the final results. A qualitative analysis will be carried out in the future to try to explain the causes of the difference. The average results are as follows:

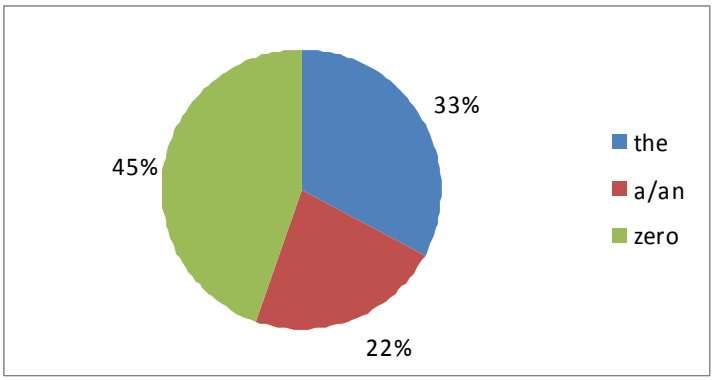

Fig. 1: The distribution of articles

Fig. 1 illustrates the general distribution of the individual articles in the texts, expressed in percent.

Tab. 1: Articles in the individual parts

article/part of the text	the	a	zero
Abstract	31	18	51
Introduction	18	30	52
Experiment	49	12	39
Conclusion	19	32	49

Tab. 1 shows the distribution of individual articles in the four basic parts of scientific texts: abstract, introduction, experiment and conclusion. The values are in percent and they mean that, for example, in the abstracts the majority of nouns were used with the zero article (61 %), the second most frequent was the definite article (30 %), and the least frequent was the indefinite article (9 %). The following data illustrate the distribution of the set kinds of reference for each article in the individual parts of a scientific text. The data for the abstracts, for example, can be interpreted in the following way: out of the nouns used with the zero article (which represented 51 % of all the articles used in the abstracts – see *Tab. 1*) 26 % expressed generic reference, 14 % specific indefinite reference and 40 % were terminology expressions. Regarding the distribution of reference within the nouns used with the definite article (in the abstract there were 30 % of these nouns), 21 % represented definite anaphoric reference, 29 % definite cataphoric reference with of phrase, 7 % were other cataphoric cases and 43 % were examples of terminology. As far as the indefinite article is concerned, it was used with generic reference in 25 % of its occurrence, 25 % with specific indefinite reference, 25 % in non-refering function and 25 % represented its use with terms.

ABSTRACT					
	the		**a**		**zero**
anapho	21	gener	25	gener	26
catapho of	29	specif	25	spec	14
catapho other	7	non ref	25	non-ref	
theory		theory		theory	
quantity		quant		quant	
other term	43	other term	25	other term	60
situational				def ref	

INTRODUCTION					
	the		**a**		**zero**
anapho	14	gener	16	gener	22
catapho of	27	specif	52	spec	35
catapho other	5	non ref	8	non-ref	1
theory		theory		theory	
quantity		quant		Quant	
other term	36	other term	24	other term	42
situational	18			def ref	

EXPERIMENT					
	the		**a**		**zero**
anapho	41	gener	8	gener	16
catapho of	15	specif	80	spec	42
catapho other	9	non ref	25	non-ref	
theory		theory		theory	
quantity		quant		quant	1
other term	17	other term	12	other term	41
situational	18			def ref	

CONCLUSION					
	the		**a**		**zero**
anapho	37	gener	60	gener	26
catapho of	21	specif	40	spec	22
catapho other	3	non ref		non-ref	
theory		theory		theory	
quantity		quant		quant	
other term	17	other term		other term	52
situational	22			def ref	

As can be expected from the general distribution, the zero article is the most numerous in the three out of the four typical parts of a scientific article. The only exception is the experimental part where the definite article prevails. In this part the definite article also most frequently (out of the four parts) expresses definite anaphoric reference. It is quite understandable since in this part a sequence of activities, processes, etc. is usually described where the same nouns or concepts are used repeatedly. The definite article is also frequent in the abstracts which often present the concrete outputs of the research. In absolute values it is however the most frequent with terminology. This unfortunately does not make the use of articles any easier because also the other two articles are used with terms. The high percentage of terminology is clearly connected with the nature of the texts, and it can at least be stated that when the occurrence of a certain article together with a term is studied, its proportion is the smallest in the case of the indefinite article.

Regarding the indefinite article, it is frequent in the introduction and conclusion. Nevertheless, it is worth noticing that in each case different kind of reference prevails. While in the introduction the indefinite article is the most frequent with specific reference, in the conclusion it is generic reference. The possible reason for the difference might reside in the fact that while in the introduction concrete and therefore specific circumstances of research are described, the concluding part attempts to overcome the boundaries of a particular study and present outcomes within more general context and with further reaching significance. A similar tendency, although not so significant, since the zero article unlike the indefinite article also plays a significant role in the area of terminology, can be observed in the case of zero article which is more frequent with specific reference in the introduction and experimental part and more frequent with generic reference in the abstract and conclusion.

2 Common mistakes of Czech authors of scientific texts

This part of the study represents just an initial brief insight into the problem. Based on the studies carried out on Finnish, Korean, Russian, and Japanese L1 speakers of English by researchers abroad (Haan 1997, Ionin 2007, Crompton 2011), areas of article use likely to cause the biggest difficulties were determined. The studies on respective L1 speakers were chosen since in all the cases the L1s do not use articles as all. In all the mentioned studies it has been concluded that learners of English whose L1 does not have articles use articles less often than native speakers: " The Finnish learners use fewer articles which may reflect the absence of articles in their mother tongue" (Haan 1997: 52). In all the studies two areas were delimitated where the mistakes in the use of articles with these learners seem to be most frequent. "It was found that while these learners made many errors of article (mis)use in English, their errors were not random. Rather, errors came in two types: overuse of *the* with specific indefinites and overuse of *a* with non-specific definites. Article use on specific definites and non-specific indefinites was accurate" (Ionin 2007: 563). These mistakes are supposed to be due to insufficient awareness of the notion of definiteness on the one hand, and specificity on the other. Snape (2005) believes that there is an Article Choice Parameter determining the distribution of articles, and claims that L2-learners possibly pick the wrong value, or fluctuate between definiteness and specificity when they are learning an L2 that has the features [+definite] and [+specific] (Snape 2005: 3).

The above mentioned conclusions were compared to the analysis of three scientific articles written by three different Ph.D. students at the Czech Technical University in Prague. The results of our brief study so far are generally consistent with the results discussed above. There were basically two visible differences. While in the studies mentioned the authors generally claim that the mistakes in the two main areas were of almost the same frequency, in our study the overuse of the definite article significantly prevailed. Moreover in our study one more area of frequent mistakes appeared and it was the overuse of *the* with generic reference. This was the second most frequent mistake. Similar findings were also reported in Crompton (2011: 78).

Conclusion

In the small scale research presented here we first categorized and analyzed the ways in which native speakers of American English use articles. The most frequently used article was the zero article followed by the definite article, and significantly the least frequent was the indefinite article. Certain differences can be observed between individual parts of a scientific text, the biggest difference being the frequency of the indefinite article in an introduction and conclusion on the one hand, and an abstract and experiment on the other, and the fact that in an introduction it prevails with specific reference while in conclusion with generic reference. The distribution of the zero article is to a certain extent similar, the differences however not being so significant. Although the zero article is. in general. the most frequently used, it is the definite article which prevails in an experimental part and in this part it moreover occurs most frequently with the definite anaphoric reference.

In the second part of the study the most common mistakes L1 Czech users of English make in scientific writing were studied, which resulted in the delimitation of three areas: overuse of the definite article with specific indefinites, overuse of the definite article with generic reference, and overuse of the indefinite article with non-specific definites.

As an intersection of the two studies we get the areas of potential problems. The first one is the potential overuse of the definite article in an introduction and conclusion where specific indefinite and generic reference can be expected to be frequent, and thus representing the

ground for the overuse of the definite article. Secondly it is important to draw attention of scientists writing in English to the difference between definiteness and specificity. It is necessary for them to realize that it is not uncommon that a noun can be non-specific but definite, and thus needs to be used not with the indefinite, but the definite article. This problem is most likely to be connected with an experimental part of a scientific text where from the context clearly identifiable but not necessarily specific entities, notions or events are described.

Literature

[1] BIBER, D.; GRAY, B.: *Challenging Stereotypes about Academic Writing: Complexity, Elaboration, Explicitness in Journal of English for Academic Purposes*. Vol. 9. 2-20. Orlando: Elsevier, 2010.

[2] BIBER, D.; GRAY, B.: *Grammatical change in the noun phrase: the influence of written language use in English Language linguistics*. Vol. 15, Issue 2. CUP 2009.

[3] CARTER, R.; MCCARTY, M.: *Cambridge Grammar of English*. CUP, 2006. ISBN 0-521-58846-4

[4] CROMPTON, P.: *Article Errors in the English Writing of Advanced L1 Arabic Learners: The Role of Transfer in Asian EFL Journal.* Vol 50. February 2011.

[5] HAAN, P.: *An experiment in English learner data analysis in Jan Arts (ed) Studies in English Language Research*. Amsterdam: Rodopi, 2011.

[6] IONIN, T.; MONTRUL, S.: *Article use and generic reference: parallels between L1 and L2 acquisition in M. P. García Mayo and R Hawkins (eds), Second Language Acquisition od Articles: Empirical Findings and Theoretical Implications*. John Benjamins, 2009.

[7] MASTER, P.: *The English article system: Acquisition, function, and pedagogy*. System 2 (2), 215 – 232. 1997.

[8] QUIRK at al.: *A Comprehensive Grammar of the English language*. Lomgman, 1985. ISBN 0-582-51734-6

[9] SMALLEY, L.S.; REUTTEN, M.K.: *Refining Composition Skills*. Boston: Heinle & Heinle, 2001.

[10] SNAPE, N. at al.. Spanish, Turkish, Japanese and Chinese L2. Learners' Acquisition of Generic Reference. *In Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 1-8. Somerville, MA: Cascadilla. 2009 Proceedings Project. www.lingref.com document #2251

Acknowledgement
This research has been supported by the Student Grant Competition of FP TUL, grant SGS 5848.

Mgr. Renata Šimůnková, Ph.D.