

The Role of the English Article System in Developing Dialogical Context: A View from Russian Science

Thomas A. Beavitt, Natalia G. Popova

Institute of Philosophy and Law of the Ural Branch of the Russian Academy of Sciences

Correspondence concerning this article should be addressed to Natalia G. Popova, Institute of Philosophy and Law, Ural Branch of the Russian Academy of Sciences, Sofia Kovalevskaya str., 16, Ekaterinburg, 620108, Russian Federation. E-mail: ngpopova@list.ru

A common issue arising in international scientific communication concerns the use of zero, indefinite- and definite articles in English. While existing pedagogical approaches are successfully used to teach near-native competency in the use of English articles, the final stage of native-equivalent competency continues to evade even advanced ESL / EAP students, especially those whose first language (e.g. Russian) does not possess articles. Therefore, the present work is aimed at developing a conceptual approach to article use based on a consideration of the development of context in scientific communication. A literature review of some existing theoretical approaches shows a clear development from semantic and transformational-grammatical attitudes towards more pragmatic explicatory strategies based on dialogic communication. A qualitative content analysis of article errors appearing in the text of research papers written by Russian scientists and corrected by a native English editor revealed the presence of nine major article-related error types, of which the use of zero article with singular NPs instead of the definite article (SxØvD) was the most common. NPs in sentences containing article errors were also evaluated according to Bickerton's NP semantic function typology, Hawkins' location theory and Liu and Gleason's major types. The hypothesis that Russian scientists, being highly educated, would generally tend to overuse *the*, *a* and *an* in an overcompensation for the tendency of less-well-educated Russians to drop articles when using English was not confirmed by the results of the content analysis. The analysis of article-use error types appearing in different sections of an IMRaD research paper showed that errors related to the overuse and underuse of the definite article *the* are particularly characteristic of the Introduction section. However, the largest number of article-related errors were found in the Discussion section, where the SxØvD error significantly outweighed other error types as compared with other sections. While existing conceptual approaches to explicating the function of the English article system have limited utility, a comprehensive system has yet to be developed. The authors recommend a dialogic approach for teaching the English article system to non-native speakers in the context of scientific communication.

Keywords: dialogic WE, definite article, contextual definiteness, noun phrases, Russian science, scientific communication, anaphor

Introduction

In the context of scientific communication, much attention has been focused on the structure of articles written for publication in academic journals. In particular, discussions have centred on the standardisation of such texts into the IMRaD¹ format, the identification of certain rhetorical "moves and steps" essential to the structure of styles of argumentation deemed "scientific" (Swales, 1990), as well as investigations into various linguistic factors (e.g. the appropriate use of different tenses in different sections (Hinkel, 2004), modality (Yang et al., 2015), pronouns (Taylor & Goodall, 2019), phraseology (Oakey, 2020), phrasal verbs (Alangari et al., 2020; Liu, 2012), hedges (Zanina, 2016), existential 'there' (Jiang & Hyland, 2020; Rasskazova et al., 2017), etc.) affecting the perceived publishability of scientific research papers in international journals, especially those indexed by the Web of Science and SCOPUS citation databases.

¹ Introduction, Methods, Results and Discussion

In previous research, the authors of the present article focused on sociological issues associated with the standardisation of English for Academic Purposes (EAP) in the context of globalisation, problematising the apparent linguistic imperialism inherent in the IMRaD format, CLIL teaching technologies and English academic writing centres (Popova & Beavitt, 2017), as well as developing an understanding of the various definitions of “science” used in the framework of contemporary sociological analysis (Popova et al., 2018).

In this paper, we present a more detailed analysis of a particular linguistic problem affecting communication in Russian science, where English is widely used as a second language (ESL). The need to regularly publish research papers in English presents a considerable challenge to Russian scientists (Kirchik et al., 2012). Despite receiving instruction from linguistic specialists embedded in their respective institutions, and/or attending writing centres, Russian scientists continue to struggle with some aspects of the English language, resulting in the often-unavoidable requirement to pay for expensive services provided by native English editors prior to submitting their work for publication.

In particular, Russian EAP users struggle with the English article system (Chrabaszcz & Jiang, 2014), whose misuse, according to Randal Whitman, “is one of the most evident grammatical signs that a person is not a native speaker of English”, thus comprising “one of the most formidable problems to overcome in teaching English grammar to foreigners” (Whitman, 1974, pp. 253–262). As Peter Master observes, while various teaching methodologies have been proposed in order to make this feature of English more accessible to the learner, “a few errors always remain and it seems in general that it is almost impossible for non-native speakers of English to arrive at the point where all article errors disappear, especially when their first languages do not contain an article system” (Master, 1995, pp. 183–184). At the same time, *the* ability to confidently deploy the most frequently-used word in the English language² comprises one of the most important stages in the transition to an advanced level of English competency; moreover, for writers who want to publish their work without relying on professional editing services, its mastery is non-optional.

Literature Review

In order to trace the development of understandings concerning the acquisition of competency in using the English article system, we present a review of the relevant literature.

Oliver Grannis notes that traditional accounts of definite article usage typically state it to follow “previous mention, generally in conjunction with the indefinite article” (Grannis, 1972, p. 279). However, he quotes the earlier work of Paul Christopherson to object that “there is a certain aversion to the use of a *the* form immediately after the word is introduced; a demonstrative is more usual in such cases” (Christophersen, 1939). Therefore, use of the definite article is “not a necessary consequence of previous mention” but rather “an option, subject to certain complicated stylistic constraints”. Discussing these constraints in more detail, Grannis uses the example of the difference in meaning between “a bee stung me on the nose” and “the bee stung me on the nose” to problematise so-called “familiarity theory”, i.e. the definite article is used to refer to that with which the speaker (or writer) is already familiar. Grannis shows that, in order to account for the semantic distinction between these examples, it is necessary to invoke a dialogue, since “the bee” is necessarily already familiar to *both* parties to the communication. Considering the distinction between determiners used in “have you seen the / my / a cat?” to show different degrees of familiarity, he observes that “the cat” is likely to be used with a family member, “my cat” with a neighbour, while “a cat” (followed by a description) is most appropriate in communication with a stranger (pp. 283–284). Grannis concludes that it is necessary to account for non-generic, productive uses of the definite article on the basis of “uniqueness within a given conversational situation”, which must be thought of “in relative rather than absolute terms”. Thus, definite article use involves a “complicated assessment of [the] listener’s actual knowledge and probable expectations” as well as an invitation to “share in an attitude of uniqueness in respect to the concept in question” (p. 286).

Tracing the origins of the definite article in modern languages, Alan Sommerstein refers to its emergence in Ancient Greek grammar, where its main function was to be “preposed to a great variety of adverbial expressions” in such a way that “an expression so prefixed could act as a noun phrase” (Sommerstein, 1972, p. 197). Sommerstein takes a transformational-grammatical (TG) approach to critique the earlier work of Paul Postal, in which English pronouns are considered as determiners whose following noun or noun phrase has been elided

² According to the Oxford English Corpus (OEC) and Corpus of Contemporary American English COCA

(Postal, 1969). While agreeing with Postal that such surface syntactic forms as pronouns should not be assumed to provide direct insight into corresponding deep syntactic forms, Sommerstein concludes that such an assumption would be just as misplaced with regard to the definite article.

Randal Whitman considers the article as performing a dual function of quantifier and determiner, “each of which is optional” (Whitman, 1974, pp. 253–262). For this reason, four distinct roles may be played by an article in a noun phrase (NP): (1) neither quantifier nor determiner; (2) quantifier alone; (3) determiner alone; and (4) both quantifier and determiner (p. 256). For this reason, the author recommends that the uses of articles should be taught in the following order: (1) quantity; (2) generic plural; (3) non-count nouns; (4) determiners; (5) generic articles (pp. 258–261).

In his book *The Roots of Language*, Derek Bickerton presents a study of the evolution of Hawaiian Creole English (HCE) via its antecedent form Hawaiian Pidgin English (HPE). Bickerton explains the transition from pidgin to creole according to the assumption that “speakers of different languages at first evolved some form of auxiliary contact-language, native to none of them (known as a pidgin), and that this language, suitably expanded, eventually became the native (or creole) language of the community which exists today” (Bickerton, 1981).

The evolution of creoles from English-based pidgins is of particular interest when considering article use. Bickerton demonstrates that, while in HPE articles only appear “sporadically and unpredictably”, in HCE “the definite article *da* is used for all and only specific-reference NPs that can be assumed known to the listener [...], the indefinite article *wan* is used for all and only specific-reference NPs that can be assumed unknown to the listener (typically first-mention use) [while] all other NPs have no article and no marker of plurality. This category includes generic NPs, NPs within the scope of negation – i.e., clearly nonspecific NPs – and cases where, while a specific referent may exist, the exact identity of that referent is either unknown to the speaker or irrelevant to the point at issue.” (pp. 22–23).

John Hawkins also takes a TG approach to explain the conversion of underlying NP indefiniteness structures to corresponding surface definiteness structures (Hawkins, 1980, pp. 41–42). He shows that the surface definite articles in the italicised NPs in (1)–(3)

- (1) The aphasic resents the fact that everyone speaks a language.
- (2) I dislike the name *Algernon*.
- (3) I recalled the sweet little child that Harry used to be.

can be derived from the underlying indefinite structures italicised in (4)–(6)

- (4) That everyone speaks a language is a fact (which the aphasic resents).
- (5) *Algernon is a name* (which I dislike).
- (6) I recalled that Harry used to be a sweet little child.

Examining whether the definite article in such constructions can be shown to be “a surface definite article only, derivable from an underlying indefinite”, Hawkins concludes that the conversion of the underlying indefinite to a surface definite is semantically motivated (pp. 42–63).

In an article published a decade later, the same author acknowledges that in the intervening period there has been a “growing understanding of the relationship between semantics and pragmatics”, such that “many aspects of natural language understanding that were previously thought to be part of the conventional meaning of a given expression can now be shown to be the result of conversational inference” (Hawkins, 1991). Remaining within the TG paradigm, Hawkins cites implicature theory, originally proposed by Paul Grice and developed by Stephen Levinson in terms of interaction between Q(uality)- and I(nformativeness)-implicatures, as “providing the right theoretical background for the kind of pragmatic-semantic-syntactic approach” initiated in his earlier work (p. 406). In order to discuss the property of uniqueness, which he shows, along with existence, to be “consistently required with definite descriptions” (p. 412), Hawkins designates a “general kind of knowledge of associative relationships between entities, such as the fact that any class at a university will have a (unique) professor, which sanctions a first-mention *the professor* (even *Who is the professor?*) when students arrive for a new class at the beginning of the semester. Similarly, following mention of *a class*, the speaker can

immediately talk of *the professor, the textbook, the final exam*.” Thus, members of a relevant linguistic community can predict that a particular class will have such unique members as “professor” and “textbook” based on the knowledge that such entities are typically included in the set that makes up such classes. Hawkins refers to such sets, used to define pragmatic parameters for the uniqueness of definite descriptions, as “P-sets” (pp. 408–409).

By the beginning of the 21st century, a more explicitly dialogical approach emerges in the work of Richard Epstein, who cites the accessibility theory of Ariel (1990) and the mental spaces theory of Fauconnier (1994) to argue that the function of *the* is to “signal the accessibility of a discourse referent” – or, more specifically, to denote “the availability of an access path through a configuration of mental spaces” (Epstein, 2002, pp. 341–342). Thus, for Epstein, the definite article serves as part of a set of “instructions for the construction of spaces, the introduction of elements into the spaces, the distribution of information over a given set of spaces and the accessibility of knowledge in a given space with respect to other spaces.” From this perspective, it can be seen that all discourses have a starting point referred to in terms of “origin space”, which, although “canonically identified with speaker reality”, may unfold to include alternate “base spaces” that represent the respective realities of an interlocutor or third person. In a given discourse, such a base space provides an initial “viewpoint space”, from which other spaces, such as the “focus space” to which structure is currently being added, can be accessed or created (p. 341).

Interpreting the role of the definite article according to the concept of accessibility³ turns out to have counterintuitive consequences. While one might expect the definite article to demarcate NPs referring to highly accessible referents, Epstein convincingly demonstrates that this is not so using the following example:

- a. There’s a cat in the yard. It’s eating a mouse.
- b. There’s a cat in the yard. #The cat is eating a mouse.

According to accessibility theory, although both *it* in (a) and *The cat* in (b) refer to the same uniquely identifiable entity introduced with the NP *a cat*, the unnaturalness of the definite article in (b) is due to the high accessibility of the referent. Therefore, in order to refer back to the cat with a NP in the subject position of the following sentence, it is necessary to use an anaphoric⁴ expression – e.g. a pronoun – that indicates higher accessibility: the reason why the definite article is inappropriate in (b) is because it would connote lower accessibility than its referent (p. 340).

Citing Ariel (Ariel, 1990, pp. 22–30), Epstein lists four factors affecting accessibility: recency of mention, saliency, competition and unity (p. 344). He shows that in the example of

- (c) There’s a cat and a dog in the yard. The cat is eating a mouse.

it is competition due to the presence of the dog that lowers the cat’s accessibility, making the definite article a more natural choice than a pronoun, which would increase ambiguity due to competition between the two referents (p. 345)

Situating the question in a dialogic context, Epstein considers what specific factors motivate the choice of a definite description in a given situation and what interpretations an addressee is motivated to construct on encountering such an expression (p. 347). He goes on to discuss the role of the definite article in designating the prominence of entities in a discourse (pp. 349–357), assigning roles and values (pp. 357–363), contributing to shifts in points of view with respect to spaces (pp. 363–367), as well as underspecification (pp. 367–371), concluding that the definite article “is essentially a discourse processing instruction signalling that the means for interpreting the NP in which it occurs is available somewhere in the configuration of mental spaces, as long as the appropriate spaces, elements and connections – i.e., access path – can be constructed by the addressee” (p. 371).

³ High accessibility corresponds to the immediacy with which items can be retrieved from short-term memory, while low accessibility characterises the retrieval of items from less immediate long-term memory

⁴ In this context, *anaphora* (Greek: ἀναφορά, “carrying back”) denotes structures used to refer to information previously introduced in the same text (or intertext).

Taking a dynamic, speaker-oriented approach to describing the role of the definite article (as opposed to the hearer-oriented paradigm described by Hawkins (Hawkins, 1978), in which use of *the* depends not only on the speaker's referential intent, but also – crucially – on the speaker's assessment of the hearer's knowledge of the referent), Epstein shows that “speakers do not simply choose articles in a relatively passive way, responding chiefly to what they think hearers know, [but] often choose *the* even when they know that the hearer is not yet able to pick out the referent in question”. Thus, the selection of articles is part of an “active, dynamic process of referent construction”, in which interlocutors are prompted to accept certain referents into the discourse. As such, while the estimation of an interlocutor's knowledge is certainly relevant to the choice of article, it remains among other factors to be taken into consideration (pp. 371–372).

In his comparison of article-related material in corpora and ESL grammars, Isaiah WonHo Yoo categorises uses of *the* into three main groups: referential, nonreferential and generic (WonHo Yoo, 2009). Referential forms are subdivided into anaphoric (i.e. second-mention), associative (anaphoric), situational, cataphoric (postmodification), unique reference and sporadic reference (pp. 268–269). Yoo also refers to the existence of two separate forms of zero article in English. Referring to his own earlier work (WonHo Yoo, 2008), as well as that of Andrew Chesterman (Chesterman, 2005) and Paul Christophersen (Christophersen, 1939), the author states that “the zero article, or \emptyset_1 , precedes mass nouns and plural count nouns, whereas the null article, or \emptyset_2 , precedes singular proper nouns and some singular count nouns” (p. 269). Examples of these quite distinct usages are provided as follows:

Athena loves \emptyset_1 milk and \emptyset_1 cookies.
Matthew was \emptyset_2 best man at my wedding.

Here, although zero article is used as expected with the uncountable mass noun “milk” and the plural count noun “cookies” to signify indefinite reference, the use of null article with “best man” connotes a greater degree of definiteness than if the definite article had been used. (Although the definite article could have been used here, it would tend to change the meaning of the proper noun or title “best man” to connote a superlative, i.e. although there were many good men at my wedding, Matthew was the best one).

In an effort to elucidate the English article system and pave the way for its successful acquisition by ESL learners, Russian linguists have developed a number of original concepts.

Taking a cognitive approach to investigating the article + noun structure in English, Alexander Smirnitsky suggests that the English article system serves to express the noun's logical-grammatical category, which is based on an opposition between abstraction and concretisation (Smirnitsky, 1959). Developing this idea, Yekaterina Dolgina analyses the choice of articles in terms of the human ability to perform categorisation procedures involving such cognitive operations as abstraction, classification and individualisation (Dolgina, 2008). According to Elena Kubryakova and Olga Iriskhanova, abstraction, which is based on the generalisation and dichotomisation of various categories (man and woman, writer and reader, good and evil), is generally indicated by a zero article (Kubryakova & Iriskhanova, 2007). Classification, which involves the identification of features typical of a certain class of notions or subjects, is marked by a noun with an indefinite article (*this chapter is an attempt to...*), while to denote cognitive processes of individualisation, which implies a narrowing of the meaning of a noun by distinguishing it among other subjects in a group, a definite article tends to be used.

Another productive research direction was to treat English articles as means of theme / rheme actualisation. Advancing the idea that they might play a “diagnostic informational” role, Boris Ilyish proposed that the indefinite article in “*The door opened and a young girl came in*” transforms the noun “girl” into a semantic predicate (Ilyish, 1948). Despite certain limitations, this theme / rheme concept continues to attract the attention of Russian linguistic researchers (Chernyakhovskaya, 1976; Khlabutina, 2011; Nikitin, 1997; Shevyakova, 1976).

According to the semantic approach taken by Vladimir Arakin, the definite article is treated as a lexical unit that includes the seme of individualisation, uniqueness or generalisation, while the seme of the indefinite article is stated to be that of classification (Arakin, 1989). However, according to Natalia Ogurechnikova, since

articles merge functionally with speech events to mark speech units or “article syntagms”, they can only convey information on the content and quantity of the thus-designated units.

In terms of empirical research, Dilin Liu and Johanna Gleason’s study of the acquisition of proficiency in using the English article system evaluated the competency of 128 ESL low-, intermediate- and advanced students enrolled on an intensive English programme at a university in correctly using *the* when it is obligatory and avoiding its overuse when the indefinite or zero article would be more appropriate (Liu & Gleason, 2002). Liu & Gleason refer to Bickerton’s (1981) typology of the semantic function of a NP, which analyses article use in terms of the combination of two discourse features: (I) whether a noun phrase (NP) denotes a specific referent (\pm SR) and (II) whether (it can be assumed that) the hearer knows the referent (\pm HK). Thus, NPs can be described as falling into four major semantic types:

- 1) Type 1 [–SR, +HK] “generic”; either the indefinite, definite – or, if plural, zero – article is used (*a / the tiger is a fierce animal / tigers are fierce animals*).
- 2) Type 2 [+SR, +HK] requires the definite article:
 - a) unique (or conventionally assumed to be unique) referent (*the Pope*);
 - b) referent physically present (*Pass me the pepper please*);
 - c) referent previously mentioned in the discourse (*the referent I previously mentioned*);
 - d) specific referent assumed to be known to the hearer (e.g., a resident in a small village with one church tells another resident – *My wife is at the church*).
- 3) Type 3 [+SR, –HK], where the indefinite – or, if plural, zero article – is used:
 - a) first mention in the discourse of a [+SR] NP assumed to be unknown to the hearer (*Tom bought a car*);
 - b) first mention of a [+SR] NP that follows “existential have” and is assumed to be unknown to the hearer (*My computer has a new sound card*).
- 4) Type 4 [–SR, –HK], where the indefinite – or, if plural, zero article – is required:
 - a) equative NPs (*She is a single parent*);
 - b) NPs in a negation statement (*I don’t have a car*);
 - c) NPs in an interrogative sentence (*Do you have a pen?*);
 - d) NPs in hypothetical statements (*If I had had more money, I would have bought a new car*).

Bickerton’s classification indicates appropriate usage of articles as informed by user knowledge of their role in marking the various constraints governing the four main semantic NP types (pp. 2–3). Setting aside NP Types 1, 3 and 4, Liu & Gleason then refer to Hawkins’ (1978) Location Theory classification of eight types of non-generic (p. 6) usages of the definite article: (1) ‘anaphoric’; (2) ‘visible situation’; (3) ‘immediate situation’; (4) ‘larger situation requiring specific knowledge’; (5) ‘larger situation relying on general knowledge’; (6) ‘associative anaphoric’; (7) ‘unfamiliar usage with explanatory modifiers’; (8) ‘unfamiliar usage with nonexplanatory modifiers’. Combining Hawkins’ eight categories into four major types, Liu & Gleason distinguish between ‘cultural’ (larger situation usage requiring general knowledge), ‘situational’ (visible, immediate and requiring specific knowledge), ‘structural’ (unfamiliar usage with and without explanatory modifiers) and ‘textual’ (anaphoric including associative) usages of Bickerton Type 2 NPs where the non-generic definite article *the* is required (pp. 6–7).

To summarise, a number of distinct theoretical approaches have been taken to try to explain the uses of articles in English, including traditional, TG, evolutionary, pragmatic, cognitive, theme / rheme and semantic. Potentially useful classification schemes for structuring empirical research into article use based on these theoretical approaches include Bickerton’s NP Semantic Type, Liu and Gleason’s Major Type and Hawkins’ Location Theory Type. The question of definiteness, which justifies the use of the definite rather than indefinite or zero article with a NP, is variously explicated in terms of familiarity, uniqueness, common knowledge and accessibility. Since all these theoretical approaches to explaining the question of definiteness turn out to reference a dialogue between speaker or writer and interlocutor or reader, in which context definiteness emerges, it makes sense to refer to a dialogic paradigm when considering didactic methodologies for teaching English article use to Russian scientists. As far as we are aware, no study has attempted to account for specific English article-use problems arising in the context of Russian science on the basis of existing theory. We selected the existing theoretical models provided by Bickerton, Liu and Gleason, and Hawkins as the most applicable to our enquiry.

Research problem and hypotheses

In order to further examine the research problem concerning an explication of the incorrect or inappropriate use of English articles by Russian authors of research papers, we carried out a content analysis study based on a corpus of sentences derived from seven research papers written in the IMRaD format by Russian scientists in the fields of Organic Chemistry, Chemical Physics, Nanoscience, Marine Biology and High-Temperature Electrochemistry. These texts, aimed at publication in SCOPUS- and/or Web of Science-indexed international journals, had either been originally written in English or translated into English from Russian by a Russian translator. As such, although competently written, the texts were generally of a linguistic standard below that required for publication in international scientific journals and for this reason had been sent for correction by a professional native English editor⁵. The research problem was expanded to include the need to test the applicability of existing classification schemas to the problem of English article use by Russian scientists.

We hypothesised that Russian scientists, being highly educated, would generally tend to overuse *the*, *a* and *an* in an overcompensation for the tendency of less-well-educated Russians to drop articles when using English since the Russian language, in common with other Slavic languages, does not have an article system. In addition, we conjectured that NPs appearing in the Introduction section of a research paper, where there is likely to be a greater amount of confusion as to the definiteness of concepts, would be more likely to involve errors involving the definite article *the* (either overuse or underuse) than in other sections where the definiteness (or not) of NPs should already be clearer to both reader and writer.

Methodology

In order to test these hypotheses and elucidate the research problem, we decided to perform two separate analyses. The first of these was carried out at the level of the sentence in order to test the applicability of Bickerton's NP Semantic Type, Liu and Gleason's Major Type and Hawkins' Location Theory Type schemas, respectively. This schema analysis was performed on 77 sentences containing only one identified article-related error. The second analysis was carried out at the level of the individual NP in order to obtain quantitative information concerning the types of article-related errors in scientific reports written by Russian authors. For this purpose, we used a series of codes that capture all possible permutations of article-related errors, e.g. SxØ√D represented a singular NP taking zero article when it should take the definite article. This error-code analysis was performed on 177 NPs derived from 123 sentences (i.e. including sentences containing more than one article-related error).

First, 123 sentences containing examples of article use deemed incorrect or stylistically-inappropriate⁶ were extracted from the corrected versions of the manuscripts and pasted into a separate Microsoft Word file, leaving intact the tracked changes carried out by the editor. All extracted sentences included a main verb and at least one NP in which an article-related problem⁷ required correction. Sentences containing additional corrected errors were only included if these other errors did not affect the use of articles. Each example sentence was then numbered and listed in two versions: original and corrected. Tracked changes to non-article related errors in the original sentence were accepted and the sentence was copy-pasted below to form the corrected version. Tracked article-related changes were then rejected in the original sentence and accepted in the corrected sentence. The relevant NPs were next underlined in both variants and any article (*the*, *a*, *an*) added by the editor was highlighted in italics in the corrected version.

Identification of NPs was carried out both with respect to the main verb and any verbs governing additional clauses, as well as relative to any genitive constructions e.g. with "of". According to the qualitative analysis approach, all NPs containing article-related errors were identified in the context of the research paper from

⁵ The research papers forming the basis for the content analysis reported here had all been corrected by Thomas Beavitt, one of the co-authors of the present paper, who has over ten years' experience working as a professional academic translator / editor / proof-reader.

⁶ Article errors were categorised according to a process whereby the co-authors of the present article – each having relevant expertise, but in complementary areas – double-checked the categorisation of identified errors according to the logic of the presented schemas.

⁷ Defined here as the incorrect or inappropriate use of definite, indefinite, zero (or null) articles, i.e. either an NP should have taken a definite or indefinite article but didn't, or a definite or indefinite article was used with an NP when a zero (or null) article would have been more appropriate.

which they were derived. Where NPs could be extended without involving any problematic genitive constructions, we did this in order to include the maximum quantity of explicit contextual information in the NP structure. Thus, in the following example, the underlined NP extends all the way to the verb

“The results of mass density, molar volume and number density measurements are shown in Figure 1 along with the respective calculations.”

whereas, in cases where issues were present in genitive constructions, e.g. with “of”, the problematic NPs were separated, as in the example

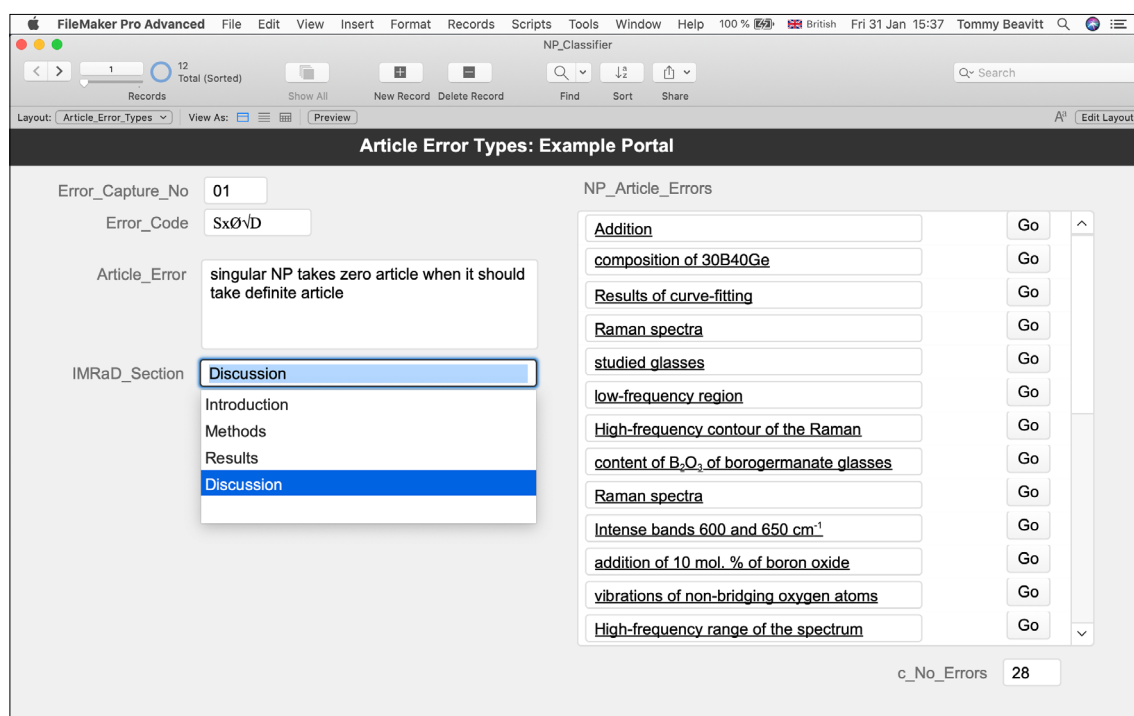
“The dependencies of the relative intensities of high-frequency characteristic bands of Raman spectra on composition are shown in Figure 9.”

The example sentences were then analysed in a specially-designed database solution *NP-Analyser* based on the Filemaker Pro DBMS. First, both the original problematic and corrected variants of each sentence were entered into two respective fields of a database record in a table *Example_Sentences*. Each record was given a unique identifier, linked to its parent document and categorised according to whether it appeared in the Introduction, Methods, Results or Discussion section of the paper. The NPs in 77 sentences containing only one problematic NP were categorised according to three semantic type classification schemas as shown in Tables 1, 2 and 3.

All 123 sentences containing article-related errors, including those sentences containing more than one article-related error, were then analysed according to the type of article-related error occurring in the NP. This exercise yielded a total of 177 problematic NPs featuring nine major error types, which were then imported into a new database table *NP_Article_Errors*. The use of relationships between tables allowed records from *Example_Sentences* and *NP_Article_Errors* tables to be viewed and analysed in portals on different layouts.

For example, the screenshot in Figure 1 shows the use of a portal to display records from the table *NP_Article_Errors* related to the table *Article_Error_Type* by the criterion of error code SxØvD (singular NP takes zero article when it should take definite article). The drop-down list IMRaD_Section further restricts the display of related records in the portal to those that also conform to the filter *Discussion*.

Figure 1
Use of portal to display related records via multi-criteria relationship



Results

The results of NP semantic function categorisation according to Bickerton's NP Semantic Type are as shown in Table 1.

Table 1

Sentence Analysis – Bickerton's NP Semantic Type

Type 1	Generic	0
Type 2	Unique, present, modified, previously mentioned or (assumed to be) known	39
Type 3	First mention in discourse or following existential have, assumed unknown to the interlocutor	14
Type 4	Equative, negative, interrogative and hypothetical objects	9
Unclassified		15
Total		77

It can be seen that none of the 77 NPs identified as problematic involved the generic definite article. 14 (18%) article problems related to NPs of Type 3 (first mention, assumed unknown), while 9 (12%) were associated with NPs of Type 4 (equative, negative, interrogative or hypothetical); in both of these semantic types, a singular NP should take an indefinite article and a plural should take zero article. The majority of identified article-related problems therefore concerned NPs of Type 2 (referents are unique, physically present, previously mentioned or specific and assumed to be known to the reader), for which a definite article is required. 15 of the sentences resisted Bickerton's classification and therefore fall outside the scope of the present work.

Those 39 sentences falling into Bickerton's NP Semantic Type 2 classification were then categorised according to Liu & Gleason's Major Type as shown in Table 2.

Table 2

Sentence Analysis – Liu & Gleason Major Type

Structural	<i>The</i> is used with a first-mention noun that has a modifier.	24
Textual	<i>The</i> is used with a noun that has been previously referred to or is related to a previously mentioned noun.	7
Cultural	<i>The</i> is used with a noun that is a unique and well-known referent in a speech community.	6
Situational	<i>The</i> is used when the referent of a first-mention noun can be sensed directly or indirectly by the interlocutors or the referent is known by the members in a local community, such as the only dog in a family or the only bookstore in a town.	2
Total		39

Here it can be seen that nearly 62% (24) of problematic NPs were those whose semantic type was identified as 'structural', i.e. the definite article should be used with a first-mention noun that has a modifier (typically a genitive phrase with 'of'). 18% (7) problematic NPs were categorised as 'textual'; in other words, they refer back anaphorically to a referent that has already been mentioned in the text. 15% (6) of article errors concerned NPs categorised as 'cultural' (larger situation requiring general knowledge), while 5% (2) were associated with NPs whose semantic function was categorised as 'situational'.

To test the soundness of Liu and Gleason's Major Type construct, we also categorised sample Bickerton Type 2 NPs according to Hawkins' Location Theory Type as shown in Table 3.

From the data it can be seen that Hawkins' category of 'unfamiliar usage in NPs with explanatory modifiers' is fully captured by Liu & Gleason 'structural' type, with both categories accounting for 62% (24) of problematic NPs. Liu & Gleason's 'textual' type also fully captured Hawkins' 'anaphoric' and 'associative anaphoric' categories involving 18% (7) of problematic NPs. However, the 'cultural' and 'situational' uses in Liu & Gleason's typology seemed less well-suited for capturing article issues in scientific writing, where some discrepancies emerged with respect to Hawkins' location theory construct along with various internal consistencies.

Table 3

Sentence Analysis – Hawkins’ Location Theory Type

Type 1	Anaphoric usage	Use of <i>the</i> when something is mentioned a second time and subsequently	6
Type 2	Visible situation usage	Use of <i>the</i> with a noun mentioned the first time to refer to something that both the speaker and the listener can see	1
Type 3	Immediate situation usage	Very similar to Type 2, the only difference being that the referent may not be visible	0
Type 4	Larger situation usage relying on specific knowledge	Use of <i>the</i> with a first-mention noun because it is known in the community	1
Type 5	Larger situation usage relying on general knowledge	Use of <i>the</i> with something that one can assume people from a country or around the world should know	1
Type 6	Associative anaphoric usage	Basically, the same as Type 1, the only difference being that first-mention <i>the</i> is used with a noun that is related to a previously mentioned noun, rather than being the same noun	1
Type 7	Unfamiliar usage in NPs with explanatory modifiers	Use of <i>the</i> with a first-mention noun that has an explanatory or identifying modifier in the form of a clause, prepositional phrase or noun	24
Type 8	Unfamiliar usage in NPs with nonexplanatory modifiers	Similar to Type 7, the only difference being that the modifier does not provide explanatory information	0
Total			34

The results of analysing 177 problematic NPs according to nine major error types are as shown in Table 4.

Table 4

Types of article-related errors in scientific reports written by Russian authors

SxØ√D	Singular NP takes zero when it should take definite article	59
SxIv√D	Singular NP takes indefinite when it should take definite article	12
SxØ√I	Singular NP takes zero when it should take indefinite article	21
SxD√I	Singular NP takes definite when it should take indefinite article	18
SxD√Ø	Singular NP takes definite when it should take zero article	21
SxIv√Ø	Singular NP takes indefinite when it should take zero article	5
PxØ√D	Plural NP takes zero when it should take definite article	26
PxD√Ø	Plural NP takes definite when it should take zero article	8
PxIv√D	Plural NP takes indefinite when it should take definite article	1
SxPoss	Inappropriate use of singular possessive	1
PxPoss	Inappropriate use of plural possessive	1
UnCat	Uncategorisable error	4
Total		177

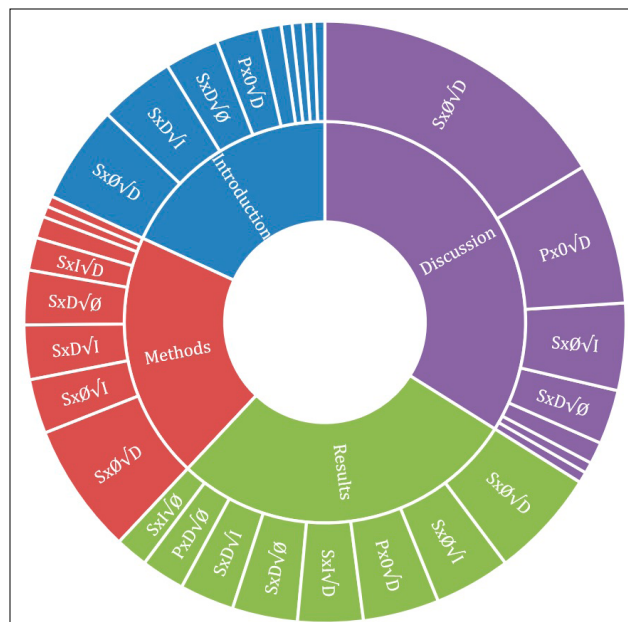
From Table 4, it can be seen that the most common (40%) type of article-related error produced by the authors of the analysed research papers involved a failure to use a definite article with a singular NP when required. Of these, 59 (33%) involved an incorrect use of zero article, while 12 (7%) involved an indefinite article use error. Another common error involved a failure to use an indefinite article with a singular NP when this was required (22%); here, zero article was used incorrectly in 21 cases (12%), while in 18 cases (10%) the definite article was incorrectly used. A slightly less frequent error affecting singular NP constructions concerned the overuse of definite (12%) and indefinite articles (3%) when a zero article would be preferred (total 15%). Article errors also applied to plural NP constructions; of these, 15% involved an incorrect usage of the zero article when a definite article would be preferred, 5% used a definite article when a zero article would have been more appropriate, while less than 1% involved the use of an indefinite article when a definite article was required. We also identified errors involving possessive constructions (a singular or plural possessive was used when a genitive construction with “of the” or an adjectivised NP would have been more appropriate), while around 2% of identified article-related errors were uncategorisable by the methods used in the present study.

In order to test the conjecture that NPs appearing in the Introduction section of a research paper would be more likely to involve errors involving the definite article *the* (either overuse or underuse) than in other

sections where the definiteness of NPs should already be clearer to both reader and writer, we analysed the incidence of article-related errors by IMRaD section. The results are presented in Figure 2.

Figure 2

Article errors by IMRaD section



As can be seen from Figure 2, the largest number of article-related errors were found in the Discussion sections of the analysed research papers. Here, the most frequently encountered error was the use of zero article in singular NPs when a definite article should have been used (SxØvD). Although this was the most frequently encountered error type in all IMRaD sections, it was markedly predominant in the Discussion section, while here the second-most frequent error type – the use of zero article in plural NPs when a definite article should have been used (PxØvD) – accounted for three times fewer errors. Conversely, in both Methods and Results sections, the second-most frequent error was the use of zero article in singular NPs when an indefinite article should have been used (SxØvI). In the Introduction section, however, the SxØvD-type error was closely followed by incorrect use of the definite article with singular NPs when the indefinite article should have been used (SxDvI).

Discussion

The hypothesis that Russian scientists would tend to overuse *the*, *a* and *an* in an overcompensation for the general tendency of Russians to drop articles when using English was not supported by the results of NP error analysis. On the contrary, by far the most frequent article-related error in the analysed sample involved zero article use when (in the estimation of the native English editor) the definite article would have been more appropriate; this finding applies to both singular (33%) and plural (15%) NP constructions. Conversely, only 12% (singular NP) and 5% (plural NP) of errors involved overuse of the definite article, while overuse of indefinite articles accounted for less than 3% of the errors affecting singular NPs. Perhaps unsurprisingly, given its well-understood role in enumerating the singular form, there was only one error (< 0.5%) involving erroneous use of the indefinite article in a plural NP.

Analysis of errors where singular NP takes zero instead of definite article (SxØvD)

Since the most common source of article-related errors consisted in a failure to correctly use the definite article with NPs, we decided to analyse these errors in more detail. In terms of Liu & Gleason's functional NP typology,

we found that almost all of them *could*⁸ be explained in terms of ‘structural’ usage, i.e. where *the* is used with a first-mention noun that has a modifier. Here, by far the most common type of modifier employed in such NP constructions involved genitive phrases with “of”. For example, in

the design of new materials having a low Ba concentration

the NP headword design is sufficiently modified (i.e. definitised) by the genitive complement of new materials having a low Ba concentration to warrant *the*.

In the case of

the characteristic feature of which

the NP headword feature is sufficiently definitised both by the adjective characteristic and the genitive complement of which, where which functions as a pronoun that refers anaphorically to a previous NP.

In addition to NPs definitised by genitive complements with ‘of’, a number of other modifier constructions may be identified. For example, in the case of

the so-called “germanate anomaly”

“germanate anomaly” can be seen as having been sufficiently definitised by the adjective phrase so-called to warrant *the*.

In the case of

the other side

it is not so clear whether ‘structural’ is the most clarifying explanation. Perhaps it is also possible to think of *the* being used in a ‘textual’ or even ‘cultural’ sense? In any case, it seems clear that the headword side is definitised by the adjective other to the extent that it exists in opposition to a ‘this side’ that already exists in the present dialogic context.

Similarly, in the case of

the studied glasses

studied functions not so much as a definitising modifier, but as a reminder that the particular glasses to which the discourse presently refers have not only been previously mentioned in the text, but indeed are of central concern to the entire communicative endeavour.

While

the low-frequency region

may likewise be interpreted in terms of the headword region being ‘structurally’ definitised by the adjectival phrase low-frequency, here a better explanation may be in terms of a ‘situational’ NP function (since the referred-to region is present to both writer and reader in the form of a diagram).

Again, with

the intense bands at 600 and 650 cm⁻¹

⁸ We once more emphasise that many of the article use choices made by professional editors with English as a first language are stylistic rather than strictly grammatical; this principle also applies to the classifications of article errors presented in this research paper, which are often open to more than one possible interpretation.

it is possible to consider the headword bands as being definitised by the modifier intense and the locative prepositional phrase at 600 and 650 cm⁻¹ ('structural'); however, here there is also a 'situational' aspect produced by the visual presence of a graph.

In the case of

the replacement of germanium by boron

we can see a very typical 'structural' use of the in scientific language, where the headword replacement is actually a rather indefinite verbal noun; therefore, its definitising modification by of germanium by boron is highly 'structural', not to say formulaic.

Again, with

the 40B30Ge composition

the definitising modification of the headword composition by the abbreviation 40B30Ge is clearly and typically 'structural'.

One or two examples of 'textual' (i.e. anaphoric) usage can be identified, e.g. the NP

the glass

is used to identify a referent that appeared earlier in the text (if the definite article were dropped, as per the uncorrected sample, the statement predicated on this NP would appear to apply to glass in general).

Likewise, in

the current paper

current functions not so much as a definitising modifier ('structural') as a discourse marker, whose function does not appear to be fully captured by either Liu & Gleason's 'textual' category or Hawkins' 'anaphoric' or 'associative anaphoric' descriptors.

Although in the cases of

the sol-gel method

and

the non-trivial Berry phase of π

the presence of modifiers definitising the headwords method and phase may be noted, it is probably better to explain this use of the definite article in terms of Hawkins' 'larger situation usage relying on specific knowledge' (simplified by Liu & Gleason into the broader 'cultural' category). Here both sol-gel method and non-trivial Berry phase are terms with which a specialist in the field of the respective research paper would be expected to be familiar.

Finally, we also found evidence of what Hawkins describes in terms of 'larger situation usage relying on general knowledge', also simplified by Liu & Gleason into 'cultural'. Here

the coastal zone

the open sea

the continental shelf in the southern part of the North Sea, English Channel and Irish Sea

appear to be definitised by general (i.e. geographic) knowledge that any non-specialised reader might be expected to possess.

Analysis of error types by frequency of occurrence in IMRaD sections

The conjecture that NPs appearing in the Introduction section of a research paper would be more likely to involve errors involving the definite article *the* (either overuse or underuse) than in other sections due to the definiteness of NPs being less clear received partial confirmation by our analysis of the incidence of article-related errors by IMRaD section (see Figure 2). In the Introduction section, the five most common article-related error types involved overuse or underuse of the definite article, while in the Methods and Results sections, the second-most frequent error was the use of zero article in singular NPs when an indefinite article should have been used (SxØvI); in the Discussion section, this error (SxØvI) was the third most frequent type.

A few examples of the kinds of errors involving overuse or underuse of the definite article in the Introduction section are discussed below:

Since *the conductivity of these materials* is lower than those of BaCeO₃ or BaZrO₃, *the application of yttrates* is unlikely to be practical under real conditions.

Here it appears that, although the SxØvD error concerning *application of yttrates* is purely ‘structural’, the failure to correctly use *the* with *conductivity of these materials* (SxØvD) may be partially attributed to the author not realising that the anaphoric *these* sufficiently definitises *materials* (referred to earlier in the Introduction section).

The summer maximum phytoplankton development period begins with *a bloom of nitrogen-fixing blue-green algae*.

Here, the author’s use of a definite article with the NP *bloom of nitrogen-fixing blue-green algae* has been corrected by the editor to indefinite article (SxDvI). The reason for this is that the reader is being provided with new information about a seasonal bloom that occurs regularly. It should not be supposed that even a specialist reader will know the exact nature of such blooms; besides, even if he or she does, the way such information is conventionally presented in the Introduction section prohibits the use of the definite article.

A similar error occurs with

The recent discovery of *[the] type-I Weyl semimetals (WSMs) belonging to the TaAs family of compounds* has stimulated research efforts to study new WSM candidates exhibiting breaking of either time-reversal or spatial inversion symmetry.

In this case, due to the NP being plural, the author’s overuse of the definite article has been corrected by the editor to zero article (SxDvØ). Again, since the reader is being provided with new information about a recent discovery, the way such information is conventionally presented in the Introduction section forbids the use of the definite article.

In the case of

In such a situation, QAHE vanishes due to the Berry curvatures contributed from *the two Weyl pairs* cancelling each other out.

the author’s use of zero article with a plural NP has been corrected with the definite article (Px0vD). Since a specialist reader may be supposed to already know that Weyl nodes always come in pairs, the use of the definite article can be seen here as ‘cultural’; moreover, a non-specialist native English reader already knows that interacting pairs are already sufficiently (‘structurally’?) definitised to take a definite article.

Similarly, with

During autumn and winter seasons, a meandering frontal zone forms between the relatively cold coastal and warm marine waters; a similar phenomenon also occurs between the relatively fresh coastal and relatively saline marine waters during spring and summer.

the author's use of *a* with relatively fresh coastal and relatively saline marine waters has been corrected to *the* (SxI√D); however, in this case it may be the competition between relatively fresh coastal and relatively saline marine waters that results in the definitisation of the combined NP.

The unexpectedly significant preponderance in the Discussion section of the error type involving use of zero article in singular NPs when a definite article should have been used (SxØ√D) also seems worthy of discussion. Here it would appear that Russian authors may not be comfortable with using *the* to refer to concepts already definitised within the present discourse (i.e. in the Introduction, Methods and Results sections).

For example, in

The results of curve-fitting applied to the Raman spectra of the studied glasses are presented in Tables 2-4 for compositions 0B70Ge, 70B0Ge and 35B35Ge respectively.

the three NPs corrected with the addition of *the* (SxØ√D) have all been sufficiently definitised by their previous appearance in the earlier sections of the paper.

A similar consideration applies to plural NP constructions such as

The obtained dependencies allow us to state that initially there is a weak change in properties until the ratio of B/Ge in the glass becomes equal to 2 (35B35Ge).

Here, the headword dependencies is not only definitised 'structurally' by the adjective obtained but has additionally been definitised 'textually' by previous discussion, i.e. it is an example of anaphoric usage.

Conclusion

The results of our review of the literature on the English article system show that, while much useful theoretical work has been done, a comprehensive explicatory system – or universal 'algorithm' – has yet to be developed. Our analysis of article-related NP errors produced by Russian scientists according to the categories developed by Bickerton, Hawkins and Liu and Gleason showed that, while useful, these typologies neither account for all errors nor provide hard-and-fast distinctions between the possible rationales according to which errors produced in non-native scientific writing are to be corrected. In particular, Liu and Gleason's attempted simplification of Hawkins' Location Theory construct into 'structural', 'textual', 'cultural' and 'situational' uses was seen to result in a reduced ability to capture article-related issues in scientific writing. Regardless of the typology used, however, a significant number of article errors were observed to occur whose solution could be explained according to more than one NP functional typology.

It can be seen that, especially as used in the Introduction and Discussion sections of an IMRaD research paper, the choice of definite, indefinite or zero (Ø) articles implies the emergence of a "dialogic WE" – i.e. in the terms of Epstein and Fauconnier, a shared mental space – in which the emerging contextual definiteness of specialised terms and concepts becomes part of the implicit – and, through the process of communication, explicit – knowledge shared between writer and reader (or speaker and interlocutor). In our opinion, while the majority of the theoretical work reviewed in this paper has some pedagogical value, the most productive approach to developing an understanding of the English article system for scientists would consist in drawing attention to the rhetorical role played by articles in the development of context according to a *dialogic paradigm*, in which, as proposed by Richard Epstein, articles function in the construction of sets of common mental spaces and the

distribution of information within these mental spaces, as well as helping to signal the accessibility of knowledge in a given space relative to other spaces.

In future work, we aim to carry out a larger-scale content analysis of English article-related errors produced by Russian scientists, as well as to develop a theoretical basis for categorising and explicating such errors as informed by an understanding of their role in structuring the developing context of scientific communication according to the paradigm of the “dialogic WE” of writer and reader. An additional relevant direction for future research would be to carry out a detailed examination of how the Russian language encodes the referentiality function of the English article system.

An Appendix, containing the list of article errors derived from research papers written by Russian scientists and forming the basis for the empirical research detailed in this paper, is available on request.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Alangari, M., Jaworska, S., & Laws, J. (2020). Who's afraid of phrasal verbs? The use of phrasal verbs in expert academic writing in the discipline of linguistics. *Journal of English for Academic Purposes*, 43, 100814. <https://doi.org/10.1016/j.jeap.2019.100814>
- Arakin, V. (1989). *Comparative typology of the English and Russian languages*. Prosveshhenie.
- Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. Routledge.
- Bickerton, D. (1981). *Roots of language*. Karoma Press. <http://langsci-press.org/catalog/book/91>
- Chernyakhovskaya, L. (1976). *Translation and the structure of meaning*. Mezhdunarodnye otnosheniya.
- Chesterman, A. (2005). *On definiteness: A study with special reference to English and Finnish* (Digitally printed 1st pbk. version). Cambridge University Press.
- Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, 30(3), 351–379. <https://doi.org/10.1177/0267658313493432>
- Christophersen, P. (1939). *The Articles: A Study of Their Theory and Use in English*. Einar Munksgaard.
- Dolgina, E. (2008). The study of the articles in modern English: A cognitive approach. *Issues of Cognitive Linguistics*, 25–32.
- Epstein, R. (2002). *The definite article, accessibility, and the construction of discourse referents*. <https://doi.org/10.1515/cogl.2002.007>
- Grannis, O. C. (1972). The Definite Article Conspiracy in English. *Language Learning*, 22(2), 275–289. <https://doi.org/10.1111/j.1467-1770.1972.tb00088.x>
- Hawkins, J. (1978). *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. Routledge.
- Hawkins, J. (1980). ON SURFACE DEFINITE ARTICLES IN ENGLISH ‘Proving Underlying Indefiniteness and Explaining the Conversion to Definiteness’. In *The Semantics of Determiners (RLE Linguistics B: Grammar)* (Vol. 24, pp. 41–66). Routledge.
- Hawkins, J. (1991). On (In)definite Articles: Implicatures and (Un)grammaticality Prediction. *Journal of Linguistics*, 27(2), 405–442. JSTOR.
- Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research*, 8(1), 5–29. <https://doi.org/10.1191/1362168804lr132oa>
- Ilyish, B. (1948). *Modern English language*. (2nd ed.). Izdatelstvo literatury na angliiskom yazyke.
- Jiang, F. (Kevin), & Hyland, K. (2020). “There are significant differences...”: The secret life of existential there in academic writing. *Lingua*, 233, 102758. <https://doi.org/10.1016/j.lingua.2019.102758>
- Khlabutina, L. (2011). Indefinite article as rheme actualization means in scientific discourse. *Philological Sciences. Issues of Theory and Practice*, 2, 182–186.
- Kirchik, O., Gingras, Y., & Larivière, V. (2012). Changes in publication languages and citation practices and their

- effect on the scientific impact of Russian science (1993–2010). *Journal of the American Society for Information Science and Technology*, 63(7), 1411–1419. <https://doi.org/10.1002/asi.22642>
- Kubryakova, E., & Iriskhanova, O. (2007). Linguistic abstraction in the names of categories. *The Bulletin of the Russian Academy of Sciences: Studies in Literature and Language*, 66(2), 3–12.
- Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, 31(1), 25–35. <https://doi.org/10.1016/j.esp.2011.07.002>
- Liu, D., & Gleason, J. L. (2002). Acquisition of the article ‘the’ by non-native speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition*, 24(1), 1–26. JSTOR.
- Master, P. (1995). Consciousness raising and article pedagogy. In D. Belcher & G. Braine (Eds.), *Academic writing in second language: Essays on research and pedagogy* (pp. 183–204). Norwood, NJ: Ablex.
- Nikitin, M. (1997). *Course on linguistic semantics*. Nauchnyj centr problem dialoga.
- Oakey, D. (2020). Phrases in EAP academic writing pedagogy: Illuminating Halliday’s influence on research and practice. *Journal of English for Academic Purposes*, 44, 100829. <https://doi.org/10.1016/j.jeap.2019.100829>
- Popova, N. G., & Beavitt, T. A. (2017). English as a means of scientific communication: Linguistic imperialism or interlingua? *Integration of Education*, 21(1), 54–70. <https://doi.org/10.15507/1991-9468.086.021.201701.054-070>
- Popova, N. G., Biricheva, E. V., & Beavitt, T. A. (2018). Three Aspects of the Phenomenon of Science: In Search for Unity among Sociologists. *The Education and Science Journal*, 20(9), 35–55. <https://doi.org/10.17853/1994-5639-2018-9-35-55>
- Postal, P. (1969). On So-Called ‘Pronouns’ in English. *Modern Studies in English*.
- Rasskazova, T., Guzikova, M., & Green, A. (2017). English language teacher development in a Russian university: Context, problems and implications. *Ampersand*, 4, 1–9. <https://doi.org/10.1016/j.amper.2017.01.001>
- Shevyakova, V. (1976). *Actual division of the sentence*. Moscow: Nauka. Nauka.
- Smirnitsky, A. (1959). *Morphology of the English language*. <http://www.libex.ru/detail/book52642.html>
- Sommerstein, A. H. (1972). On the So-Called Definite Article in English. *Linguistic Inquiry*, 3(2), 197–209. JSTOR.
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Taylor, H., & Goodall, J. (2019). A preliminary investigation into the rhetorical function of ‘I’ in different genres of successful business student academic writing. *Journal of English for Academic Purposes*, 38, 135–145. <https://doi.org/10.1016/j.jeap.2019.01.009>
- Whitman, R. L. (1974). Teaching the Article in English. *TESOL Quarterly*, 8(3), 253–262. JSTOR. <https://doi.org/10.2307/3586169>
- WonHo Yoo, I. (2008). A Corpus Analysis of (The) Last/Next + Temporal Nouns. *Journal of English Linguistics*, 36(1), 39–61. <https://doi.org/10.1177/0075424207306710>
- WonHo Yoo, I. (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. <https://doi.org/10.1016/j.jeap.2009.07.004>
- Yang, A., Zheng, S., & Ge, G. (2015). Epistemic modality in English-medium medical research articles: A systemic functional perspective. *English for Specific Purposes*, 38, 1–10. <https://doi.org/10.1016/j.esp.2014.10.005>
- Zanina, E. (2016). Strategic Hedging: A Comparative Study of Methods, Results and Discussion (and Conclusion) Sections of Research Articles in English and Russian. *Journal of Language and Education*, 2(2), 52–60. <https://doi.org/10.17323/2411-7390-2016-2-2-52-60>