

The Use of the English Article System by Russian Learners

Olena Martynchuk



In Partial Fulfilment of the Requirements for the MA Degree

UNIVERSITY OF OSLO

May 2010

Content

Chapter 1	Introduction	1
1.1	Previous research	2
1.2	Methodology	2
1.3	Material	3
1.4	Structure	3
Chapret 2	Theoretical framework	5
2.1.	Functions and meanings of the English articles	5
2.1.1	The definite article	8
2.1.2	The indefinite article	11
2.1.3	Zero/ null article	13
2.2	Definiteness and specificity in Russian	14
2.2.1	Other lexical marking on indefinites in Russian	15
2.2.2	Word order, definiteness, and specificity in Russian	16
2.3	Definiteness and specificity in Chinese	17
2.4	Huebner's classification and the Fluctuation Hypothesis	18
2.5	Previous studies based on Huebner's classification	22
2.6	Acquisition orders revealed in previous studies for the [-Art] group of learners	23
Chapter 3	The Method	28
3.1	The data	28
3.2	Tagging system	30
3.3	Untagged	32
3.4	Classification scheme	34
Chapter 4	The Study	42
4.1	Scope	42
4.2	Hypothesis	42
		42

4.3 Research questions	43
4.4 Structure	44
4.5 Measure systems	45
4.6 Results and discussion	46
4.6.1 Article acquisition order for Russian learners based on SOC	
4.6.2 Misuse	47
4.6.3 Overuse	52
4.6.4 Acquisition order based on the TLU measure system	54
4.7 Which of the two acquisition orders is the more reliable one?	55
4.7.1 Conclusions	59
4.8 Comparison of article use by Russian and Chinese learners	59
4.8.1 Inconsistencies of the article acquisition orders of Russian and Chinese learners	60
4.8.2 Comparison of Russian learners to Chinese and Spanish according to the accuracy rates revealed in the studies based on TLU	63
4.8.3 Misuse	65
4.9 Articles' accuracy in accordance with four types of semantic context	70
Chapter 5 Conclusions	78
Attachment 1	85
Attachment 2	86

Tables

Table 1. Article usage for the non-generic reference	7
Table 2. Functions of the zero article	13
Table 3. Article Grouping Cross-linguistically: Two-article Languages	20
Table 4. Accuracy rates for the Chinese learners based on TLU	26
Table 5. Tagset for the articles used correctly	30

Table 6.	Tags for incorrect use of English articles in each semantic context	31
Table 7.	Accuracy rates obtained by the SOC measure system for the definite, indefinite and zero articles	46
Table 8.	Misuse percentages of the definite, indefinite and zero articles based	48
Table 9.	Overuse percentages of the definite, indefinite and zero articles in non-obligatory contexts based on the SNOC measure system	53
Table 10.	Accuracy percentages for the definite, indefinite and zero articles based on the TLU measure system	54
Table 11.	Percentages of accuracy for the definite, indefinite and zero articles relative to the number of the instances of the definite, indefinite and zero articles supplied in the data	57
Table 12.	Percentages of accuracy for Russian, Chinese and Spanish learners revealed by the TLU measure system	63
Table 13.	Article misuse in obligatory contexts for definite, indefinite and zero contexts based on the TLU measure system	66
Table 14.	Percentages of accuracy for the Russian, Chinese and Spanish learners in the four semantic contexts	72

Figures

Figure 1.	Bickerton's (1981) semantic wheel for NP reference	18
Figure 2.	Percentages of the article accurate use and misuse in obligatory contexts for each of the articles	50
Figure 3.	Percentages of the accuracy and overuse rates of the articles	53
Figure 4.	Accuracy rates for each of the articles calculated based on the SOC and TLU measure systems	56
Figure 5.	Accuracy percentages for the definite, indefinite and zero articles in obligatory contexts and relative to the total amount of instances	58

supplied for each of the articles.

Figure 6. Accuracy percentages for the definite, indefinite and zero articles based on TLU for the Russian, Chinese and Spanish learners	63
Figure 7. The diagram shows the percentage of the zero and indefinite articles used inappropriately in definite contexts	67
Figure 8. Percentages of the definite and indefinite articles inappropriately used in zero contexts	68
Figure 9. Percentages of the zero and definite articles used inappropriately in indefinite article contexts	70
Figure 10. Accuracy rates for the Russian, Chinese and Spanish learners of English by semantic context	73

Chapter 1 Introduction

This thesis is a corpus-based study of learners' use and acquisition of the English articles. Although English articles are highly frequent and traditionally introduced at the early stages of English teaching, they cause many difficulties for learners and especially for those whose first language lacks articles. Due to the fact that article choice is context-specific, difficulties in appropriate article use persevere till the very last stage of English language acquisition.

A very interesting example of the difference an article can make is Neil Armstrong's famous first words from the moon in 1969, "That's one small step for man, one giant leap for mankind"¹. Some historians and grammarians have criticized Armstrong for not using the more grammatically correct version, "One small step for a man...". According to English grammar (considering missing *a*), what Armstrong did say was, "One small step for mankind, one giant leap for mankind." since *a man* is used to refer to one representative of a kind while *man* is generic reference in which all or most members of a set are referred to. Thus, as it turns out even native speakers are not always 100 % correct in their use of articles.

For example, Russian learners are capable of making such a mistake as using the definite article with *mankind* (e.g. Perhaps it would have been better for the mankind if [...]), which is semantic nonsense since the meaning of the word *mankind* is the human race: the totality of human beings (<http://www.merriam-webster.com/dictionary/mankind>) and the meaning of the definite article with plural or mass nouns is to "[...] refer to the totality of the objects or mass in the relevant shared set." (Hawkins 1978:159). This proves that learners whose first language does not have an article system do not feel how an article can change the meaning of noun phrase since for them *mankind* and *the mankind* has the same meaning.

Being a Russian native speaker myself I have always been interested in the acquisition process of the articles by learners who belong to the [-Art] group of speakers, i. e. languages that lack articles. Do all learners who belong to this group acquire articles in the same order?

¹ The quote is taken from Monika Ekiert's study (2007) "The Acquisition of Grammatical Marking of Indefiniteness with the Indefinite Article *a* in L2 English".

1.1 Previous research

For a long period of time many linguists have studied article acquisition by non-native speakers of English with different language backgrounds. One of their aims has been to define which of the articles is acquired first. According to their findings learners who belong to the [-Art] group acquire the zero article first as it is used most accurately in its obligatory contexts, the definite article second and the indefinite article last.

Most of the previous studies (Yamada and Matsuura 1982; Master 1987; Parrish 1987; Thomas 1989) have used the SOC (Supplied in Obligatory Contexts) measure system to estimate acquisition order. However, as Pica (1984) mentions in her “Methods of morpheme quantification” SOC does not take into account overuse of morpheme in non-obligatory contexts. Pica (1984) claims that if the morpheme is overused or overgeneralized SOC will overestimate the learner's accuracy and offers TLU (Target Like Use) as a more reliable calculation method for the estimation of the article acquisition order.

Unlike previous researchers Díez-Bedmar and Papp (2008) have used Pica's measure system (TLU) in their study of the use of articles by Chinese and Spanish learners. Their results go against those found in the previous studies as the indefinite article seems to be less difficult to acquire for both [-Art] and [+Art] groups, and it is the definite article that is problematic, even at an advanced stage of acquisition. They believe that their results contradict previous studies due to the appliance of another measure system (TLU).

The aim of my study is to answer the following questions: what do SOC and TLU tell us about the use of articles by Russian learners in terms of acquisition order and accurate use?; which of the two measure systems is more appropriate for the definition of the acquisition order?; what is the difference in article use between Chinese and Russian learners?

1.2 Methodology

There are two main methodologies used for article acquisition studies: longitudinal and cross-sectional. The first is a kind of observational study where the same people are studied for a

period of time and thus reveals changes more accurately. The second is also a type of observational study, however, it concerns a single point in time. Cross-sectional research is based on the assumption that the overall finding revealed in a target group is typical for the whole group. Thus, only cross-sectional methodology is appropriate for corpus based studies on article acquisition. Since the present study will be corpus based I will apply the cross-sectional methodology.

1.3 Material

The only corpus I have come across which contains appropriate data for the present study is An International Corpus of Learner English (ICLE). The ICLE is a computerized corpus of argumentative essays on different topics written by advanced learners of English (university students of English mainly in their second or third year) (<http://web.abo.fi/fak/hf/enge/research/ICLE.html>). That is why the data for my investigation is retrieved from the Russian part of the ICLE corpus (RICLE) which consists of 276 essays and makes all together 229,584 words. The RICLE is presented by argumentative essays which have an average length of 832 words. All of the Russian learners, whose essays compile RICLE, studied at Moscow Lomonosov State University and were about 21 years old, 84% of whom are girls. They all have learned English primarily in a classroom setting and English is rather a foreign language for them than a second language as most of them have never been to any English speaking country.

1.4 Structure

The thesis will be structured as follows:

Chapter 2 first considers the description of the English article system, in particular, the meanings and functions of the articles. Then follows an account on the way definiteness and specificity are realized in the Russian and Chinese languages. A review of theoretical approaches to and previous research on the acquisition of English articles will follow.

Chapter 3 is dedicated to the methodology, specifically, data retrieval, tagging system, classification scheme.

Chapter 4 gives a detailed account of the findings with regard to article accuracy, overuse, acquisition order and semantic types of articles.

Chapter 5 is

Chapter 2 Theoretical framework

In this Chapter I will provide some background to the use of articles in English and realization of definiteness and specificity in Russian and Chinese to pave the way for the current research. Two theoretical approaches to research on article acquisition will be presented: Huebner's classification and Ionin's Fluctuation Hypothesis, which have been used previously in studies concerning the [-Art] group of learners. I will also give evidence of why Huebner's classification is more appropriate for the present study. Then I will present previous findings on article use and acquisition orders for the [-Art] group of learners. Possible reasons for the inconsistencies in acquisition orders as to proficiency levels will be suggested. A fuller account of how accuracy and acquisition rates can be measured will be given in connection with the analysis of the data in Chapter 4 .

2.1. Functions and meanings of the English articles

[Articles'] meanings are abstruse. They refer to no particular object, class of objects, or class of actions, as do, for example, *mommy*, *dog*, or *push*, or even a consistent internal feeling such as is nominated by *want*. Their meaning inheres in the semantically abstract notions of specificity of reference and the specificity of a reference for their listener. Each of these presents what abstractly seem like severe problems of conceptualization. (Maratsos 1976:7-8)

According to Quirk et al. (1985) articles belong to the central determiners which occur before the noun acting as a head of the noun phrase and are used to refer to the linguistic or situational context. Depending on which of the three articles is used the noun phrase will have definite or indefinite reference. Definite reference is typically indicated by the definite article and indefinite reference by the indefinite or zero article.

1. a) Have you seen the bicycle? (definite)

b) Have you seen a bicycle (bicycles)? (indefinite) (Quirk et al. 1985:253)

Apart from definite or indefinite reference articles can also be used for specific or generic reference. For example, in 2a the reference is specific since we have in mind particular representatives of the class lions and tigers. When the reference is generic it is to the whole class of tigers and not to one or some representatives of the class of tigers. *Tigers* in 2b refers to the whole class of tigers and provides us with a general characteristic of the class.

2. a) A lion and two tigers are sleeping in the cage. (specific)

b) Tigers are dangerous animals. (generic) (ibid. 1985:265)

Quirk et al. (1985:265) think that the distinctions between definite and indefinite, and between singular and plural which are crucial for specific reference are of less importance for generic reference, as it is used to denote all or most members of the class or species.

3. a) A lion is a noble beast.

b) The lion is a noble beast.

c) Lions are noble beasts.

d) The lions are noble beasts. (ibid. 1985:265)

While Quirk et al. (1985) distinguish between specific and generic reference, Celce-Murcia and Larsen-Freeman (1999) propose generic vs. non-generic distinctions where non-generic reference is further subdivided into specific reference and non-specific reference. Specific reference is a reference to a particular member of the class signaled by the definite article. Non-specific reference is reference to no particular member of the class or to no member at all, which is signaled by indefinite articles *a/an* in the singular and no article in the plural.

However, there is another important aspect of the specific/non-specific distinction that influences our choice of the articles that is the speaker/listener's common knowledge and the shared situation of utterance (see Table 1).

Table 1. Article usage for the non-generic reference (cited from Ekiert 2007:9)

Listener/ Reader	Speaker/Writer	
	Specific referent	Specific referent Can I have the car?
Nonspecific referent	I saw a funny looking dog today.	I don't have a car. I need a new belt.

According to the table above the definite article is used only in cases where a referent is known to both speaker/listener or writer/reader as only in such case the reference can be defined as specific. If only one of the participants of the shared situation of utterance has a definite referent in mind the reference is non-specific since the referent is indefinite for one of the participants. In cases where the referent is unknown for both of the participants the reference is non-specific and refers to no member of the class at all, but to the notion of one as in *I don't have a car/I need a belt*.

Fodor and Sag (1982) propose a different way of looking at specific/non-specific distinctions for the indefinite article. They argue that English indefinites are ambiguous between referential (specific) and quantificational (non-specific) interpretations.

4. a) A man just proposed to me in the orangery.
(though I'm much too embarrassed to tell you who it was)
- b) A man is in the woman's bathroom.
(but I haven't dared to go in there to see who it is)

According to Fodor and Sag's theory, in the first example *a man* refers back to a particular man as the speaker is clearly aware of the identity of the man who proposed to her, and simply does not mention who it was to the hearer. In the second example it is clear that the speaker has no

particular referent in mind, as the woman reports the presence of some, unfamiliar to her, man in the woman's bathroom. The crucial distinction between the two examples is that in the first the speaker intends to refer to a particular individual, whereas in the second no particular person is meant. For Fodor and Sag, example 4a contains a referential (specific) indefinite while 4b contains a quantificational (non-specific) indefinite. Thus, for Fodor and Sag (1982) speaker's knowledge of the referent is the crucial determinant of specificity. Consequently, *a man* is [+specific] in (4a) and [-specific] in (4b) from the speaker's point of view.

However, as I understand it, in both examples the speaker refers to a particular man. In 4a it is *a man* that proposed to her and in 4b it is *a man* who is in the woman's bathroom at the moment of reference. The only difference between the two men is that in 4a the speaker knows the identity of a man while in 4b the identity of a man is not known to the speaker. Even though in 4b any man can be in the bathroom since the speaker does not know who he is, it is obvious that at the moment of utterance there is a particular man in the bathroom whose identity is not known. I would define both references as specific indefinites as I think that an example of a non-specific indefinite reference would be: *I need a man*. In such a sentence the reference is indeed to any member of a class rather than to any specific representative of a class.

Even though there are some inconsistencies in the interpretation of the uses and functions of the articles, most grammarians agree that articles can be used for specific and generic reference and that common knowledge and shared situation of utterance play an important role in the definition of the reference of an article.

2.1.1 The definite article

Since articles can be used for both generic and non-generic reference further discussion will provide a better understanding of the meanings conveyed by each of the articles when used generically and non-generically.

Quirk et al. (1985:256) summarize the overall function of the definite article as follows: “the definite article *the* is used to mark the phrase it introduces as definite, i.e. as referring to

something which can be identified uniquely in the contextual or general knowledge shared by speaker and hearer.” Based on the assumption that the use of *the* relies on shared knowledge, Quirk et al. suggest several contexts in which the identity of the referent may be 'recovered' by the hearer. These contexts are immediate situation, larger situation (general knowledge), anaphoric reference (direct and indirect), cataphoric reference, sporadic reference, the 'logical' use of *the*, the use of *the* with reference to body parts.

Hawkins (1978), in his attempt to develop a theory of the referential meaning of the English noun phrase containing the definite and the indefinite articles, identifies three uses of the definite article: the anaphoric use and two situational uses: the immediate situation use and the larger situation use, relying on specific knowledge or on general knowledge about the referent and Hawkins argues that:

The use of the definite article involves a pragmatic strategy which consists in matching a referent with a whole set of objects rather than in identifying an object as such in the world. Set identification and referent location are characteristic of all uses of *the*, even though the process of actually locating the referent within the set may require different pragmatic abilities from one usage type to another. (Hawkins 1978:129)

He also compares the effect of the definite article to that of placing a 'pragmatic blanket' over some of the infinite number of possible or potential referents of a referring predicate. The objects falling under this pragmatic blanket are those existing in the shared speaker-hearer sets.

Maria Belen Díez-Bedmar and Szilvia Papp (2008) mention that Hawkins' important notion of the meaning of the definite article is that of “uniqueness” with which they disagree since it has been pointed out by Lyons (1999:278) that definiteness with plurals and mass nouns involves not uniqueness but inclusiveness or even identifiability. I have to disagree with their interpretation of Hawkins' notion of 'uniqueness' of the definite article as the following quotation from Hawkins proves that he does not consider 'uniqueness' to be a central or important notion for the description of the definite article.

I shall now argue that uniqueness cannot be any part of the meaning of *the* as such, even though definite descriptions with singular count nouns do refer

uniquely. It is, instead, crucially the singularity of the count noun, in conjunction with *the*, which makes the total reference unique. Thus, uniqueness results from a fusion of the meaning of the definite article with singularity or oneness, but it is not a part of the meaning of *the* itself. (Hawkins 1978:158).

So far we have considered only the non-generic meaning of the definite article but as I have mentioned before it can also be used generically. Hawkins (1978) suggests that articles have two referential functions since the nature of specific and generic reference is fundamentally the same. According to Hawkins (1978) *the lion* refers inclusively in both examples, although the parameters relative to which inclusive reference is being made in the first example are no longer such as in the second example where a unique and singular lion is being referred to. This means that the definite article refers inclusively within pragmatically defined parameters in both examples (5a and 5b). Inclusiveness, however, in the case of generic reference spans the whole class of lions, as in 7a.

5. a) *The lion* is a noble beast.

b) *The lion* is in the cage. (Hawkins 1987:216)

[...] it seems that we can capture a generalization between non-generic and generic definite article usage by assuming that inclusiveness is common to both in a way that we certainly could not if we assumed uniqueness to be the significant generalization for non-generic definite descriptions. Inclusiveness is compatible with any number of objects being referred to, including the whole class of potential referents at one end, and a unique object at the other. Thus, the universal idea is present in all uses of *the*, both generic and non-generic. (Hawkins 1987:216).

Hawkins also finds evidence for a pragmatic blanket effect being true for both specific and generic reference.

6. a) *Italians* are lazy.

b) *The Italians* are lazy. (Hawkins 1987:216)

The difference between these two examples is that in 6a laziness is understood as an inherent trait of all Italians while in 6b laziness involves a pragmatic restriction and refers only to those Italians who currently inhabit Italy. Having discovered such striking similarities between specific definite reference and generic definite reference, Hawkins (1978) questions whether the label 'generic' should be used for the example *The Italians are lazy* as it differs from the specific definite reference only over the quantity of referents being referred to.

Master (1997:225) suggests that if the definite article is used generically, it must occur with a singular count noun, as “generic *the* can only indicate a class of entities and not representatives of the class” and that in the cases when the definite article is used with plurals “it indicates non-generic usage, even though in some cases we know pragmatically that not every single member is meant (e.g. The Eskimos are good hunters)”.

2.1.2 The indefinite article

The indefinite article is notionally the 'unmarked' article in the sense that it is used (for singular count nouns) where the conditions for the use of *the* do not obtain. That is, *a/an X* will be used where the reference of X is not uniquely identifiable in the shared knowledge of speaker and hearer. Hence *a/an* is typically used when the referent has not been mentioned before, and is assumed to be unfamiliar to the speaker or hearer. (Quirk et al. 1985:272)

In addition to the claim made by Quirk et al. in the quotation above, they also state that “unlike the definite article, the indefinite article does not signal co-reference with a preceding indefinite noun phrase” (ibid: 272).

While Quirk et al. (1985) explain the indefinite article uses as those where conditions for the use of *the* do not apply, Master (1997) contrasts the function of the indefinite article with that of the zero article:

The general function of *a* is the opposite of \emptyset : it signifies a boundary (or creates one where there was none before) that makes a formless entity

discrete and thereby countable. It thus occurs most frequently with singular count nouns. However, it also frequently occurs with mass nouns in a specialist context. For example, we generally consider steel to be a non-count noun, but a specialist may speak of *a high-grade steel*. *A* thus serves to create a boundary, which we interpret to mean *a kind* or *type of*. (Master, 1997:225)

The central referential meaning of the indefinite article, according to Hawkins (1978), is non-uniqueness. He claims that in cases where the object in question is unique within the speaker-hearer relevant shared set of objects, *a+singular count noun* cannot be used to refer to it and just in cases where the object or objects are not unique the indefinite article can be used. When a book, house or wedding are discussed we all know that there is one author to a book, one bride to a wedding and one roof to a house, thus, these objects are unique within the speaker-hearer shared sets, but there is more than one page to a book, more than one bridesmaid to a wedding and more than one window to a house. So if the definite article refers inclusively to all objects, then the indefinite article refers 'exclusively' to some only (Hawkins 1978:186).

As the definite article refers inclusively in both generic and non-generic reference, similarly the indefinite article refers exclusively in both cases. Singular generic indefinites differ from non-generic indefinites in that with non-generic indefinites the speaker may know which individual is being referred to or included and the hearer generally does not, whereas with generic indefinites both speaker and hearer do not have any particular included referent in mind. But both generic and non-generic indefinites include only one object in their reference.

The difference between a generic and non-generic indefinite is basically pragmatic, it shows whether the context leads the hearer to believe that the speaker either does or does not have a particular singular referent in mind. (Hawkins 1987:215)

To prove exclusiveness of singular indefinite generic reference in which it is similar to non-generic indefinite reference Hawkins (1987:215) suggests the following examples:

7. a) The lion is numerous in these parts.
- b) A lion is numerous in these parts.

c) Lions are numerous in these parts.

According to Master (1997) the indefinite article is the second most common way of describing a generic noun after the zero article.

2.1.3 Zero/ null article

The fact that the zero article has been neglected in much determiner research is "probably because it has neither phonologic, nor graphemic overt form" (Stephanides 1978:84).

70 years ago, Palmer (1939) suggested that there may be two types of zero article:

1. one for mass/non-count and plural nouns (e.g. milk and eggs);
2. one that occurs with certain singular count and proper nouns (e.g. lunch, London).

Chesterman (1991) proposes to refer to the first type as to the zero article and to the second type as to null article.

According to Master (1997:222) the zero article is "the most indefinite of the articles which general function is to remove the boundaries that make nouns discrete". For this reason, according to Master (1997), the zero article occurs most frequently with indefinite non-count (formless, continue) and plural count (limitless) nouns and it is the preferred means of describing generic nouns. Master also suggests the following functions of the zero article. (see Table 2)

Table 2. Functions of the zero article (cited from Master 1997:222)

Functions	Plural count	Non-count
First mention	<u>Men</u> are fools.	<u>Wine</u> was the topic that night.
General characteristics	<u>Snails</u> have shells.	<u>Mercury</u> is heavier than water.
Existential there	There are <u>holes</u> in your socks.	There is <u>lipstick</u> on his face.
Defining postmodification	<u>Cars</u> from Japan are reliable.	<u>Water</u> from a spring is pure.

Partitive of-phrases	We drank <u>gallons</u> of coffee.	Use <u>sulphate</u> of ammonia.
Intentional vagueness	<u>Capitals</u> of nations are rich.	<u>Research</u> on fusion is impressive.

Another interesting claim made by Master (1997) is that the zero article occurs with singular count nouns in alternation with indefinite *a* and renders a singular count noun to a non-count distinction. He exemplifies this alternation by four types of overlapping contrast between the zero article (\emptyset) and the indefinite article (*a*):

1. mass (\emptyset) - count (*a*)

Example: The boys ate chicken (vs The boys ate a chicken)

2. general (\emptyset) - particular (*a*)

Example: Mice like cheese (vs This is a cheese of uncommon flavor).

3. abstract (\emptyset)-concrete (*a*)

Example: They communicate by radio (vs They communicate by means of a radio).

4. "adjective" (\emptyset) - noun (*a*)

Example: He was man [i.e. manly] enough to accept his fate (vs He behaved like a man accepting his fate).(Master, 1997:222)

The null article (\emptyset), according to Master (1997:223), is the most definite of the articles and its general function is “to name a one-member set” (the members of a set comprise all the terms or entities associated with a word). As a result the null article occurs with singular count nouns in alternation with the definite article:

1. name(\emptyset) -description (*the*)

Example: Mr. Jones was appointed chairman (vs Mr. Jones was appointed the chairman).

2. familiar(\emptyset)-unfamiliar (*the*)

Example: After dinner, we'll see a movie (vs After the dinner, we'll see a movie).

2.2 Definiteness and specificity in Russian

According to Lambrecht (1994) and Lyons (1999) languages that do not have an article system

have semantic/pragmatic definiteness (i.e. topic markers, word order, classifiers). Since the Russian language has no overt article system a bare Determiner Phrase (DP) can be either definite or indefinite and specific or non-specific.

To mark definiteness Russian has demonstratives which may be used with previous-mention definites. The use of the demonstrative pronouns can be seen as a common way to express definiteness between the Russian and English languages. However, the Russian demonstratives *etot/tot* (this/that) more often fulfill the deictic function rather than mark an NP as definite.

Russian also has some lexical marking on indefinites. One of the indefinite specificity markers is the numeral *odin* (one) which has a special reading in an unstressed position. Both *odin*-marked and bare indefinites are possible but the first one brings more importance to the NP. The next example is taken from Ionin (2003: 108) to give a better understanding of what is meant by “brings more importance”.

8. a) *Ja pročitala včera odnu knigu. Ona byla očen' interesnaja!*
I read-PST yesterday one book-ACC. She was-FEM-SG very interesting!
“I read a (certain) book yesterday. It was very interesting!”
- b) *Ja pročitala knigu, a potom legla spat'.*
I read-PST book-ACC and then lay-PST sleep-INF.
“I read a book for a bit, and then went to sleep.” (Ionin 2003:108)

Since *odin* marks specificity it is logical to suppose that Russian speakers would treat *the* as the English equivalent of *odin* and therefore mark it as [+specific]. However, according to Ionin (2003) there is no real motivation to expect a mapping between *odin* and the definite article as *odin* is used only with indefinites. Moreover *odin* is only one of several lexical modifiers on Russian indefinites. On the other hand it is also possible to speculate that Russian speakers would associate the specificity marker *odin* with the indefinite article *a* and mark only specific indefinites with *a*. However, in the course of her research Ionin proves that this is not the case either.

2.2.1 Other lexical marking on indefinites in Russian

Two lexical items *kakoj-to* and *kakoj-nibud*, correspond to *some* in English. *Kakoj-nibud* is a narrow-scope quantifier, while *kakoj-to* is the wide-scope counterpart.

9. a) Lena xochet prochest' kakuju-to knigu.
Lena wants read-INF some book.ACC.
“Lena wants to read some book.”
- b) Lena xochet prochest' kakuju-nibud' knigu.
Lena wants read-INF some book-ACC.
“Lena wants to read some book or other.” (Ionin 2003:110)

Ionin (2003:111) notes that the presupposition that Russian speakers may map *kakoj-to* and *kakoj-nibud'* to *the* and *a*, respectively, is also false and as her research reveals there is no reason to believe that Russian speakers map any lexical modifier (*odin, kakoj-to, or kakoj-nibud'*) to a particular English article.

2.2.2 Word order, definiteness, and specificity in Russian

Definiteness and specificity are expressed to some extent by word order in Russian. According to Bailyn (1995) the preverbal position in Russian is associated with old information. Thus, if a bare DP subject is placed preverbally, it is interpreted as a definite. If the subject is placed postverbally, it is interpreted as an indefinite. In cases where the postverbal subject is modified it can be understood either as definite or indefinite, depending on the context.

10. a) Koshka vbezala v komnatu.
Cat-NOM ran into room.
“The cat ran into the room.”
- b) V komnatu vbezala koshka.

In room ran cat-NOM.
“Into the room ran a cat.”

c) V komnatu vbezala koshka moego soseda.

In room ran cat-NOM my neighbor-GEN
“Into the room ran a/the cat of my neighbor.” (Ionin 2003:112)

A bare object DP in a postverbal position can be interpreted either as definite or indefinite depending on the context.

11. Masha chitaet knigu.

Mary reads book-ACC

“Mary is reading the/a book. (Ionin 2003:112)

The conclusion Ionin draws is that “there is no one-way relationship between Russian word order and either definiteness or specificity thus the only generalization that can be made is that indefinites typically cannot appear preverbally unless they bear lexical modification.” (Ionin 2003:113).

Another way to mark a DP for definiteness or indefiniteness is through the case of the direct object. An accusative case in Russian is considered to be a definite object and the genitive case is known to be an indefinite object.

12. a) Daj mne den'gi.

Give me the money.

b) Daj mne deneg.

Give me some money. (Ionin 2003:113)

It is also possible to express definiteness through the use of the relative pronoun. Naylor (1983:208) says that if the theme is followed by a wh-word or a relative pronoun and a relative pronoun is in thematic position in the subordinate clause, the noun which was used in the theme is marked as definite.

13. Lakej i kucher, kotorye videli etih ludej, ne obratili na nih vnimanija.

The footman and the coachman, who saw the people, paid no attention to them. (Naylor 1983:208)

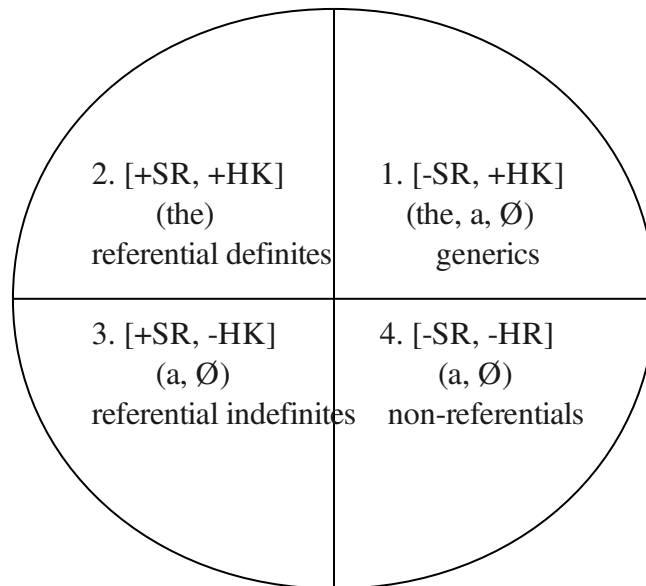
2.3 Definiteness and specificity in Chinese

This section on definiteness and indefiniteness in Chinese is included here since I will compare my results for the Russian learners with those for Chinese learners, as presented in Díez-Bedmar and Papp (2008). Díez-Bedmar and Papp (2008) base their description of Chinese on Robertson's (2000) work which as they write provides “an excellent summary of definiteness of Mandarin Nps.” (2008:153). Robertson (2000) characterizes Mandarin Chinese as a topic prominent language where definiteness and specificity are expressed by word order. This means that Mandarin Chinese speakers associate an NP in the preverbal topic position with the old information which is known by both speaker and hearer and thus is interpreted as a definite. Robertson's view on definiteness and specificity in Mandarin Chinese is that in addition to word order there are also the demonstratives *zhei* ('this') and *nei* ('that') and the numeral *yi* ('one'). However they do not mention the fact that Robertson (2000:147) has noted that “the Chinese language may be following a path which is prefigured in the development of English”. Li and Thompson (1981) point out that although Mandarin Chinese does not have the equivalents of the English articles *the* and *a*, “the demonstrative *nei* 'that' [...] is beginning to function as 'the' if it is not stressed, and the numeral *yi* 'one', if it is not stressed, is beginning to function as 'a'.” (Thompson 1981:131-2)

Snape, Leung and Ting (2006:138) argue that the universal cognitive/pragmatic category of identifiability is in the process of being grammaticalized onto the lexical items which could be treated as equivalents of the definite and indefinite articles in English (i.e., *yi* ‘one’= [-definite], *nei* ‘that’ = [+definite]) in Mandarin Chinese. They take the English facts as a point of comparison, where historically, the definite article derives from the demonstratives ‘this’ and ‘that’ and the indefinite article derives from ‘one’ (Hawkins, 2004).

2.4 Huebner's classification and the Fluctuation Hypothesis

Figure 1. Bickerton's (1981) semantic wheel for NP reference (from Huebner 1983)



Huebner's (1983) classification (which was originally based on Bickerton's (1981) semantic wheel model, see figure 1) is one of the most widely used models for the analysis of English noun phrase (NP) environments. In Huebner's model, the use of English articles is determined by two discourse features of referentiality. This system classifies each NP as plus or minus specific referent ([+/-SR]) and plus or minus assumed known to hearer ([+/-HK]). Based on these two aspects Huebner (1983) proposes four basic NP contexts that determine article use:

1. [-Specific Referent, +Assumed Known to the Hearer]: Generics
2. [+Specific Referent, +Assumed Known to the Hearer]: Referential Definites
 - a. Unique or conventionally assumed unique referent;
 - b. Referent physically present;
 - c. Referent previously mentioned in discourse;
 - d. Specific referent otherwise assumed common knowledge.
3. [+Specific Referent, -Assumed Known to the Hearer]: Referential Indefinites
First mention of NP [+SR] in a discourse and assumed not common knowledge.
4. [-Specific Referent, -Assumed Known to the Hearer]: Non-Referentials

- a. Equative noun phrases:
 - b. Noun phrases in the scope of negation;
 - c. Noun phrases in scope of questions, irrealis mode.
5. idioms/proper nouns (which are not incorporated into the classification)

According to Huebner's classification, if a specific referent is meant by the speaker depending on the knowledge of the hearer, the reference can be specific definite or specific indefinite. If no specific referent is meant based on the knowledge of the hearer the noun phrase either refers to all or most members of the class (+HK) or to no member at all, but to the notion of one (-HK). Proper nouns and idioms as well as some conventional uses are not incorporated into Huebner's classification as they can not be analyzed in accordance with four types outlined above (see Figure 1).

Ionin (2003:105) claims that all of the studies based on Huebner's framework look at definiteness and scope, but do not look at specificity as speaker intent to refer (as Ionin defines it). In her definition of specificity Ionin follows Fodor and Sag (1982) who argue that speaker's knowledge of the referent is a crucial determinant of specificity. Thus, according to Fodor and Sag (1982), a dog is [+specific] in (14a) and [-specific] in (14b).

14. a) A dog barked under my window all night – it was my neighbor's dog, a big German shepherd.
 b) A dog barked under my window all night – I wish I knew what dog it was!

(Ionin et al. 2004:19)

Consequently, Ionin and Wexler (2004) hypothesize that the overuse of *the* with indefinites in L2-English is due to the fact that L2-English learners use *the* to encode specificity. Based on such an assumption they predict that L2 English learners would overuse *the* with [+specific] but not with [-specific] indefinites.

Table 3. Article Grouping Cross-linguistically: Two-article Languages (cited from Ionin et al. 2004, p. 19)

	[+definite] (target: the)	[-definite] (target: a)
[+specific]	correct use of the	overuse of the
[-specific]	overuse of a	correct use of a

Based on the assumption that all of the learners have access to Universal Grammar (UG) they propose that L2 learners who belong to the [-Art] group will fluctuate between the two settings of the Article Choice Parameter (ACP) in the absence of L1 transfer:

Setting 1. Articles are distinguished on the basis of specificity.

Setting 2. Articles are distinguished on the basis of definiteness.

Their results prove, as was predicted, that both Russian and Korean speakers fluctuate between definiteness and specificity and thus the absence of L1 transfer.

However, Dirdal (2005:112-114) argues that the possibility of L1 transfer effect does exist since there are differences in percentages of article use between the Russian and Korean learners and variation in results obtained by means of the translation task and the gap task which cannot be explained by the ACP. Moreover, she stresses the fact that it is unclear why the ACP is based on referentiality and definiteness and not on de re/ de dicto distinctions if all the learners have access to UG.

A recent study by Hawkins et al. (2006) of intermediate L2 Japanese learners and a study by Reid et al. (2006) of intermediate Japanese and Spanish L2 learners, where they used the same forced elicitation task as Ionin et al. (2004), support the Fluctuation Hypothesis as Japanese learners were fluctuating between definiteness and specificity in both studies while Spanish L2 learners did not fluctuate as Spanish belongs to [+Article] group.

A similar study to that of Reid et al.'s (2006) was conducted by Ting (2005), which included 8 Mandarin Chinese and 5 Spanish L2 learners of English. The aim of her study was to check whether Chinese L2 learners would fluctuate between two article choice parameters as the Russian and Korean learners were previously reported by Ionin (2003). The results revealed that the Chinese, unlike the Japanese L2 learners, did not fluctuate between definiteness and specificity in [-definite, +specific] and [+definite, -specific] contexts. In fact, overuse of *a* was found in [+definite, +specific] contexts (17%) which was unexpected. The Spanish L2 learners did not fluctuate in any context as predicted.

In other words, not all of the learners who belong to the [-Art] group fluctuate between definiteness and specificity, again questioning the reliability of the ACP. Huebner's classification, on the other hand, can be used for both [+Art] and [-Art] groups of learners and thus allows

comparison between learners with various L1 backgrounds according to four semantic types of noun phrase.

2.3 Previous studies based on Huebner's classification

The first one to use a four type noun phrase classification was Huebner himself. In his study, Huebner (1983) investigated the use of the definite article *da* by an adult Hmong speaker with basic English skills. Based on his observations over one year, Huebner identified six stages in his subject's learning trajectory as the subject's proficiency was increasing. According to Huebner the second stage was marked by the definite article flooding and the six stage by strong association of the definite article with [+HK] Although Huebner did not conclude that reviled trajectory might be universal, his finding did provide evidence of systematic variability in article use with the increase of proficiency. This brings us to the conclusion that article acquisition for the [-Art] group of speakers is a process which can be divided into several stages in accordance with the change in their proficiency.

Parrish (1987), Master (1987), and Thomas (1989) as well have used Huebner's classification but unlike Huebner they investigated all three articles *the*, *a* and \emptyset . Master's (1987) pseudo-longitudinal study supports Huebner's finding that the [-Art] group would use *the* considerably in [+HK] contexts and *a* or \emptyset to a greater extend in [-HK] contexts.

In contrast to Huebner (1983) and Master (1987), Parrish (1987) and Thomas (1989) claim that *the* is initially associated with [+SR] rather than [+HK] contexts, they also did not find evidence for *the*-flooding but considerable overuse of the zero article by [-Art] rather than by [+Art] group in the definite and indefinite contexts. Thomas (1989) explains the fact of overgeneralization of the zero article as a result of L1 transfer. Like Thomas (1989), Master (1987) found that UOC (Used in Obligatory Contexts) was higher for the zero article in the [-Art] group which he suggests is due to the fact that [-Art] L1 speakers use zero article in their first languages, therefore zero can hardly be considered a use especially in the early stages of language acquisition. However, Parrish (1987) proposes that learners who belong to the [-Art] group acquire the zero article first, followed by the definite article, and finally the indefinite article

(\emptyset > the > a).

2.4 Acquisition orders revealed in previous studies for the [-Art] group of learners

Even though many researchers agree that the first article to be acquired by the [-Art] group of speakers is the zero article, the second is the definite article and the last is the indefinite article, their studies suggest variate and inconsistent acquisition orders as to proficiency levels and language backgrounds.

Master (1987), for example, conducted a cross-sectional study of 20 adult L2 learners from five native language backgrounds. His participants were drawn from three groups of [-Art] L1 speakers (Chinese, Japanese and Russian) and two groups of [+Art] L1 speakers (Spanish and German). In each language group four subjects represented four stages of interlanguage development: Baslang, Low-Mesolang, Mid-Mesolang, and High-Mesolang. As a proficiency measure Master used Cazden, Cancino, Rosansky, and Schumann's (1975) negation criterion. He analysed spontaneous speech which was elicited from each learner by means of informal interview. Based on the accuracy rates revealed in Master's study (1987) Fen-Chuan Lu (2001) presents the following acquisition orders for the [-Art] group of learners:

Baslang	$\emptyset > a > the$
Low-Mesolang	<i>the</i> > \emptyset > <i>a</i>
Mid-Mesolang,	$\emptyset > the > a$
High-Mesolang	$\emptyset > the > a$

Thomas (1989) did a cross-sectional study of 30 adult L2 learners from nine native language backgrounds divided into two groups: 23 subjects in the [-Art] group and 7 in the [+Art] group. She identified three proficiency levels for each of the groups: Low, Mid, and High and unlike Master (1987) she measured learners' proficiency by means of a one-hour in-house placement test and got the same results for all three proficiency levels for the [-Art] group:

Another cross-sectional study on article acquisition was conducted by Yamada and Matsuura (1982) who unlike Master (1987) and Thomas (1989) used as proficiency measure the division between high school and college Japanese students. Their study revealed the following sequences of acquisition:

Intermediate	<i>the > a > ∅</i>
Advanced	<i>the > ∅ > a</i>

Even though they used the same measure system (Supplied in Obligatory Contexts or SOC) differences in elicitation of the articles, measure of proficiency level plus different combinations of L1 speakers effected the results and caused inconsistencies of the acquisition orders. As we can see Thomas (1989) came up with the same acquisition order for all three levels while Master and Yamada & Matsuura present different acquisition orders depending on the proficiency level. Even though the acquisition orders are inconsistent with one another in terms of proficiency, all studies suggest that at some point learners begin to use the definite article more accurately than the zero article.

Fen-Chuan Lu (2001) in her cross-sectional study on article acquisition applied both SOC and TLU (Target Like Use). The aim of her study was to see what the three measure systems: SOC, TLU, and UOC², tell about acquisition of English articles. She used a multiple-choice cloze test³ as the testing instrument to elicit articles from the 55 Mandarin Chinese speakers who according to the TOEFL score were divided into three groups: the Advanced group; the Upper-Intermediate group; the Lower-Intermediate group.

² UOC (Under Over Suppliance) was devised by Master (1987) to be a complementary measure to observe the learner's overuse or underuse of the article:

$$\text{UOC} = \frac{\text{the total number of supplings in both obligatory and non-obligatory contexts}}{\text{number of obligatory contexts}}$$

³ Cloze test comprises 58 items in two parts: discrete sentences and a descriptive paragraph.

SOC showed the highest accuracy for the definite article in the Lower-Intermediate group and for the indefinite article in both Upper-Intermediate and Advanced groups. Zero article is the least accurately used in all three groups:

Advanced group	$a = the > \emptyset$
Upper-Intermediate group	$a > the > \emptyset$
Lower-Intermediate group	$the > a > \emptyset$

If previous studies based on the SOC measure agreed that the definite article is used more accurately than the indefinite article at the Intermediate and Advanced levels of proficiency Fen Chuan's study proposes that Chinese learners already at the Upper-Intermediate level cope better with the indefinite article than the definite article.

According to Fen-Chuan Lu (2001) TLU revealed the same hierarchy of accuracy for all three proficiency levels which is consistent with the Lower-Intermediate group acquisition order based on SOC.

Advanced	the > a > \emptyset
Upper-Intermediate	the > a > \emptyset
Lower-Intermediate	the > a > \emptyset

Since TLU is claimed to be a more reliable accuracy measure Fen-Chuan Lu concludes that the acquisition order revealed by TLU represents acquisition order for the Chinese learners. The same acquisition order was reported by Yamada & Matsuura for Japanese learners at the Intermediate level of proficiency based on the SOC measure.

Díez-Bedmar and Papp did a study of second course Chinese and Spanish students who according to the accuracy rates correspond to Chuan's Advanced group.

Table 4. Accuracy rates for the Chinese learners based on TLU and cited from Díez-Bedmar and Papp (2008:164) and Fen-Chuan Lu (2001:68)

	the	a	\emptyset
--	-----	---	-------------

Advanced (Fen-Chuan Lu)	76.88	70.02	62.17
Advanced (Díez-Bedmar and Papp)	67,4	73,18	75,94

They also used TLU in their calculations but instead of multiple close test, which Fen-Chuan Lu (2001) used to elicit the data, they analyzed written essays retrieved from the ICLE corpus. Their study unlike Chuan's revealed the following hierarchy of accuracy with the zero article being the most accurate:

Advanced $\emptyset > a > the$

However, such acquisition order was already suggested before for the [-Art] group of speakers (Chinese, Japanese and Russian) by Master (1987) at the Baslang level of proficiency based on the SOC measure. Thus, the indefinite article prior acquisition to the definite article is not that unexpected as Díez-Bedmar and Papp seem to think and it does not contradict but is inconsistent with the previous studies based on SOC as to proficiency level. However, their acquisition order for Chinese learners does contradict to that of Fen-Chuan Lu which was based on TLU.

Since inconsistency in article acquisition orders exists even among studies which have used Pica's measure system (TLU), there are other factors that influence article acquisition and which should be taken into account. For example, Brown (1983) explains article acquisition orders "variegation" by "the variety of methodological approaches and disparities in data-gathering procedures" (1983:25). Fen-Chuan Lu (2001:52-53) suggests that, apart from task effects, sample sizes and methodological approaches the fact that there is no shared placement standard for the participant's English proficiency can also contribute to inconsistencies in article acquisition orders since, in general, the acquisition orders are inconsistent with one another in terms of proficiency. Thomas (1989:53) warns that "direct comparison of L2 acquisition orders may not take into account linguistic idiosyncrasies underlying each language". Thus, if for instance Chinese and Spanish learners do acquire the indefinite article prior to the definite article it does not mean that such an acquisition order is true for English learners with other language backgrounds.

According to inconsistencies in article acquisition orders it is reasonable to assume that article acquisition is a process which can be divided into at least three stages based on the learners proficiency (beginning, intermediate, advanced). By defining article acquisition order at each of the three stages we would follow article acquisition process. Thus, Díez-Bedmar and Papp's acquisition orders for the Chinese and Spanish learners are relevant advanced level of proficiency only because it is not sure that acquisition order for Chinese learners is the same at the beginning and intermediate levels. Since I have chosen to study the first and second course Russian learners I will define acquisition order for the advance level of proficiency.

Chapter 3 Method

As mentioned before (see subsection 1.3) my study is a corpus based cross-sectional research on article use and acquisition. In order to compare article use and acquisition of English articles by Russian and Chinese learners I make use of what has been termed Contrastive Interlanguage Analysis (CIA). CIA was devised by Sylviane Granger in 1998 to identify the features common to all learners and the ones unique to learners with a particular L1 background. According to Granger (1998:47), “following the principles of CIA, a comparison between learners with different L1 backgrounds can be made”. She argues that this will enable us to distinguish “universal” errors from “L1-transfer” errors.

In order to conduct CIA of article use by Russian and Chinese learner I take the same variables as Díez-Bedmar and Papp in their study on article use by Chinese and Spanish learners. That is why I analyze essays written by second year university students comprising 40 000 words. More detailed information on the analyzed data is given in subsection 3.1.

Originally, I had also planned to apply Díez-Bedmar and Papp's tagging system and taxonomy of the English article system which they had based on Ekiert (2004) and adapted from Huebner (1985), Thomas (1989) and Goto Butler (2002). However, for the reasons explained in the following subsections I had to modify their taxonomy and tagging system. As a result, in subsection 3.2 I present a slightly modified tagging system; in 3.3 I explain in which cases articles will not be tagged and why; in 3.4 previous classification schemes based on Huebner (1983) will be discussed and the revised classification scheme will be presented.

3.1 The data

As the data was to be compared with that of Díez-Bedmar and Papp, I wanted to make sure that my data was as similar as possible to theirs as to the amount of words analyzed and proficiency of the learners. In their research Díez-Bedmar and Papp used 74 essays written by Chinese

learners comprising a total of 39,663 words and 101 essays written by Spanish learners, amounting to 39,881 words. As on average an essay of a Russian learner in the RICLE corpus was about 870 words, my data comprises 46 essays amounting to 39,982 words, written under the same external conditions as those of the Spanish and Chinese learners. The Russian learners had studied English for an average of 7-8 years at school and not more than 2 years at university. Therefore, they can be compared with the Chinese learners who have studied English for an average 7-8 years at school and at least one year at university, even though their proficiency levels still may differ as the proficiency level highly depends on the way the language was taught both at school and university.

Once the learner data was extracted it was corrected by three native speakers specifically looking at article use. Two of the native speakers were males and one female. They come from three different English-speaking countries: the USA, Australia and Great Britain. Díez-Bedmar and Papp do not specify how many native speakers were used to correct their data, thus, it is possible that they have used more than 3 native speakers, but unfortunately I did not have the possibility to get more people to look through the data.

There were some difficulties related to this process: not all the data could be corrected as there were some places where the native speakers were unable to correct the use of the articles as they could not understand the meaning of the sentence or some part of the sentence. Fortunately, there were not many places where the context prevented them from understanding which article should be used, however, all the uncorrected data had to be omitted. Here are some examples of sentences which were omitted at this stage of analysis.

15. [Who would refuse the man who dared to accept possibility of electricity a considerable amount of imagination?]

16. [Man knew a horse and he heard strange tales of eastern animals with horns on their noses, and so he created a unicorn.]

In both cases, the native speakers were unable to understand the meaning of the sentences and thus they were unable to decide which article should be used as article usage depends highly on the context; I therefore deleted all the instances of this kind. After all the uncorrected instances

were discarded I was left with 39,756 words for further analysis.

3.2 Tagging system

The next step was to tag all correct and incorrect uses of the articles in the data. For the reason of conducting a more “fine-grained” analysis, following Díez-Bedmar and Papp, all the instances of the articles were divided into four groups according to the for types of semantic contexts proposed by Bickerton (1981).

Díez-Bedmar and Papp name the four types in the same way as Huebner (1983) when they present the framework they use for article classification (see table 1, 2008:152). However, further in their analysis (see table 3, 2008:160) as the tags are devised they use slightly different definitions for the four types: Type 1. Generic, Type 2. Definite, Type 3. Specific Indefinite and Type 4. Non-Specific Indefinite. In order not to cause any misunderstandings I will use Huebner's definition for all four types in the course of my analysis as I find it more appropriate and less confusing.

When the article is used correctly by the learner, following Díez-Bedmar and Papp, I tagged it by means of the coding system outlined below (see table 5). The number indicates the type of the context to which an article belongs which is followed by the acronym of a correctly used article.

Table 5. Tagset for correct use of English articles in each semantic context

	Type 1 generics	Type 2 referential definites	Type 3 referential indefinites	Type 4 non-referential nouns
DA (definite article)	1DA	2DA		
IA (indefinite article)	1IA		3IA	4IA
ZA	1ZA		3ZA	4ZA

(zero article)				
----------------	--	--	--	--

When the article was used incorrectly by the learner, Díez-Bedmar and Papp marked it with the general tag “GA” followed by the article that would have been used by a native speaker in that specific context. Here is an example of how they used tags to mark the incorrect uses in texts:

17. It's a traditional view that only (1GAIA) man can carry the family. (ibid: 165)

In this example the general tag “GA” reflects that \emptyset man is an error of type 1, for which the number one stands, where a native speaker would have used the indefinite article which is represented by the acronym “IA” instead of the zero article which was used by a Chinese native speaker. I think that it is also important to have a tag for the article which was used by a learner. This will make it possible to register which of the articles was used by a learner instead of the target article and thus provide us with information as to which article Russian learners tend to use instead when the target article is *a*, for example.

As a result I ended up with sixteen error tags presented in the table below, where number indicates type of context which is followed by the general error tag “GA” to indicate an error in the use of the article system and finally the acronym for the target article which is followed by the acronym for the article used by the learner.

Table 6. Tags for incorrect use of English articles in each semantic context

Target article	Type 1 generics	Type 2 referential definites	Type 3 referential indefinites	Type 4 non-referential nouns
DA (definite article)	1GADA-IA	2GADA-IA		
	1GADA-ZA	2GADA-ZA		
IA	1GAIA-DA		3GAIA-ZA	4GAIA-ZA

(indefinite article)	1GAIA-ZA		3GAIA-DA	4GAIA-DA
ZA (zero article)	1GAZA-DA		3GAZA-IA	4GAZA-IA
	1GAZA-IA		3GAZA-DA	4GAZA-DA

Here is an example of how I used this tagging system to mark article use by the Russian learners.

The first tag shows correct usage of the zero article (1ZA) followed by four incorrect uses.

18. (1ZA) People who spent a few years in (1GAZA-IA) a prison also need (1GAZA-IA) a rehabilitation. They are out of (1GAZA-DA) the society, they can't find (1GAZA-IA) a work, other people don't accept them.

This reflects that *a prison*, *a rehabilitation*, *a work* and *the society* are incorrect uses of the first type, signalled by the general tag “GA”, and that in all four cases a native speaker would have used a zero article while a Russian learner used the indefinite article with *prison*, *rehabilitation* and *work*, and the definite article with *society* instead of the correct zero article. By stating which article was used by a Russian learner, in the example above we can see that Russian learners used the indefinite and definite articles where a native speaker would have used the zero article. For an example of a fully error-tagged RICLE text, see Appendix 2.

3.3 Untagged

The next step was to define what should be treated as other conventional uses which together with idiomatic expressions and proper nouns belong to type 5 and, as was mentioned in the theoretical part, will not be analyzed. Master (1987), for example, incorporates conventional expressions such as *go to school* into the [-SR, +HR] (generics) contexts while Huebner (1985) omits all commonly used expressions such as *in the morning*, *in the future* and *go home* which, as he notes, seem to be learned as “formulas”. In addition, he eliminates second and subsequent

noun phrases in a series, as in *the men and the women*. Noun phrases preceded by possessive pronouns are eliminated as well, because in such cases the [+ HK] [+ SR] distinction is neutralized. For example, we can say *my brother* without having mentioned him previously or without the hearer knowing whether the phrase *my brother* refers to a unique referent (my only brother) or to one of many brothers a person has. Parrish (1987), however, has adapted Huebner's system, accounting not only for the use of *the*, *a* and \emptyset , but also for the use of quantifiers and demonstratives since she claims that quantifiers can be used with Type 3 and Type 4 noun phrases and demonstratives can be used with Type 2 noun phrases, as in the examples she offers:

19. a) They don't have **any** children. (Type 4 NP)
b) I have **one** sister. (Type 3 NP)
c) Give me **that** book. (Type 2 NP) (Parrish 1987:26)

As in my research I am interested in how Russian native speakers cope with the English article system and as a result in the proportion of correct and incorrect uses of the articles I will exclude all noun phrases preceded by possessive pronouns as well as those preceded by quantifiers and demonstratives. I also decided to exclude such conventional uses as *in the X century*, *in the X year*, *in the morning (evening, future)*, *in the end*, *go home*, but I kept expressions such as *go to school*, *be at school/home/work*, *in prison* and other similar instances which, depending on the context, can be used with the definite, indefinite or zero articles. *A little bit*, *a lot of* and *a few* are another conventional expressions quite frequently used by the learners, which I decided to eliminate as they are extremely common and are rather subconscious uses.

All of the essays in my material had a title to which some of the learners referred back in their essays. Since the title coincided with the task set by the ICLE team and the students knew which articles should be used in this context I excluded such repetitions of the title together with all the quotations used by the learners; such instances do not depict the learner's choice of articles. I also decided to count such series of nouns as *society and technology* as one instance of article usage as well as the enumerations in which all the nouns belong to the same type.

20. Since the French revolution (1ZA) *equality, brotherhood and freedom* have been proclaimed the highest virtues of any society.

3.4 Classification scheme

Having decided what to include and what to leave out, I started classifying all the articles into four groups according to the taxonomy proposed by Díez-Bedmar and Papp (see Appendix 1) which, as was mentioned, they had based on Ekiert (2004) and adapted from Huebner (1985), Thomas (1989) and Goto Butler (2002). When it came to the classification of my data their classification scheme appeared to be a bit confusing as they did not give an example for every subcategory enumerated and the examples they used for each type were not clearly identified as belonging to a specific subcategory. I also had trouble understanding why *foreigners* in *Foreigners would come up with a better solution in this matter* is an example of type 4 and not of type 1 as, in my opinion, the reference is made to the foreigners in general. Another example of unclear classification was why they exclude *physically present referent*, which is on Huebner's list, from type 2 but still use an example *Pass me the pen* which seems to be an example of a physically present referent. As Díez-Bedmar and Papp based their classification scheme on that of Ekiert (2004) I first had a look at Ekiert's taxonomy of the English article system in order to get a better understanding of the confusing issues mentioned above. However, Ekiert's presentation of the taxonomy is almost identical apart from the fact that non-specific indefinites are not mentioned under type 4, thus I had to look through all the classification schemes referred to, starting with Huebner (1983).

As Huebner (1983) was the first one who suggested that noun phrases be classified in terms of four semantic categories, [\pm Specific Referent (\pm SR)] and [\pm Assumed Known to the Hearer (\pm HK)], his presentation of the classification scheme gives a clear understanding that type 3, which is named *referential indefinites*, consists of first-mention noun phrases and type 4, which is named *non-referentials*, consists of: a. Equative (or attributive) noun phrases; b. Noun

phrases in the scope of negation; c. Noun phrases in the scope of questions and irrealis mode.

Thomas (1989) is less specific and under type 4 enumerates attributive and non-specific indefinites and then adds “(etc)”. For non-specific indefinites (type 4) Thomas uses such examples as *I guess I should buy a new car* where *a new car* is a non-specific indefinite as any car can be meant.

Non-specific indefinite noun phrases frequently occur in so-called opaque contexts, where verbs like *believe*, *want*, *need* or *hope* introduce a description that may not correspond straightforwardly with anything in the real world. According to Lyons (1977:190-191) in the non-specific reading of the indefinite noun phrase it is actually uncertain if the indefinite noun phrase is referential or just descriptive (attributive).

21. John wants to marry *a girl with green eyes*. (Lyons 1977:190)

[...] the expression 'a girl with green eyes' can be construed as being used specifically or non-specifically. If it is taken as a referring expression (i.e., as having specific indefinite reference), then it presupposes, or implies, the existence of some individual who satisfies the description (...). If the indefinite noun-phrase 'a girl with green eyes' is construed as non-specific, there is no presupposition or implication of existence at all; and this is characteristic of descriptive noun-phrases (...) which occur after verbs denoting what (has been) called propositional attitudes. (1977: 190)

Such ambiguities can be solved when a sentence is looked at in context. Thus if *John wants to marry a girl with green eyes* is uttered having a definite referent in mind which is not known to the hearer, *a girl with green eyes* will be type 3 referential first-mention noun phrase, but if the same sentence is uttered to state that John does not know whom he wants to marry apart from that he would like her to have green eyes then *a girl with green eyes* is type 4 non-referential attributive indefinite as it is simply a description of a girl John wants to marry and is similar to that of *Alice is an accountant* where *an accountant* is simply a description of Alice.

The previously mentioned sentence *I guess I should buy a new car* offered by Thomas

(1989) as an example of a non-specific indefinite noun phrase can be also interpreted as a description of an object which does not correspond straightforwardly with anything in the real world and thus can be viewed as an attributive indefinite.

Goto Butler (2002:478-479) in her turn proposes a classification which is different from all the ones previously mentioned. Under type 2 she has exophoric, homophoric reference (*Pass me **the** pen.*), cataphoric reference (***The** idea of coming to the USA...*), anaphoric reference (*When I found a red box in front of my house, it was too late. **The** box blew up with a terrific explosion.*), connotative reference (*This book did not sell well even though **the** author was a famous writer.*), extended reference (*I won a million-dollar lottery. **The** news quickly spread all over town*) (all the examples are taken from Goto Butler's classification scheme p. 478-479). Definite noun phrases which Thomas defines as specified by definition Goto Butler (2002) defines as unexplanatory modifiers. She also defines type 1 as generics and unspecifiable: [a(n)], [the], [Ø], but she does not give an example of unspecifiable: [a(n)], [the], [Ø] and defines plural and non-count nouns as generic only if they are followed by a postpositional modifier.

22. Ø Language is a great invention of humankind. (Butler 2002:478)

That is why *foreigners would come up with a better idea in this matter* is an example of type 4 non-referentials as it does not satisfy the parameters outlined for type 1 by Goto Butler because *foreigners* is a plural noun which is not followed by a predicative phrase. However, according to Hasselgård et al. (1998:114), the plural indefinite form has the widest applicability among other generic forms. They consider *teenagers* and *cars* to have generic reference in the following two examples:

23. a) *Teenagers* like dancing and having fun.

b) *Cars* should be banned from the center of Oslo. (Hasselgård et al. 1998:115)

Thus, Hasselgård et al. (1998) would define *foreigners* in Butler's example of type 4 for non-referential noun phrases as a generic reference. They also suggest that uncountable nouns such as *life*, *nature* and *electricity* in the following examples have generic reference as uncountable nouns require the zero article to express general tendencies.

24. a) Life is short.
b) We must go back to nature.
c) What would we do without electricity? (Hasselgård et al. 1998:115)

Moreover, they note that uncountable nouns with premodification may still be conceived as generic: *computer technology*, *modern art*, *modern American fiction*, *organized crime*. Quirk et al. (1985) also mention that uncountable nouns with premodification (European history) have generic meaning but if an uncountable noun is postmodified *the history of Europe* it has a definite reference. Quirk et al. explain that cataphoric *the* is added before history because “the effect of the of-phrase is to single out a particular subclass of the phenomenon denoted by the noun, and there by to change the generic meaning into a specific or partitive one”. (Quirk et al. 1985:282)

Quirk et al. (1985:282) argue that the generic use of the zero article with both plural and noncount nouns identifies the class considered as an “UNDIFFERENTIATED WHOLE” and note that *mankind/humanity/man* in the meaning of “the human race”, *men/man/a man* in the meaning of “human being” are used generically.

25. *Hunger* and *violence* will continue to mark the future of *mankind/humanity*.
(Quirk et al. 1985: 282)

The form *all + noun* with no article usually has generic reference as well.

26. All men are created equal. (Quirk et al. 1985: 259)

In my definition of generic reference I followed Hasselgård et al. (1998) and Quirk et al. (1985) and thus such entities as *foreigners* in *foreigners would come up with a better idea in this matter*, which Díez-Bedmar & Papp placed under type 4, I have classified as type 1 (generic reference).

To get a better understanding of what the uses of the definite article are I consulted Quirk et al. (1985) and Hawkins (1978) to see which uses of the definite article they differentiate. Quirk et al. (1985) includes into *uses of the definite article*: immediate situation use, larger situation use, anaphoric reference:direct/indirect, cataphoric reference, sporadic reference (*my sister goes to **the** theater every month.*), the “logical” use of *the* (***the** first, **the** only*), the use of *the* with reference to body parts. In sporadic use of *the theatre* “refers, rather, to the theatre as an institution, so that it would be inappropriate to ask, in response: *Which theatre?*”(Quirk et al. 1985:269). This means that *the theatre* does not refer to any concrete theatre in the real world but rather to an institution in general. We can use in the same way *the cinema, television, the radio, the news, the press* and even *the train, the bus* when referring to aspects of mass communication and transport. Quirk et al. (1985:270) mention that “the sporadic use of *the* is in certain instances close to the generic use of *the*”. Those sporadic uses of *the* which refer to a concept in our mind or to an institution I have treated as generic reference as they are [-SR] but [+HK] and thus they fit into the parameters outlined for type 1.

27. Mary took *the bus/the train* or *a bus/a train* to London. (Quirk et al. 1985:270)

The fact that we can use both *the bus/train* and *a bus/train* and that *television* as a reference to mass communication is used with the zero article gives additional grounds to define such instances as generic.

Hawkins (1978) in his attempt to develop a theory of the referential meaning of the English noun phrase also differentiates a separate set of unfamiliar uses of the definite article.

Into unfamiliar uses of the definite article Hawkins (1978:131-157) includes referent establishing relative clauses (*What is wrong with Bill? Oh, **the** woman he went out with yesterday was nasty to him.*), associative clauses (***the** bottom of the sea, **the** cause of the war, **the** front page of the Guardian*), NP-complements (*Bill is amazed by **the** fact that there is so much life on earth.*), nominal modifiers (***The** number seven is my lucky number*), unexplanatory modifiers (***the** first person to sail to America was Icelander.*). Even though Butler (2002) includes into type 2 as a separate category only unexplanatory modifiers I think that all uses of *the* to which Hawkins refers as *unfamiliar uses of the definite article* can be presented as a separate category under type 2.

Since I analyze written data Butler's classification scheme is the most appropriate and matches my type of data best. Here is a slightly modified classification based on Butler (2002) which I will use for the classification of my data.

Type 1: [-SR, +HK], generic noun phrases: [a(n)], [the], [Ø]

Type 2: [+SR, +HK], referential definites: [the]

Exophora, homophora	Pass me <i>the pen</i> .
Cataphora	<i>The idea</i> of coming to the U.S. was. . . .
Anaphoric reference	When I found a red box in front of my house, it was too late. <i>The box</i> blew up with a terrific explosion.
Connotative reference	This book did not sell well even though <i>the author</i> was a famous writer.
Extended reference	I won a million-dollar lottery. <i>The news</i> quickly spread all over town.
Unfamiliar uses of <i>the</i>	<i>The first</i> person to jump into the cold water was my brother.
Unique in all contexts	There are nine planets traveling around <i>the sun</i> .

Type 3: [+SR, -HK], referential indefinites: [a(n)], [Ø]

first mention of referents
noun phrases following existential *have* and *there is/are*

I have bought *a new car*.
Our house has *a garage*.

Type 4: [-SR, -HK], nonreferentials: [a(n)], [Ø]

attributive noun phrases
noun phrases in the scope of negation
noun phrases in the scope of question
noun phrases in the scope of irrealis mood

Alice is *an accountant*.
I do not see *a pencil*.
Do you see *a pencil*?
If I had a million dollars, I'd buy *a big yacht*.

Type 5: Idioms, other conventional uses and numerical function of a: [a(n)], [the], [Ø]

Before I move further in my research I need to explain that I have included most existential constructions into type 3 as Thomas (1989) makes it clear from her discussion that she includes existential constructions under type 3 even though she does not mention them on her list. I also included noun phrases in scope of negation, question and irrealis mood into type 4, following Huebner (1983).

However, the fact that I have no knowledge of the immediate situation of students writing these essays, the HK becomes slightly problematic as the only thing I could go by for HK was context of the essays (anaphoric, cataphoric reference and a shared general knowledge of the world) I had to reconsider my classification for type 3 and 4 as not all noun phrases which followed existential constructions were referential when taken in a specific context.

28. (1ZA) Swedish prisons are very comfortable, (4ZA) cells look like (4ZA) hotel rooms, **there are** (4GAIA-ZA) TV, (4ZA) bookshelves, (4GAIA-ZA) toilet and (4GAIA-ZA) bathroom.

In the example above nouns which follow existential *there are* do not refer to any concrete objects as , for example, *there is a TV in my room* does. All such instances, which can be defined as non-

specific indefinites, will be analysed as type 4. That is why I decided to follow Butler (2004) and define type 3 simply as referential indefinites and type 4 as non-referential indefinites.

The reason why I felt I could not apply Díez-Bedmar and Papp's model as it was outlined in their article was because there were some issues which needed a more detailed specification to ensure the reliability of my findings. Nevertheless, I will compare my analysis of the Russian data with their findings for Chinese learners, such a comparison will reveal tendencies for Russian vs. Chinese learners, even if the classification of the data does not make 100 %.

Chapter 4 The Study

4.1 Scope

As was exemplified and discussed in section 2.6 previous studies on article accurate use and acquisition order are first of all inconsistent with one another in terms of proficiency levels. Since I analyze essays written by second year students, the acquisition order and accuracy of use for the Russian learners studied here will concern (refer to the) advanced level only.

4.2 Hypothesis

I hypothesize that accuracy rates for the definite and indefinite articles for Russian learners will be lower than those presented by Díez-Bedmar and Papp (2008:164 Table 7) for Chinese learners due to Chinese language specifics. As to acquisition order there are no reasons to expect the indefinite article to be acquired prior to the definite article by Russian learners.

4.3 Research questions

Based on the theoretical background presented earlier I have singled out the following research questions:

1. What do the SOC (Supplied in Obligatory Contexts) and TLU (Target Like Use) measure systems tell us about the use of English articles by Russian native speakers in terms of accurate use and acquisition order?
2. Which of the two systems is more reliable when describing acquisition order and why?

3. What is similar and what is different in article use by Russian and Chinese learners in terms of acquisition order, accuracy and overuse?
4. Do Russian and Chinese learners tend to associate the definite article with [+HK] or [+SR]?

4.4 Structure

In accordance with the research questions my analysis will be divided into four main parts: presentation of the accuracy rates and acquisition orders based on the SOC and TLU measure systems; evaluation of the measure systems in terms of reliability; comparison of acquisition order, article accuracy and overuse between Russian and Chinese learners; comparison of article use in four semantic types by Russian and Chinese learners (see section 2.4).

The subsections of the first part are dedicated to:

- definition of the article acquisition order based on SOC; (subsection 4.6.1)
- definition of the article acquisition order based on TLU; (subsection 4.6.4)

The second part of the analysis is dedicated to the discussion of the two measure systems with regard to accuracy rates and their reliability.

The third part deals with the comparison of article use by Russian and Chinese learners in terms of:

- accuracy rates revealed by TLU;(subsection 4.8.2)
- misuse rates revealed by TLU;(subsection 4.8.3)

The last part of the analysis is an attempt to explain how the use of the articles depends on the speaker/listener's common knowledge and the shared situation of utterance (see Table 1). The reasons why Russian and Chinese learners use articles differently in four semantic types will be discussed with regard to the ways definiteness and specificity is realized in the two languages and the way the English article system is taught.

Finally, some conclusions will be presented

Note: As noted in Chapter 2 a fuller account of how acquisition rates can be measured will be given before the results are presented.

4.5 Measure systems

Article acquisition order is defined by the accuracy rates of the articles. The article which is used most accurately is assumed to be acquired first, the one which stands second in the accuracy hierarchy is acquired next and, finally, the one which is used least accurately is acquired last. As mentioned there are two measure systems which can be applied for article accuracy rate estimation: SOC (Supplied in Obligatory Contexts) and TLU (Target Like Use).

Most researchers (Dulay & Burt, 1973, 1974; Bailey, Madden, & Krashen, 1974; Larsen-Freeman, 1975; Andersen, 1976, 1977, 1978; Hakuta, 1976; Master, 1987; Parrish, 1987; Thomas 1989) have used the SOC measure to estimate accuracy rates of articles.

$$\text{SOC} = \frac{\text{number of correct suppliance in obligatory contexts}}{\text{number of obligatory contexts}}$$

When applying the SOC measure system we estimate how accurately an article is used relative to its obligatory contexts, i. e. contexts in which a native speaker would use it. This explains, perhaps, why many of the studies based on the SOC measure system propose that the zero article is used most accurately at the beginning level of English learning and thus is acquired first by the [-Art] group of learners. However, some linguists (Master, 1997; Thomas 1989) report that learners who belong to this group do not just use the zero article most accurately but also overuse

it markedly in non-obligatory contexts, i. e. contexts where another article would be used. In order to redress this potential inflation of SOC Pica (1984) designed a slightly different measure system which is known as TLU:

$$\text{TLU} = \frac{\text{number of correct suppliance in obligatory contexts}}{(\text{number of obligatory contexts}) + (\text{number of supliances in non-obligatory contexts})}$$

This measure system estimates article accuracy relative to the number of obligatory contexts plus the number of non-obligatory contexts (overuse). As a result the zero article at the beginning level would probably be less accurate if its overuse was taken into account in accuracy estimation.

However, it does not necessarily mean that accuracy numbers revealed in previous studies are incorrect or unreliable. It is important to understand that the main difference between SOC and TLU is that SOC depicts article accuracy based on article use in obligatory contexts while TLU presents correct article relative to article use in obligatory and non-obligatory contexts.

I apply both measure systems to estimate article accuracy for Russian learners as it seems unclear why Master (1987), Parrish (1987), Thomas (1989) would chose SOC over TLU which was devised by Pica already in 1984 and is claimed to be more reliable.

4.6 Results and discussion

The following section is dedicated to the presentation of the findings obtained based on both measure systems: Supplied in Obligatory Contexts and Target Like Use. First I present acquisition order for Russian learners based on the SOC measure system and discuss its consistence and inconsistency with previous findings. Then I calculate percentages of article

misuse in definite, indefinite and zero contexts and article overuse in non-obligatory contexts and based on these findings make some general comments on article use by Russian learners. Finally, accuracy hierarchy and thus acquisition order revealed by the TLU measure system will be reported and discussed.

4.6.1 Article acquisition order for Russian learners based on SOC

Following the previously presented SOC formula I have counted the number of correct article uses in obligatory contexts relative to the sum of all obligatory contexts (see section 4.5) and as a result obtained accuracy rates presented in Table 7.

Table 7. Accuracy rates obtained by the SOC measure system for the definite, indefinite and zero articles

Articles	Percentages
the	92.1
a	83.6
Ø	88.4

As we can see from Table 7 according to the rates revealed by the SOC measure system the definite article is used correctly in 92.1 percent and thus is the most accurately used article relative to obligatory contexts. The definite article is followed by the zero article which is used correctly in 88.4 percent of obligatory contexts, while the indefinite article is used correctly in 83.6 percent and thus causes most difficulties for the Russian learners. Example 29 illustrates

correct uses of the indefinite, definite and zero articles in my material.

29. It is widely recognised now that (4IA) a country could be considered civilised only if (2DA) the equality of all (1ZA) human beings translated into (1ZA) social relationships clearly implies that all (1ZA) men should have (1ZA) equal opportunities.

According to the hierarchy of accuracy revealed by SOC, acquisition order for Russian learners at the advanced stage of proficiency is the following:

the > Ø > a

The same acquisition order was reported by Yamada & Matsuura (1982) for the Advanced group of Japanese learners. Their Advanced group was represented by college students and thus their subjects' proficiency corresponds to the first and second academic year, which suggests that Russian and Japanese learners acquire articles in the same order.

This acquisition order is also consistent with the acquisition sequence reported by Master (1987) for the [-Art] group of learners but it corresponds to the Low-Mesolang group (see section 2.6), probably because Master applied a different proficiency measure and each group consisted of learners with three different L1 backgrounds (Russian, Chinese and Japanese). Even though the revealed acquisition order does not correspond to that of Master as to proficiency level it still supports my finding.

Thomas' acquisition order (1989), which is the same for all three levels of proficiency (Low, Mid, High), supports my findings as well. Thus, according to the SOC measure system the definite article is used most accurately and is closest to the native like use at the advanced level of proficiency, while the indefinite article is the one that is acquired last.

4.6.2 Misuse

Since in the tag for incorrect article use I have included not only the target article but also the one which was used incorrectly by the learner, I am able to count percentages of article misuse applying the same formula which I have used for the estimation of accuracy. To estimate article misuse I have counted the number of incorrect article uses in obligatory contexts relative to the sum of all obligatory contexts and as a result obtained accuracy rates presented in Table 8.

$$\text{SOC} = \frac{\text{number of incorrect suppliance in obligatory contexts}}{\text{number of obligatory contexts}}$$

Table 8. Misuse percentages of the definite, indefinite and zero articles based on SOC

Articles	Percentages		
the	7.9	a	3.4
		∅	4.5
a	16.4	the	7.1
		∅	9.2
∅	11.6	the	9.7
		a	1.9

According to Table 8 the definite article is misused in more or less the same proportion for the zero and indefinite articles. The indefinite article is misused a little bit more for the zero article than for the definite article, while the zero article is being misused almost exclusively for the definite article.

It is a bit surprising that Russian learners use the definite article instead of the zero article five times more often than the indefinite article. I had a look at those contexts where the definite article was used instead of the target zero article and came to the conclusion that Russian learners

have trouble understanding how to use articles with abstract nouns. Most of the cases where Russian learners chose to use the definite article were when it was used with such abstract nouns as *mankind, society, labour, service, growth, entertainment* (example 30).

- 30.** If (1GAZA-DA) the slave labour were effective, Russia would have (2DA) the highest living standards in (2DA) the world., and (2DA) the communist regime would have prospered.

Russian learners at the advanced level of English proficiency underuse the indefinite article twice more than the definite article. The indefinite article is often underused by Russian learners in such partitive of-phrases as *a part of, a style of, a source of, a sort of* (example 31); with non-referential nouns (example 32); abstract nouns (example 33).

- 31.** Kuwait was never (4GAIA-ZA) part of the USA, it is (4GAIA-ZA) part of the world.
- 32.** If we take (4GAIA-ZA) TV, for example I don't even have one.
- 33.** And (1ZA) films are full of (1ZA) cruelty and violence, (2DA) the presentation of which on (1ZA) television might lead (1ZA) young people to regard (1ZA) violence as (4GAIA-ZA) acceptable behaviour, which should never happen.

I have noticed two peculiarities in incorrect definite article use in indefinite contexts. Russian learners once again have difficulties with abstract nouns (example 34).

- 34.** What we have today is (4GAIA-DA) the environmental crisis which becomes more obvious.
- 35.** When one wants to read (4IA) a book, he usually chooses it himself. When one wants to go to (4GAIA-DA) the exhibition, he also chooses it himself, and appreciating (2DA) the works of (1ZA) art he should also make certain

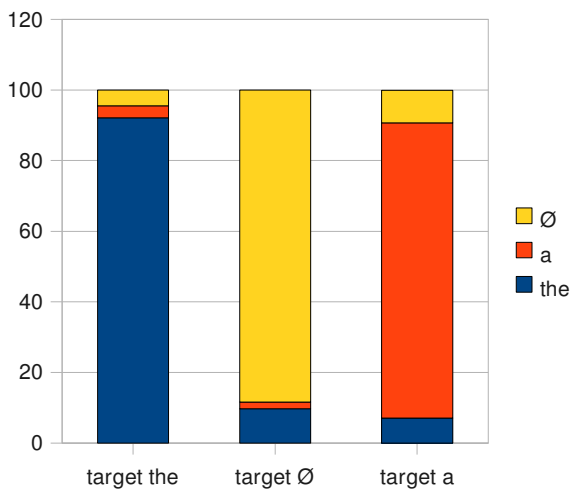
efforts and try to reveal its sense, to comprehend it; when one wants to go to (4GAIA-DA) the concert, he also makes (2GADA-ZA) choice himself

However an even more typical mistake made by Russian learners is misuse of the indefinite article for the definite article with non-referential nouns (example 35). In Example 35 no specific *book*, *exhibition* or *concert* is meant but a concept of one: a set of written, printed, or blank sheets bound together into a volume; a public showing (as of works of art, objects of manufacture, or athletic skill); a public performance (as of music or dancing) (<http://www.merriam-webster.com>). However, a learner used the indefinite article with book and the definite article with exhibition and concert. Again *book* is a concrete noun while exhibition and concert are abstract nouns.

Consequently, Russian learners overuse the definite article in zero and indefinite contexts because they seem to be confused by the fact that abstract nouns are concepts and not objects. They tend to use the definite article with such abstract nouns as *society* or *exhibition*, probably because nouns that name an idea or concept are assumed to be known to both hearer and speaker since everyone knows what *society* and *exhibition* stand for.

Figure 2 gives a better visual presentation of the misuse percentages presented in Table 8 along with accuracy rates from Table 7.

Figure 2. Percentages of the article accurate use and misuse in obligatory contexts for each of the articles



In Figure 2 each column represents obligatory contexts for each of the articles and the colors represent which article has been used.

When we look at the obligatory uses of the definite article we see that in those 7.9 percent of the cases when the definite article is used incorrectly it is misused a bit more for the zero article than for the indefinite article (examples 36 and 37). This means that even though Russian learners use the definite article most correctly relative to its obligatory contexts they still both misuse it for the indefinite article and underuse it since zero article signals non-suppliance of the target article. Consequently, Russian learners have both grammatical and pragmatic deficit when using the definite article. By grammatical deficit I mean article underuse and while pragmatic deficit stands for article overuse and misuse.

36. **(2GADA-ZA)** Russian Mafia is, of course, less impressive than Stalin now but **(1ZA)** commercial crimes and, especially, **(1ZA)** business fraud have hit their highest level last year and continue to increase now.
37. Apart from what has been said above, it is **(2GADA-IA)** a competition between **(1ZA)** national economies that matters.

As we move to the zero article it is misused more than the definite article since it is used less accurately in obligatory contexts. As mentioned the zero article is misused almost exclusively for the definite article as blue color represents the definite article in zero article contexts and reaches up to ten percent while red color, which represents the indefinite article, comprises about two percent of the obligatory uses of the zero article (examples 38 and 39).

It is hard to comment on the grammatical deficit of the zero article as it is impossible to define when the zero article was used correctly consciously and when it was used correctly unconsciously in my material. As a result the accuracy of the zero article still can be overestimated as it is not sure that all of the correct uses of the zero article are conscious. Consequently, when analyzing the use of the zero article we look only at pragmatic deficit as grammatical deficit is not taken into account.

38. (1ZA) European problems seem to be very local in comparison to (1GAZA-DA) the Russian organised criminality.
39. But (2DA) the best way to prevent (4IA) a new crime, she believes, is to put this person on (1GAZA-IA) a probation.

The indefinite article which is used least accurately in its obligatory contexts is as a result misused most. Since the indefinite article misuse for the zero article comprises almost ten percent of its obligatory uses it is clear that the indefinite article unlike the definite article is still significantly underused (examples 40 and 41).

40. In most countries (1GAZA-DA) the military service is (4GAIA-ZA) rather honoured and prestigious business .
41. There are some ideas, which are intended to modify (2DA) the political structures of (2DA) the EU to turn it into (4GAIA-ZA) federal and more democratic union.

4.6.3 Overuse

Since we know already how accurately articles were used in obligatory contexts it will be easier to comment on the reliability of TLU if overuse in non-obligatory contexts will be estimated as well.

However, I was not sure how to estimate article overuse in non-obligatory contexts. According to SOC, we estimate article accuracy in obligatory contexts by dividing the number of correct uses in obligatory contexts by the number of all obligatory contexts (see section 4.5). Consequently, we estimate article overuse in non-obligatory contexts by dividing the number of incorrect uses in non-obligatory contexts by the number of all non-obligatory contexts. Thus the formula to estimate article overuse in non-obligatory contexts (Supplied in Non-Obligatory

Contexts) is:

$$\text{SNOC} = \frac{\text{number of incorrect suppliance in non-obligatory contexts}}{\text{number of non-obligatory contexts}}$$

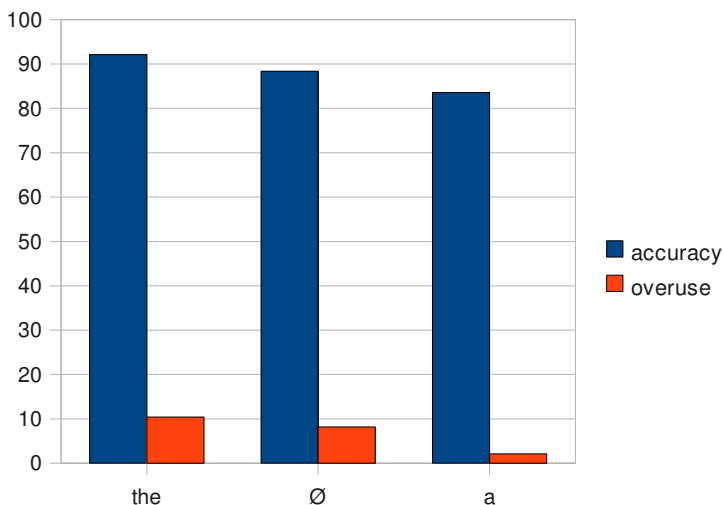
Following the presented SNOC formula I have counted the number of incorrect article uses in non-obligatory contexts relative to the sum of all non-obligatory contexts and as a result obtained overuse rates presented in Table 9.

Table 9. Overuse percentages of the definite, indefinite and zero articles in non-obligatory contexts based on the SNOC measure system

Articles	Percentages
the	10.4
a	2.1
Ø	8.2

The next figure depicts accuracy percentages together with overuse percentages of the articles. As mentioned above overuse rates were calculated based on the SNOC which was devised from SOC.

Figure 3. Percentages of the accuracy and overuse rates of the articles



Now when we have accuracy and overuse for each of the articles it is obvious that the definite article is not just used most accurately in its obligatory contexts but that it is overused most in non-obligatory contexts. There seems to be a correlation between article accuracy and overuse. The more accurately an article is used in obligatory contexts the more it is overused in non-obligatory contexts.

Consequently, the definite article is not just used most accurately in obligatory contexts but it is also overused most. Does it mean that its accuracy was overestimated by SOC as it was claimed by Pica (1984)? Apparently Pica was right because we cannot assume the definite article to be the most accurately used by Russian learners when they are not able to restrict its uses to obligatory contexts.

4.6.4 Acquisition order based on the TLU measure system

Following the calculation formula proposed by Pica (1984) (see section 4.5) I have counted the number of correct uses in obligatory contexts relative to the sum of obligatory and non-obligatory contexts. Table 10 presents article accuracy rates revealed by the TLU measure system.

Table 10. Accuracy percentages for the definite, indefinite and zero articles based on the TLU measure system

Articles	Percentages
the	79.1
a	77.3
Ø	82.3

According to the accuracy rates presented in Table 10 after article overuse was taken into account it appears that Russian learners use most accurately the zero article, followed by the definite article while the indefinite article stands last in acquisition sequence, therefore, the acquisition order is the following:

Ø > the > a

This acquisition order supports the one revealed by SOC (the > Ø > a) in that Russian learners acquire the definite article prior to the indefinite article, but contradicts it as to the zero article acquisition order. On the other hand, according to the TLU measure system, the zero article is the one which is used most accurately, which confirms Díez-Bedmar and Papp's findings as to the zero article's prior acquisition. One thing which is perfectly clear is that Russian learners, unlike Chinese learners, acquire the indefinite article last. The only question is which one do they acquire first? So was Pica (1984) right and does her measure system depict a more reliable acquisition order?

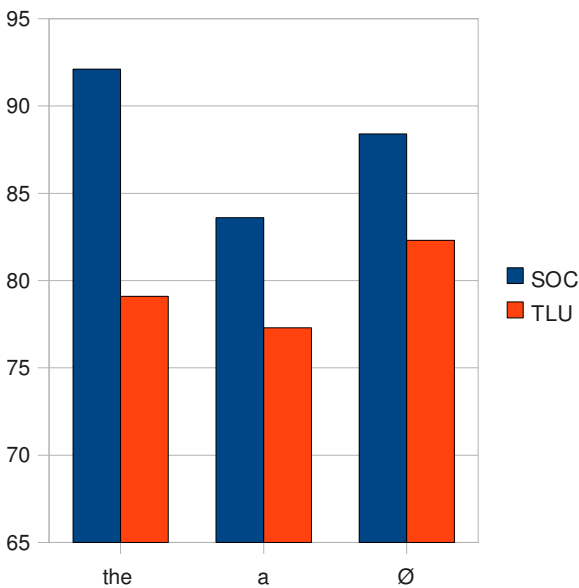
4.7 Which of the two acquisition orders is the more reliable one?

As was explained before both SOC and TLU depict article accuracy: SOC shows how accurately

Russian learners use articles relative to obligatory contexts while TLU shows how many of the instances of the articles supplied are used correctly relative to obligatory and non-obligatory contexts together. Previously I have presented evidence supporting both acquisition orders but one of them depicts article accuracy more precisely. After TLU was applied it is not just the acquisition order that has changed, but also accuracy rates for each of the articles became lower.

Figure 4 depicts the change in accuracy rates for each of the articles after overuse was taken into account.

Figure 4. Accuracy rates for each of the articles calculated based on the SOC and TLU measure systems



An interesting tendency is revealed as soon as we calculate accuracy relative to obligatory and non-obligatory contexts together. First of all, accuracy for all three of the articles drops when overuse of the articles is considered. Secondly, accuracy drops more for the definite article than for the zero or indefinite articles. Due to the drop in the accuracy rate for the definite article the zero article becomes the most accurate with the accuracy rate 82.3 percent and the definite article as a result becomes the second in the accuracy hierarchy (79.1%). The indefinite article is still the least accurate with a rate of 77.3 percent.

The drop in the accuracy rate for the definite article is more significant than for the zero and indefinite articles as a result of the higher overuse of the definite article in non-obligatory contexts (see Figure 3). As a result the definite article loses almost 10 percent from the previous accuracy rate while the zero and the indefinite articles lose about 6 percent each. Even though the zero article was overused twice more than the indefinite article it is interesting to observe that they both dropped about 6 percent.

Since it was stated that the definite article is overused most in non-obligatory contexts the change in acquisition order is caused by the overuse which before was not regarded as a factor which defines article accuracy as well. To show how important article overuse can be in estimation of accuracy I propose to have a look at statistical data. Let us concentrate on the total amount of instances supplied for each of the articles and estimate how many of those were used correctly.

number of correct supliances

total number of supliances

Thus, accuracy rates presented in Table 11 provide us with information on how many of the articles were used correctly relative to the total number of instances supplied by the Russian learners.

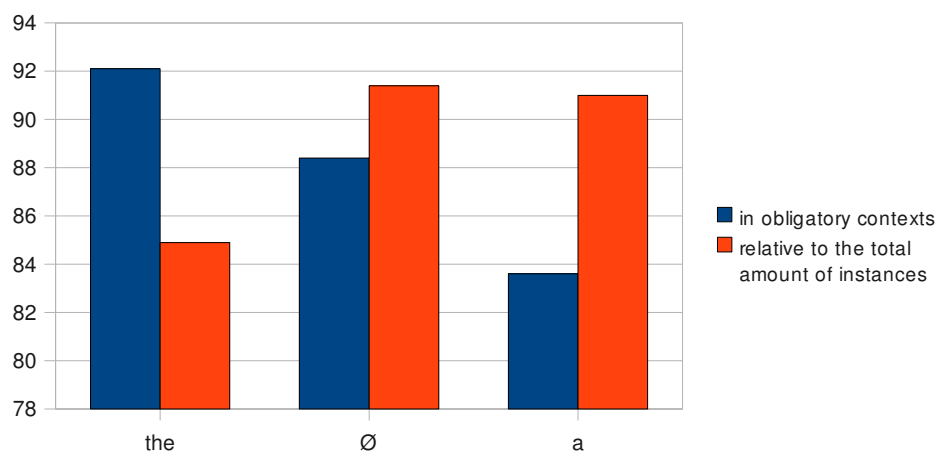
Table 11. Percentages of accuracy for the definite, indefinite and zero articles relative to the number of the instances of the definite, indefinite and zero articles supplied in the data

articles	Percentages of accuracy
the	84.9
a	91

∅	91.4
---	------

Comparing these accuracy rates with those based on SOC (see Figure 3) we can see that Russian learners used the definite article in 92.1 percent of the cases where the definite article is appropriate but only 84.9 percent of all the instances supplied by the Russian learners were correct. They used the indefinite article in 83.6 percent of the cases where it should be used, however, 91 percent of all instances of the indefinite article supplied by Russian learners were correct. As for the zero article Russian learners have used it in 88.4 percent of the appropriate contexts but 91.4 percent of all uses of the zero article in the data were correct. Figure 5 gives a visual picture of the just commented tendencies.

Figure 5. Accuracy percentages for the definite, indefinite and zero articles in obligatory contexts and relative to the total amount of instances supplied for each of the articles



The fact that so many uses of the definite article were incorrect can be a proof of the definite article flooding. This term was first used to describe article overuse by Huebner (1983) who was investigating definite article acquisition and based on his observations over one year identified six stages in his subject's learning trajectory as the subject's proficiency was increasing. According to Huebner second stage in definite article acquisition was marked by a phenomenon

which he named *flooding* and which now is called overuse.

It is interesting that the indefinite article is used less accurately in obligatory contexts than the zero article but equally accurately when it comes to the total amount of the instances supplied for each of the articles. Since this accuracy rate present statistical data we should not forget that indefinite article is the one supplied least accurately due to the low accuracy in obligatory contexts and thus is under represented. Consequently, the indefinite article is used correctly in 91 percent of the cases because of its lower frequency.

Thus, Russian learners at the advanced stage of proficiency use the definite article most accurately in obligatory contexts but at the same time they overuse it most and that is why only 84.9 percent of all the instances were correct. The zero article is used less accurately in obligatory contexts but since it is also less overused it was used correctly in 91.4 percent of all the instances of the zero article. Since a higher percentage of the zero article uses is correct it supports my finding that the zero article is used more accurately.

The indefinite article which is used least accurately in obligatory contexts is also overused least and due to the low overuse rate it is used correctly in 91 percent of all the instances of the indefinite article.

Based on the data obtained it will be reasonable to suppose that Russian learners are going through a stage of definite article flooding as only 84.9 of all its uses are correct while its accuracy in obligatory contexts is 92.1 percent. What again supports TLU that the definite article is acquired second.

4.7.1 Conclusions

TLU is more reliable for the definition of the acquisition order as it takes into account both correct and incorrect uses of the article and as a result depicts true accuracy rates. Thus, the acquisition order for the Russian learners at the advanced level of proficiency is:

∅ > the > a

Since Chinese and Russian learners are representatives of the [-Art] group of learners the fact that they use the zero article most accurately gives grounds to conclude that the zero article is acquired first. However, following Díez-Bedmar and Papp, there is still uncertainty as to how conscious those correct uses of the zero article are.

Chinese learners seem to use the zero article most correctly, as has already been stated (Master 1997:216), but it is difficult to ascertain whether this is a correct use of the zero/null article or non-suppliance, i.e. underuse, of any articles in those contexts. (2008:164)

However, the fact that learners have been studying the English language and thus its article system for 8 years at school and two years at the university may suggest that this is a conscious choice.

To my mind, since TLU is a more accurate system for article accuracy estimation it is possible to conclude that the zero article at the advanced level of proficiency is used more accurately since its oversuppliance was taken into account as well. Conclusively, both Chinese and Russian learners do use the zero article more correctly relative to the other two at the advanced level of proficiency. It is also possible to speculate that the zero article is used most accurately by all the representatives of the [-Art] group at the advanced level of proficiency. We should not forget however that in obligatory contexts it is the definite article which is used more correctly and that the zero article stands first in acquisition sequence because of the high definite overuse.

4.8 Comparison of article use by Russian and Chinese learners

This section is dedicated to the comparison of my findings to those of Díez-Bedmar and Papp, i. e. article use by Russian vs. Chinese learners. I will start with a discussion of the article acquisition orders and accuracy rates for Russian and Chinese learners and the reasons why

Chinese learners acquire the indefinite article prior to the definite article. Then, following Díez-Bedmar and Papp, I calculate misuse percentages in each article context for Russian learners and compare them to misuse percentages for the Chinese learners.

4.8.1 Inconsistencies of the article acquisition orders of Russian and Chinese learners

According to Díez-Bedmar and Papp's (2008:167) findings, both Chinese and Spanish learners acquire the indefinite article prior to the definite article.

in contradiction to previous studies, the indefinite article seems to be less difficult to acquire, and it is rather the definite article that presents a problem, even at an advanced stage of acquisition, to both Chinese and Spanish learners. (Díez-Bedmar and Papp 2008:167)

Since Chinese learners, similar to Spanish learners use the indefinite article more accurately than the definite article, while Russian learners seem to follow previously proposed acquisition order (the>a) this similarity between Chinese and Spanish learners can be explained as a result of Chinese language specifics.

As I have mentioned in the theoretical part, Snape et al. (2006) claim that the Chinese language is on its way to grammaticalization of the universal cognitive/pragmatic category of identifiability onto the lexical items which could be treated as equivalents of the definite and indefinite articles in English. Since previous studies focusing on cross-linguistic differences in the acquisition of the articles revealed that learners whose first languages contained an article system differ in English article acquisition from those whose first languages do not contain such a system, showing that English article usage is clearly influenced by the first language. This gives us grounds to think that if Chinese is on its way to grammaticalization of the universal cognitive/pragmatic category of identifiability onto the lexical items it would explain why Russian learners did not perform better for the indefinite article than for the definite article, while the Chinese did.

However, this very likable explanation is groundless as Spanish learners belong to the [+Art] group and thus as Díez-Bedmar and Papp mention themselves, need to adjust their article

system to the English one.

[...] we can conclude that there is one major difference in the use of articles in English and Spanish: the zero article can not be used to express generic reference in Spanish. Apart from this, some minor differences can also be noticed in contexts where Spanish does not require any article and English requires an indefinite one. (Díez-Bedmar and Papp 2008:155)

This means that they would not undergo the same acquisition process as Russian or Chinese learners. Moreover, Díez-Bedmar and Papp, based on the inconsistencies between the Spanish and English article systems, hypothesize that Spanish learners will mainly have difficulties using the zero article to convey generic meaning and in those cases where English requires an indefinite article and Spanish none. Their hypothesis was right as the zero article is the least accurately used by Spanish learners. It seems reasonable to assume that Spanish learners use the indefinite article most accurately because its uses are most common in English and Spanish. Consequently, French learners of English can use most accurately another article depending on consistencies and inconsistencies of the two languages.

As to the Chinese learners, based on the lack of an article system in the Chinese language, Díez-Bedmar and Papp hypothesize that:

[...] Chinese speakers will underuse the definite and indefinite articles in English, overuse the zero article, and even misuse them until they manage to learn to restrict the use of them to the specific semantic contexts and pragmatic functions in which they are employed in English. (Díez-Bedmar and Papp 2008:154)

Thus Chinese and Spanish learners are supposed to undergo two different processes on their way to article acquisition and yet Díez-Bedmar and Papp claim that since both Chinese and Spanish learners cope better with the indefinite article it is the definite article which causes most problems for the learners of English. My results did not support their hypothesis as Russian learners use the definite article more accurately than the indefinite article.

Another possible reason for the higher accuracy of the indefinite article by Chinese learners is a significantly high overuse of the definite article in non-obligatory contexts. As it was

proofed in the present study on the example of Russian learners article overuse in non-obligatory contexts influences article accuracy greatly. Since Russian learners use the definite article less accurately than the zero article due to the definite article overuse in non-obligatory contexts it is possible to hypothesize that Chinese learners use the definite article least accurately as a result of the even higher definite article overuse in non-obligatory contexts.

4.8.2 Comparison of Russian learners to Chinese and Spanish according to the accuracy rates revealed in the studies based on TLU

As was mentioned before, accuracy rates based on TLU depict how accurately articles were used relative to both obligatory and non-obligatory contexts. Table 14 is based on the accuracy rates revealed in Díez-Bedmar and Papp's study (Díez-Bedmar and Papp 2008, p. 164 see Table 7 and 8) and on the accuracy rates obtained in the course of the present study. I would like to remind that I hypothesized that Chinese learners would use the definite and indefinite articles more accurately than Russian learners due to the Chinese language specifics (see section 4.2). Since Russian learners use articles more accurately than Chinese learners, in contradiction to my hypothesis, I have also included accuracy rates obtained for Spanish learners to compare with Russian learners' accuracy.

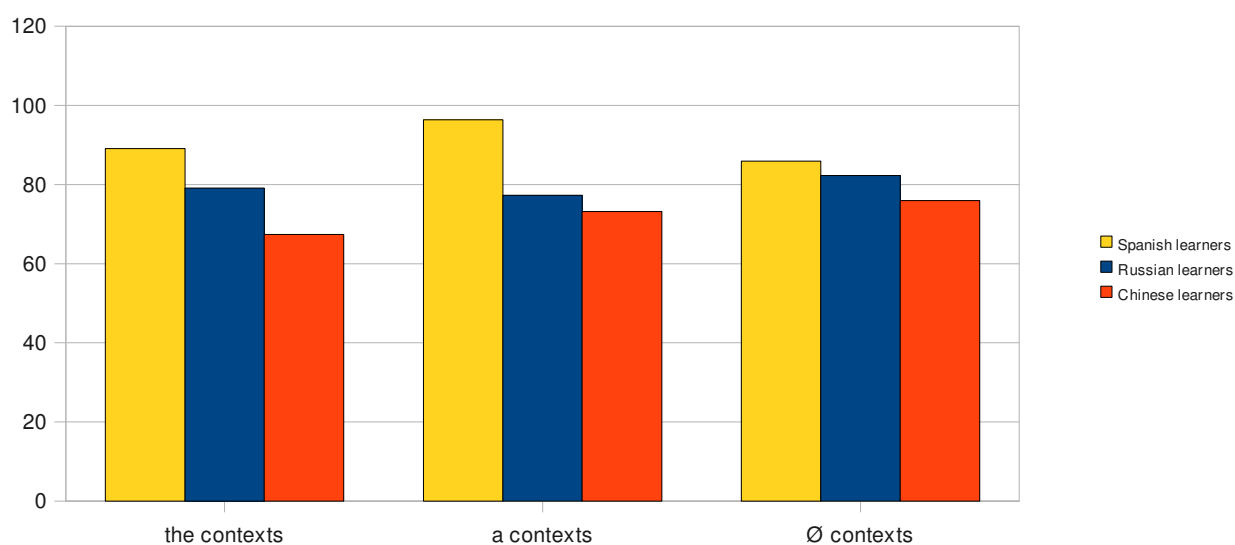
Table 12. Accuracy percentages for Russian, Chinese and Spanish learners revealed by the TLU measure system

Articles	Percentages		
	Russian	Chinese	Spanish
the	79.1	67.4	89.07
a	77.3	73.18	96.39

∅	82.3	75.94	85.94
---	------	-------	-------

Figure 6 is based on Table 12 and thus visualizes the difference in accuracy percentages of Russian, Chinese and Spanish learners.

Figure 6. Accuracy percentages for the definite, indefinite and zero articles based on TLU for the Russian, Chinese and Spanish learners



Spanish learners, as was predicted by Díez-Bedmar and Papp, do use articles more accurately than Chinese or Russian learners. The reason why Spanish learners use articles most accurately is that they belong to the [+Art] group and need just to adjust their article system to the English one what results in a better performance for all of the articles. However, it is hard not to notice that even though Russian learners belong to the [-Art] group as well as Chinese learners, they perform better for all of the articles, including the definite and indefinite articles, which Chinese learners were supposed to use more accurately than the Russian learners.

It is possible that Russian learners use articles more accurately because they are at the higher level of English proficiency as, each country has their own specific educational program

for English language teaching, or due to the language specifics Russian learners acquire the English article system faster. However the last assumption is hard to prove as the Chinese language is claimed to be on its way to article grammaticalization and thus Chinese learners are supposed to use the definite and indefinite articles more accurately in comparison with the Russian learners. Thus, could be speculated that Russian learners use articles more accurately as a result of a different and more efficient article teaching system.

Spanish learners cope best in comparison with Chinese and Russian learners with the indefinite article and least well with the zero article, which they were predicted to have most problems with. Chinese and Russian learners are close to the Spanish learners in accuracy rates for the zero article since the zero article is used least accurately by the Spanish learners.

As we can see the main difference in article use between Russian and Chinese learners lies in the use of the definite article. My presupposition is that since Chinese learners are less accurate in article use than Russian learners they probably are at the earlier stage of the definite article flooding, which resulted in a more severe definite article overuse and lowered its accuracy rate. If it is so, then following Huebner (1987), they are going through the second stage of definite article acquisition.

It would be easier to tell if Chinese learners use the definite article least accurately because they overuse it in other contexts or simply use it less accurately than the indefinite article in obligatory contexts if we had accuracy rates based on SOC. Even though SOC overestimates learners' accuracy it shows how accurately the article is used in its obligatory contexts and overused in non-obligatory contexts separately what appears to be more important for the Contrastive Interlanguage Analysis. Applying SOC we can compare both accuracy in obligatory and overuse in non-obligatory contexts, while applying TLU we simply compare accuracy in general.

That is why I think that TLU is more suitable for the definition of the article acquisition order while SOC is more appropriate for the Contrastive Interlanguage Analysis. Unfortunately I am able to compare only accuracy rates for the Spanish, Chinese and Russian learners based on TLU, as Díez-Bedmar and Papp did not apply SOC in their investigation of the article use by the Spanish and Chinese learners of English.

4.8.3 Misuse

Proceeding with the analysis of article overuse by Chinese and Russian learners I propose to have a look at the overuse percentages in each of the three article contexts. It will provide the possibility to get a better understanding of how Chinese and Russian learners overuse articles depending on semantic context.

Following Díez-Bedmar and Papp, I have estimated article misuse in obligatory contexts relative to the sum of obligatory and non-obligatory contexts presented in Table 13.

number of incorrect suppliance in obligatory contexts

number of obligatory contexts and non-obligatory contexts

Misuse percentages for the Chinese learners, which are presented in Table 13, were taken from Díez-Bedmar and Papp's study (Díez-Bedmar and Papp 2008, see Table 7 p. 164).

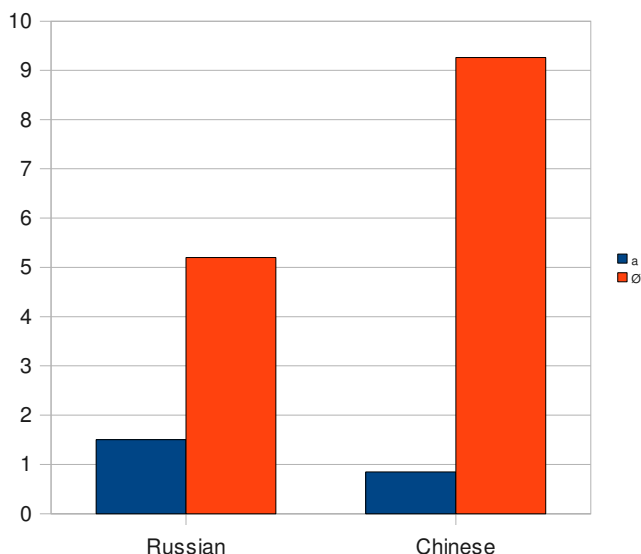
Table 13. Article misuse in obligatory contexts for the definite, indefinite and zero contexts based on the TLU measure system

Articles	Contexts	Percentages	
		Russian	Chinese
the	∅	5.5	9.3
	a	1.5	0.85
∅	a	2.2	0.74
	the	8.9	15.8
a	the	7.3	7.5
	∅	9.6	14.8

Immediately from Table 13 it is understood that Chinese learners use the definite article less accurately because they underuse it twice more than Russian learners. The zero article is used less accurately because Chinese learners misuse it for the definite article twice more than Russian learners. As to the indefinite article Chinese learners again underuse it more than Russian learners.

Figure 7 depicts article overuse in definite context and thus makes it possible to see to what extent the zero and indefinite articles are used in place of the definite article.

Figure 7. The diagram shows the percentage of the zero and indefinite articles used inappropriately in definite contexts



As we can see both Chinese and Russian learners tend to misuse the definite article more for the zero article (see example 42) than for the indefinite article (see example 43).

42. (1ZA) People simply do not have enough time to examine what is (1GADA-ZA) truth and what is (1IA) a lie.
43. I think this is (2GADA-IA) an evidence that (1ZA) human imagination is much more extensive than its result – (1ZA) technology, and that is why it cant be limited.

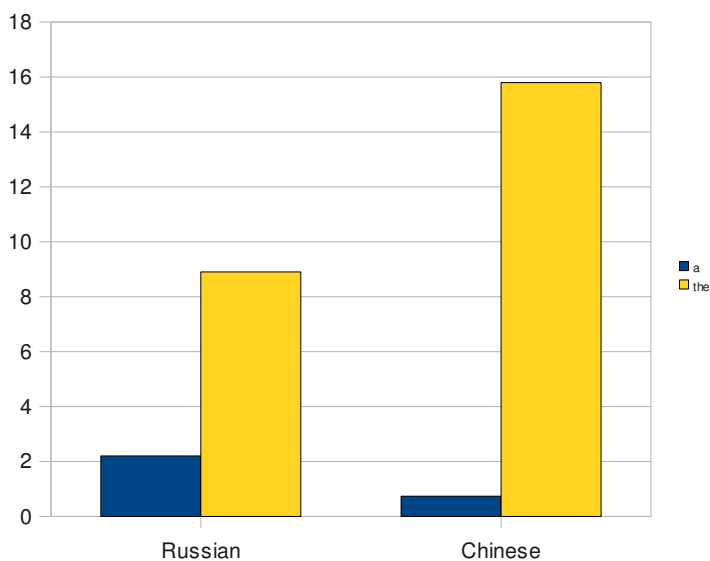
However, the proportion of the zero and indefinite articles used instead of the definite article is different for the Russian and Chinese learners. Russian learners use the zero article instead of the definite article about four times more often than the indefinite article while Chinese learners use it about nine times more often than the indefinite article.

The difference between zero and indefinite article overuse is less significant for the Russian learners when compared to Chinese learners. This means that Russian learners do not just underuse the definite article but also misuse it for the indefinite article. The fact that Chinese learners mostly underuse the definite article and just in some cases misuse it for the indefinite article give grounds to think that they are at the lower stage of the definite article acquisition as the definite article is being significantly underused.

On the other hand, if article flooding is the first stage in article acquisition the fact that Russian learners overuse the indefinite article more can be seen as the first stage in the indefinite article acquisition and thus a shift from definite article flooding to indefinite article flooding. This, however, is just an assumption which needs to be proved by further analysis of the acquisition process for the Russian learners.

When the zero article is the target learners tend to misuse it for the definite article rather than for the indefinite article (Figure 8).

Figure 8. Percentages of the definite and indefinite articles inappropriately used in zero contexts



While the percentages of the use of the indefinite article for the definite article (see Figure 7) were closer between Chinese and Russian learners, as was expected as they both belong to the [-Art] group, in the zero context the rates of indefinite article overuse are closer between Chinese and Spanish learners. Russian learners again overuse the indefinite article most as they have more trouble restricting indefinite article uses to its obligatory contexts.

This means that there are cases where Russian learners associate zero article uses with definite (example 44) and indefinite article uses (example 45), while Chinese learners due to the higher definite article overuse, misuse the zero article almost exclusively for the definite article.

44. But how can we decide whose contribution to **(1GAZA-DA)** the society is more significant?
45. (1ZA) People who spent a few years in **(1GAZA-IA)** a prison also need **(1GAZA-IA)** a rehabilitation.

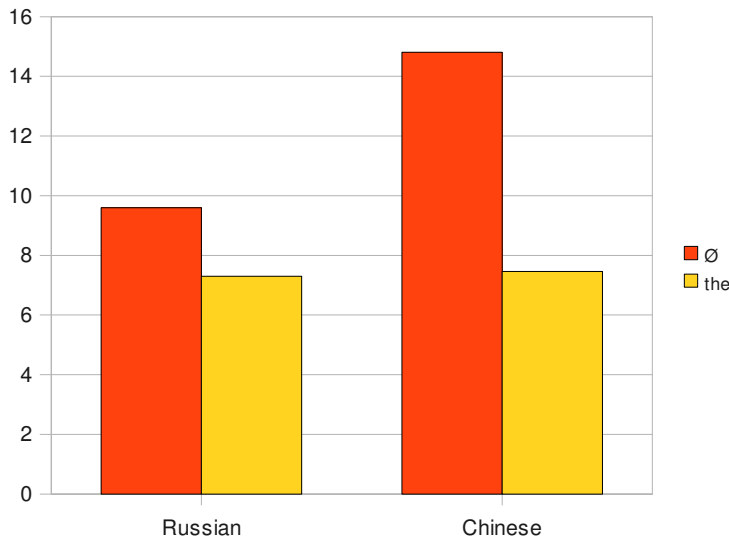
Since Chinese learners overuse the definite article in zero contexts almost twice more than the Russian learners also means that they have more trouble with correct zero article uses in obligatory contexts. The same can be said about the obligatory definite article contexts where Chinese learners overused the zero article almost twice more than the Russian learners. The fact that Chinese learners overuse the zero article twice more in definite contexts and the definite article twice more in zero contexts compared to the Russian learners also means that Chinese learners have more trouble both using the definite and zero articles in obligatory contexts due to the misuse of one for another. Conclusively, Chinese learners have more trouble differentiating the definite article in obligatory contexts from the zero article in obligatory contexts.

As to the indefinite article it has been shown that it was the Russian learners who had more trouble restricting its uses to obligatory contexts. Now we will see which of the articles is used instead of the indefinite article. Examples 46 and 47 illustrate the zero and the definite article used instead of the indefinite, respectively.

46. I can't imagine **(1GAIA-ZA)** really interesting life without (1ZA) imagination!

47. [...] (2DA) the possibility of (1ZA) moral humiliation and physical abuse ,
 (2DA) the aggravation of (1ZA) health or even (2DA) the death of (1GAIA-
 DA) the servicemen.

Figure 9. Percentages of the zero and definite articles used inappropriately in indefinite article contexts



Russian and Chinese learners overuse the definite article in the indefinite contexts equally but Chinese learners overuse the zero article in indefinite contexts more when compared to the Russian learners. Thus, Chinese learners use the indefinite article a bit less accurately than the Russian learners due to the higher zero article overuse in indefinite contexts. This, to my mind, is a result of the indefinite and definite articles underuse since the zero article was also represented more in definite contexts by Chinese learners.

Consequently, Chinese learners overuse the definite article more than the Russian learners because they have more problems differentiating zero article contexts from definite article contexts. This in its turn has caused less accurate use of the definite and zero articles by Chinese learners. They use the indefinite article less accurately due to its higher underuse in the obligatory contexts as it was more often misused for the zero article by Chinese learners.

4.9 Articles' accuracy in accordance with the four types of semantic context

The next part of my analysis is dedicated to the use of the articles in the four semantic contexts described in section 2.4. As stated earlier, in the theoretical part, Huebner (1983) classifies each NP as plus or minus specific referent ([+/-SR]) and plus or minus assumed known to the hearer ([+/-HK]), based on these two discourse features Huebner proposes four basic NP contexts that determine article use:

- Type 1. [-SR, +HK]: Generics
- Type 2. [+SR, +HK]: Referential Definites
- Type 3. [+SR, -HK]: Referential Indefinites
- Type 4. [-SR, -HK]: Non-Referentials

In order to master the English article system a learner would need to understand how to operate with these two discourse features. Learners who belong to the [-Art] group often associate definiteness with only one of the two discourse features, causing article mistakes. Based on their findings Huebner (1983) and Master (1987) claim that learners who belong to the [-Art] group associate the definite article considerably with [+HK] contexts, and use the indefinite or zero articles to a greater extent in [-HK] contexts. Parrish (1987) and Thomas' (1989) findings suggest that the definite article is initially associated with [+SR] rather than [+HK] contexts. Ionin claims that L2 English learners would overuse the definite article with [+specific] but not with [-specific] indefinites and explain the overuse of the definite article with indefinites in L2-English as a result of the fact that L2-English learners who belong to the [-Art] group use the definite article to encode specificity. Analyzing accuracy rates for each of the articles in the four semantic contexts I hope to shed more light on how Russian learners operate with the two discourse features and which of the two is crucial for definite article usage.

As mentioned in Díez-Bedmar and Papp's study (2008:159-160), Contrastive Analysis helps us to predict possible learners' mistakes based on the inconsistencies between the two languages. For example, Díez-Bedmar and Papp predict that Chinese learners would both underuse and misuse articles due to the absence of the article system in their own language. On the other hand based already on the mistakes made we can see which of our presuppositions were

true and try to explain the reasons for such mistakes .

Díez-Bedmar and Papp using Contrastive Analysis also predict the highest number of mistakes for the Spanish learners in generic contexts as a result of the inconsistencies in the two article systems. The accuracy rates as we can see from table 14 are lowest for the zero and indefinite article in generic contexts and thus as predicted Spanish learners do have troubles with generic contexts due to the fact that in the Spanish language the zero article is not used for generic reference. My hypothesis is that based on the accuracy rates for each of the articles in the four semantic contexts we can try to define the initial factors that cause such accuracy rates.

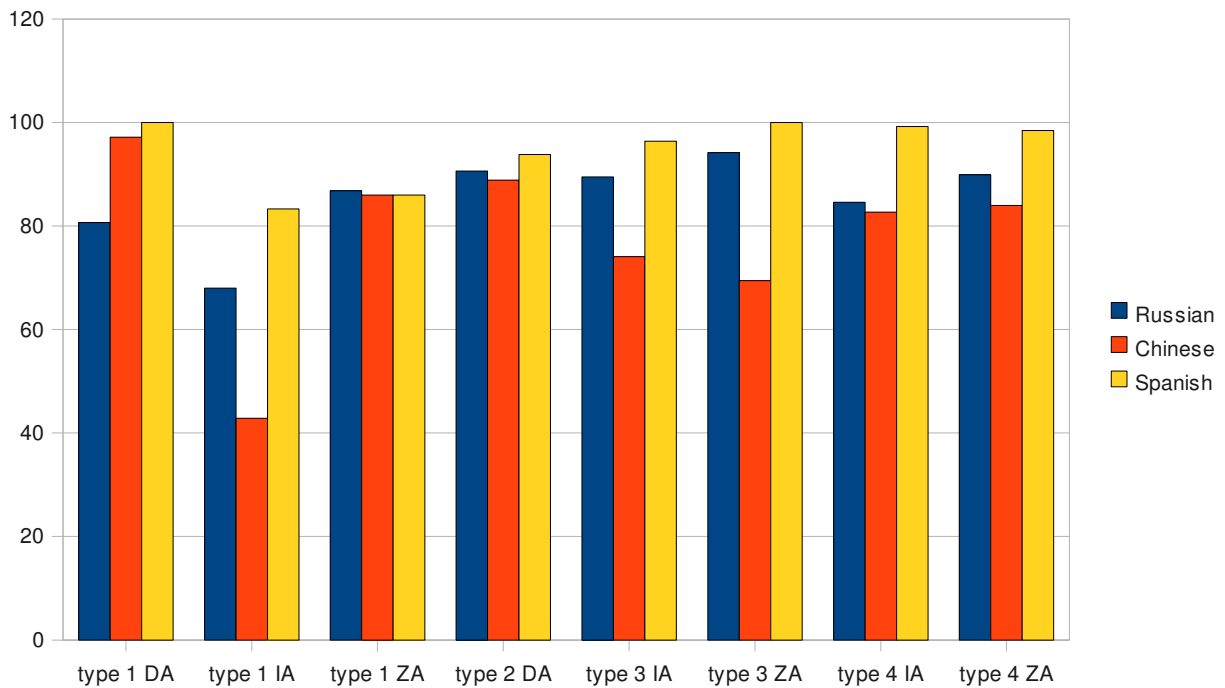
Consequently, comparing accuracy rates for each of the articles in different semantic contexts we can define in which of the contexts an article is used most correctly and in which least correctly. Since we can state in which of the four types of semantic contexts each of the articles is used most and least accurately we can determine if the article is associated with +/-SR or +/-HK.

Table 14 is based on Table 9 from Díez-Bedmar and Papp (2008:165) and represents Russian, Chinese and Spanish learners' accurate use of English articles according to the four types of NP in obligatory contexts. Figure 10 gives a visual representation of the data presented in Table 14.

Table 14. Accuracy percentages for the Russian, Chinese and Spanish learners in the four semantic contexts

	Type 1 Generics			Type 2 Referential Definites			Type 3 Referential Indefinites			Type 4 Non-Referentials		
	Chi	Russ	Spa	Chi	Russ	Spa	Chi	Russ	Spa	Chi	Russ	Spa
DA	97.14	80.7	100	88.9 0	90.6	93.87						
IA	42.86	72.9	83.33				74.10	89.5	96.37	82.68	84.6	99.2
ZA	85.99	86.8	85.99				69.42	94.2	100	83.98	89.9	98.46

Figure 10. Accuracy rates for the Russian, Chinese and Spanish learners of English by semantic context



Already from figure 10 based on the accuracy rates for the Russian (marked with blue) and Chinese (marked with red) learners we can specify main consistencies and inconsistencies in article usage by semantic context. The most notable difference in article accuracy is in type 1: accuracy rates of the definite and indefinite articles, and type 3: accuracy rates of the indefinite and zero articles.

I think that the discrepancy in accuracy rates of the Russian and Chinese learners is a result of the inconsistencies in the way definiteness and referentiality are marked in the two languages. My aim is to determine the reasons which cause Chinese or Russian learners to use one of the articles more accurately in one but not another type of NP context.

Based on the accuracy rates for the indefinite article Chinese learners have most trouble using the indefinite article in generic contexts [-SR;+HK] and less in specific [+SR;-HK] and non-specific [-SR;-HK] contexts. Based on this information we can assume that Chinese learners associate the indefinite article with [-HK] since the main difference between type 1(generics) and type 3 and 4 (referential and non-referential indefinites) is that type 1 relies on the assumption

that the referent is known to the hearer while in types 3 and 4 that the referent is unknown to the hearer. Thus, Chinese learners tend to use the indefinite article more accurately when the referent is not known to the hearer and less accurately when the referent is known to the hearer. This gives us grounds to assume that Chinese learners associate the indefinite article with [-HK].

On the other hand, we should not forget that generic reference is a reference to the whole class and thus it is possible that Chinese learners have more trouble using the indefinite article when the whole group or class is meant. Another option is that, following the rules of the English grammar, Chinese learners associate the indefinite article with -HK since the primary rule for the indefinite article's successful use is first mention of the referent which is assumed not known to the hearer.

It is also possible that the indefinite article in Chinese, possibly due to the language specifics is less associated with [+HK] and more with [-HK]. Even though a more detailed study of how referentiality and definiteness are realized in the Chinese language is needed this could mean that Chinese learners of English are more likely to interpret the reference as indefinite when it is not known to the hearer.

As to the zero article Chinese learners seem to associate it with [-SR] as it is used most correctly in types 1 and 4. Thus, Chinese learners use the zero article most accurately when no specific referent is meant. This is not surprising since only referential nouns can be definite or indefinite in Chinese and thus the fact that Chinese learners use the zero article more accurately when no referent is meant can be seen as a result of how definiteness and specificity are being realized in the Chinese language. This assumption is, however, in direct contradiction to Ionin's Fluctuation Hypothesis according to which no transfer effect is possible.

Before I move further I should mention that accuracy rates for type 1 and 4 for the Russian and Chinese learners are not directly comparable as I have included some of the instances which Díez-Bedmar and Papp identified as type 4 into type 1. This can be the reason why Russian learners seem to use the zero article a bit more accurately in type 4 than type 1 while Chinese learners use it more or less equally.

The definite article seems to be more associated with [+HK] rather than with [+SR] when it comes to Chinese learners since it is used a bit more accurately in type 1 [-SR;+HK] than in type 2 [+SR,+HK]. This supports Huebner's finding that the definite article is associated with

[+HK].

Consequently, Chinese learners use the indefinite article more accurately when it is assumed to be not known to the hearer and the definite article when it is assumed to be known to the hearer while the zero article is used most accurately when no specific referent is meant.

This can be due to language transfer or due to the appliance of grammatical rules for the use of the articles: we use the indefinite article when we mention something for the first time and the definite article when the referent is already known to the hearer. However, since Chinese learners have most trouble using the zero article in type 3 it perfectly goes with the fact that Chinese learners associate the zero article more with [-SR] as only referential nouns can be definite or indefinite. The low accuracy rate for the indefinite article in type 1 contexts can be explained as non-association of the indefinite article with generic reference or [+HK].

Now if we move on to the Russian learners they, when it comes to the indefinite article, as similar to the Chinese learners, have most trouble using the indefinite article in generic contexts, however, they use it more accurately when compared to the Chinese learners. Even though both Russian and Chinese learners have troubles using the indefinite article in type 1 [-SR;+HK] context, when we look at type 3 [+SR;-HK] and 4 [-SR:-HK] contexts Russian learners still have more trouble using the indefinite article while Chinese learners have more trouble using the zero article in type 3 and the indefinite article in type 4.

To begin with, since the indefinite article is less accurate in type 1 [-SR;+HK] than type 4 [-SR;-HK] we can assume that Russian learners use it more accurately when the referent is unknown to the hearer since both type 3 and 4 are marked as [-HK] while type 1 is marked as [+HK]. Thus, Russian learners as well as Chinese learners use the indefinite article more accurately when the referent is unknown to the hearer (examples 48 and 49) while the fact whether a specific referent is meant or not does not seem to cause trouble.

48. Now I want to draw your attention to **(3IA)** an article about Swedish prisons published in 'spid-info', N9 1995.
49. Alcoholism is **(4IA)** an illness, and if somebody is ill, **(2DA)** the best decision is to cure him.

As to zero article use, Russian learners supply it with more or less the same accuracy, which is higher than the accuracy of the indefinite article in all three types of NP contexts. Since the accuracy for the zero article is a bit lower in the first type when compared to the third and fourth types, I believe that Russian learners use the zero article more accurately when the referent is not known to the hearer as the common trait between the third and fourth types is [-HK], as in examples. However, type 1 is not just marked as [+HK,-SR] as it also differs from types 3 and 4 in that it is a generic reference. Thus, it can be speculated that Russian learners have more trouble using the zero article in generic contexts. Surprisingly enough they misuse the zero article in generic contexts for the definite article in 90 % of the cases, as in example 50.

- 50.** Thus (1ZA) humanities study everything that is connected with (1GAZA-DA) the human beings, (2DA) the most important facts in (2DA) the development of (1GAZA-DA) the human civilization.

Thus, Russian learners are guided by the same discourse feature both for the zero and indefinite articles, however, since they use the zero article more accurately in all three contexts, the indefinite article is more associated with the [-HK] while the zero article is used quite accurately in types 1, 3 and 4 irrespective of the fact that type 1 presupposes that the referent is known to the hearer. However even though type 1 is generic and due to Russian language specifics should be more associated with the zero article, as in the Russian language generic reference is not marked for definiteness or specificity, Russian learners use the zero article in type 1 less accurately than in types 3 and 4 because they misuse it for the definite article.

Russian learners, unlike Chinese learners, associate the definite article more with a specific referent than hearer knowledge as it is used more accurately in type 2 [+SR;+HK] than type 1 [-SR;+HK], however, as example 50 shows Russian learners misuse the zero article in type 1 almost exclusively for the definite article. This means that even though the definite article is used less accurately in type 1 (examples 51) than in type 2 (example 52) it is overused in type 1 (example 53 and 50) in zero article contexts. However, the main difference between the Russian and Chinese learners would be that Russian learners tend to use the definite article when they have a specific referent in mind while Chinese learners are guided by the hearer knowledge.

51. For example, why should one go to (**1GADA-IA**) a theatre if he can watch it at home.
52. Even (**2DA**) the opponents of (**2DA**) the idea of (1ZA) human equality are presently forced to admit that this concept is widely recognised as just and stimulating.
53. Hopefully, one of (2DA) the lessons (**1GAZA-DA**) the mankind has learned in this century is that (4IA) a society based on (1ZA) social or racial inequality – (1ZA) slavery, fascism, communism, apartheid - is not only immoral, but is doomed to (1ZA) backwardness.

If Chinese learners associate the definite article more with hearer knowledge than a specific referent it might explain why Russian learners fluctuate between the two settings of the Article Choice Parameter (see section 2.4) while Chinese learners do not. On the other hand, there is also reason to believe that L1 transfer does play a role in article acquisition and cannot be rejected.

Chapter 5 Conclusions

I started my investigation wondering if all learners of English whose mother tongue lacks articles acquire them in the same order. In order to restrict this I decided to conduct a corpus based cross-sectional study on article use and acquisition by Russian learners. After a revision of the theoretical material on acquisition order estimation I realized that acquisition is defined by accuracy which changes as the proficiency increases and can be calculated with the help of two different measure systems: SOC (supplied in Obligatory Contexts) and TLU (Target Like Use). Díez-Bedmar and Papp claimed that TLU is more reliable in accuracy estimation, however, I decided to apply both. As a result I had four questions to be answered:

1. What do the SOC (Supplied in Obligatory Contexts) and TLU (Target Like Use) measure systems tell us about the use of English articles by Russian native speakers in terms of accurate use and acquisition order?
2. Which of the two systems is more reliable when describing acquisition order and why?
3. What is similar and what is different in article use by Russian and Chinese learners in terms of acquisition order, accuracy and overuse?
4. Do Russian and Chinese learners tend to associate the definite article with [+HK] or [+SR]?

According to the SOC measure system Russian learners at the advanced level of proficiency use the definite article most accurately followed by the zero article while the indefinite article is acquired last. TLU, unlike SOC, revealed that if overuse is taken into account it is the zero article which is used most accurately and that the definite article is overused to such extent that it stands second in acquisition sequence. The indefinite article is acquired last according to both measure systems.

The next conclusion I came to in the course of my investigation was that TLU is a more reliable measure system for the acquisition order estimation. However, in my opinion, applying SOC one

can conduct Contrastive Interlanguage Analysis of article use with regard to accuracy in obligatory contexts, article misuse in obligatory contexts and article overuse in non-obligatory contexts while TLU allows comparison of acquisition orders only. I think that by comparing learners' article accuracy and misuse in obligatory contexts and overuse in non-obligatory contexts we will be able to define common and different traces in article use with regard to both accuracy and overuse. To estimate article overuse in non-obligatory contexts I have devised SNOC (Supplied in Non-Obligatory Contexts)

My findings do not support Díez-Bedmar and Papp's claim that the indefinite article is acquired prior to the definite article in [-Art] languages since Russian learners use the indefinite article least accurately according to both measure systems. Moreover, I hypothesize that Chinese learners use the definite article less accurately because of the severe overuse in non-obligatory contexts, which, according to Huebner (1983), is a typical stage in article's acquisition process. The fact that Chinese learners overuse the definite article more than the Russian learners supports my hypothesis.

Another interesting finding is that Russian learners, unlike Chinese learners, associate the definite article more with a specific referent than hearer knowledge as it is used more accurately in type 2 [+SR;+HK] than type 1 [-SR;+HK], however, Russian learners misuse the zero article in type 1 almost exclusively for the definite article, probably because they assume generic reference to be a common knowledge.

The direction for the present study was conditioned by previous studies. Having analyzed theoretical material on article acquisition I have read many studies on article acquisition and most of them deal with acquisition orders, accuracy hierarchies, measure systems. My personal opinion of a non-native speaker of English is that the knowledge of the acquisition order does not help to improve article mistakes or to understand which aspect of article use should be paid more attention. For the further research on article acquisition by Russian learners I would propose to concentrate more on the analysis of the mistakes and the reasons causing these mistakes. Pursuant to this aim I would like to refer to the following peculiarities of article use by Russian learners revealed in the present study:

1. They have trouble understanding how to use articles with abstract nouns;
2. They underuse the indefinite article twice more than the definite article. The

indefinite article is often underused in of-phrases as *a part of, a style of, a source of, a sort of* and abstract nouns and both underused and misused for the definite article with non-referential nouns.

Since Russian learners at the advanced level of proficiency had most trouble in correct article use with abstract nouns, of-phrases and non-referential nouns my suggestion is to teach the English article system in several stages. From my own experience I know that articles are usually introduced at the early stages of English teaching and are never referred to at the later stages.

Thinking of what have been done before in research on article acquisition by Russian learners I would say that my main finding is that Russian learners still have trouble using articles with abstract and non-referential nouns at advance level of proficiency and that they tend to associate the definite article with [+SR].

Literature

Primary sources

ICLE/RICLE

Secondary sources

Altenberg, B. and Granger, S. (eds). 2002. *Lexis in Contrast. Corpus-based Approaches*. Amsterdam/Philadelphia: John Benjamins [Studies in Corpus Linguistics 7] 3–48.

Bailyn, John. 1995. *A configurational approach to Russian 'free' word order*. Ph. D. dissertation, Cornell University.

Benson, C. 2002. *Transfer/Cross-linguistic Influence*. *ELT Journal* 56(1): 68–70.

Bickerton, D. 1981. *Roots of Language*. Ann Arbor, MI: Karoma Press.

Brown, J. D. (1983). An exploration of morpheme-group interactions. In K. M. Bailey, M. H. Long, & S. Peck (Eds.), *Second language acquisition studies* (pp. 25-40). Rowley, MA: Newbury House Publishers.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL teacher's course*. Boston: Heinle & Heinle.

Chaudron, C. & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, 12, 43-65.

Chesterman, A. 1991. *On Definiteness: A study with Special Reference to English and Finnish*. Cambridge: Cambridge University Press.

Christophersen, P. 1939. *The articles: A study of their theory and use in English*. Munksgaard, Copenhagen.

Dirdal Hidlegunn 2005. *Second language acquisition of articles and plural marking by Bengali learners of English*. PhD thesis, University of Cambridge.

Downing, E. and Locke P. 2002. *A University Course in English*. Second edition. London: Routledge.

Ekiert, M. 2007. *The Acquisition of Grammatical Marking of Indefiniteness with the Indefinite Article a in L2 English*. Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, Vol. 7, No. 1.

- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Fen-Chuan, C. 2001. *The acquisition of English articles by Chinese learners*. *Second Language Studies*, 20(1): 43-78.
- Fodor, J. & Sag, I. (1982). *Referential and quantificational indefinites*. *Linguistics and Philosophy* 5, 355-398.
- Gilquin, G. 2001. *The Integrated Contrastive Model. Spicing up your data*. *Languages in Contrast* 3.1: 95-123.
- Gilquin, Gaëtanelle, Szilvia Papp and María Belén Díez-Bedmar (Eds.) 2008. Linking up contrastive and learner corpus research. Amsterdam/New York, Rodopi, (pp. 147-175).
- Goto Butler, Y. 2002. *The Use of English Articles: An Analysis of the Metalinguistic Knowledge Used by Japanese Students in Acquiring the English Article System*. University of Pennsylvania
- Granger, S. 1996. 'From CA to CIA and back. An integrated approach to computerized bilingual and learner corpora'. In K. Aijmer, B. Altenberg & M. Johansson (eds) *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*. Lund 4–5 March 1994 (pp. 37–51). Lund: Lund University Press.
- Granger, S. 1998. *A bird's eye view of computer learner corpus research*. In S. Granger and J. Hung (eds) 1998, pp. 45-48.
- Granger, S., Hung, J. and Petch-Tyson, S. (eds). 2002. *Computer Learner Corpora, Language Acquisition and Foreign Language Teaching*. Amsterdam/Philadelphia: John Benjamins [Language Learning & Language Teaching 6].
- Granger, S., Lerot J. and Petch-Tyson, S. 2003. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam – New York, NY: Rodopi.
- Hasselgård, H., Johansson, S., Lysvåg, P. 1998. *English Grammar: Theory and Use*. Oslo: Universitetsforl.
- Hawkins J.A. 1978. *Definiteness and indefiniteness*. Guildford, London and Worcester, UK.
- Hawkins, R., Al-Eid S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Leung, I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N. & K. Velasco-Zarate. 2006. *Accounting for English article interpretation by L2 speakers*. In Foster-Cohen, S. H. (ed.), *EUROSLA Yearbook*, 6, (pp. 7–25). Amsterdam: John Benjamins.
- Huebner, T. 1983. *A Longitudinal Analysis of the Acquisition of English*. Ann Arbor, MI:

Karoma.

Humphrey, J. 2007. *Acquisition of the English Article System: Some Preliminary Findings*. Nagoya University of Foreign Studies Journal of School of Foreign Languages. Vol 32. 301-325.

Ionin, T. R. 2003. *Article semantics in second language acquisition*. Ph. D. dissertation. Dissertation Abstracts International, A: The Humanities and Social Sciences.

Ionin, T., Ko., H., & Wexler, K. 2004. *Article semantics in L2-acquisition: the role of specificity*. *Language Acquisition*, 12, (1), 3–69.

Lado, R. 1957. *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.

Larca, B. 1999. *Gramatica Descriptiva de la Lengua Espanola*. Madrid: Editorial Espasa Calpe, S. A. 891-928

Li, C. & Thompson, S. 1981. *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.

Leech, G. 1998. "Preface". Granger, S. (ed.) 1998. *Learner English on Computer*. London/New York: Longman.

Ludlow, P. and Neale, S. 1991. *Indefinite descriptions: In defense of Russell*, *Linguistics and Philosophy*, 14, 171-202

Lyons, C. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Maratsos, M. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition*. Cambridge, UK: Cambridge University Press.

Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, UCLA.

Master, P. (1993, April 15). *A contrastive study of determiner usage in EST research articles*. Paper presented at the TESOL Convention, Atlanta, Georgia.

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

Naylor, Kenneth E. 1983. *On Expressing Definiteness in the Slavic Languages and English*. American Contribution to the Ninth International Congress of Slavists. Columbus: Slavica. *Linguistics* vol. 1. 203-220.

Palmer, H. 1939. *A grammar of spoken English on a strictly phonetic basis*, 2nd. edn. Heffer, Cambridge.

Parrish, B. 1987. A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37(3), 361-383.

Pica, T. 1984. *Methods of morpheme quantification: their affect on the interpretation of second language data*, *Studies in Second Language Acquisition*, 69-78.

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1985) *A Comprehensive Grammar of the English Language*. London: Longman.

Reid, J., P. Battaglia, M. Schuldt, E. Narita, M. Mochizuki & N. Snape 2006. *The article choice of learners of English as a second language*. Unpublished term paper, University of Essex, UK.

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

Sinclair, J. M. (ed.) (1991) *Corpus, concordance, collocation*. Oxford University Press, Oxford, UK.

Snape, Neal, Y.-k. Ingrid Leung, and Hui-Chuan Ting. 2006. *Comparing Chinese, Japanese and Spanish Speakers in L2 English Article Acquisition: Evidence against the Fluctuation Hypothesis?* In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, ed. Mary Grantham O'Brien, Christine Shea, and John Archibald, 132-139. Somerville, MA: Cascadia Proceedings Project. www.lingref.com, document #1495.

Stephanides, Eva H. 1978. *The use of the article in English and Hungarian: a contrastive analysis*. *Papers and studies in Contrastive Linguistics* 8. Poznan: Adam Mickiewicz University.

Stig, Johansson and Knut Hofland. 1994. *Towards an English-Norwegian parallel corpus*. In U. Fries, G. Tottie, and P. Schneider (eds), *Creating and using English Language Corpora*, 25-37. Amsterdam and Atlanta, GA: Rodopi.

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

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

Ting, H. C. 2005. *The acquisition of articles in L2 English by L1 Chinese and L1 Spanish speakers*. Unpublished MA dissertation, University of Essex, UK.

Van Deusen-Scholl, N. and Hornberger, N. H. (eds) 2008. *Encyclopedia of Language and Education*. 2nd Edition, Volume 4: *Second and Foreign Language Education*, 337-351. Springer Science+Business Media LLC.

Internet sources

<http://www.merriam-webster.com/>

http://en.wikipedia.org/wiki/Contrastive_analysis

Attachment 1

Table 5. Taxonomy of the English article system with examples

Category	Article	Environment	Examples
Type 1. [−SR,+HK]	<i>the, a, Ø</i>	Generic nouns	Ø Fruit flourishes in the valley. Ø Elephants have trunks. The Grenomian is an excitable person. They say the elephant never forgets: A paper clip comes in handy. An elephant never forgets.
Type 2. [+SR,+HK]	<i>the</i>	Referential definites previous mentions specified by entailment specified by definition unique in all contexts unique in a given context	Pass me the pen. The idea of coming to the UK was... I found a book. The book was... The first person to walk on the moon...
Type 3. [+SR,−HK]	<i>a, Ø</i>	Referential indefinites first-mention NPs, or NPs following existential ‘has/have’ or ‘there is/there are’	Chris approached me carrying a dog. I've bought a new car. A man phoned. I keep sending Ø messages to him. I've got Ø friends in the UK. I've managed to find Ø work.
Type 4. [−SR,−HK]	<i>a, Ø</i>	Non-referential nouns attributive indefinites non-specific indefinites	Alice is an account. I need a new car. I guess I should buy a new car. A man is in the ladies, but I have not seen him. Ø Foreigners would come up with a better solution.
Type 5.	<i>a, the, Ø</i>	Idioms Other conventional uses	All of a sudden, he woke up. In the 1950s, there weren't many cars. His family is now living Ø hand to

Attachment 2

1<ICLE-RU-MOS-0003.2>

(1ZA) Social equality is one of (2DA) the most controversial values of (1ZA) modern society. Over (2DA) the centuries, (1ZA) Christianity has been promoting (3IA) an idea of all (1ZA) men being equal in (2DA) the eyes of God. Since (2DA) the French revolution (1ZA) equality, brotherhood and freedom have been proclaimed (2DA) the highest virtues of any society. It is widely recognised now that (4IA) a country could be considered civilised only if (2DA) the equality of all (1ZA) human beings translated into (1ZA) social relationship clearly implies that all (1ZA) men should have (1ZA) equal opportunities, i.e. equal starting conditions.

However, (1ZA) people are different in terms of their talents, dedication, goals in life, etc. This difference seems to deny (2DA) the very idea of (1ZA) equality. (1ZA) Human ambition makes (1ZA) people seek (1ZA) positions of (1ZA) power which promote (2DA) the sense of their exceptional value for (1ZA) society and superiority over other more humble human beings. (1IA) Excessive ambition is (4IA) a theme addressed in (2DA) the best pieces of (2DA) the world's literature. Macbeth has become (4IA) a symbol of (1ZA) disaster caused by (1ZA) excessive ambition, Raskolnikov showed (2DA) the world what happens if (4IA) a person considers his life more valuable than (2DA) the life of other persons, however useless or insignificant. (1IA) An ambitious desire of (4ZA) a person or a group of persons to prove themselves and (2DA) the entire world that they are smarter, better, more deserving than others gave rise to all racist theories, (2DA) the consequences of which in the XXth century are too well known.

Hopefully, one of (2DA) the lessons (1GAZA-DA) the mankind has learned in this century is that (4IA) a society based on (1ZA) social or racial inequality – (1ZA) slavery, fascism, communism, apartheid - is not only immoral, but is doomed to (1ZA) backwardness. If (1GAZA-DA) the slave labour were effective, Russia would have (2DA) the highest living standards in (2DA) the world., and (2DA) the communist regime would have prospered. Even (2DA) the opponents of (2DA) the idea of (1ZA) human equality are presently forced to admit that this concept is widely recognised

as just and stimulating. G. Orwell's catch phrase reflects (1DA) the need to pay at least (1ZA) lip service to (1ZA) equality even by (2DA) the worst totalitarian regimes - true, 'all animals are equal', but with 'some animals being more equal than others', (2DA) the status quo is restored. (1GAZA-DA) The appearances are saved but (2DA) the basic principle of (1ZA) inequality is in place.

(2DA) The most common arguments against (1ZA) social equality are that it looks nice on (1ZA) paper but is unattainable in (1ZA) real life, that (1ZA) natural inequality of (1ZA) men makes (1ZA) social equality (4IA) a dream, (4IA) a desired goal at best. Therefore, 'some animals' have always been and will always remain 'more equal than others'. It cannot be denied that some people are more talented, energetic, sophisticated than others. However, (2DA) the world will be (4IA) a better and more secure place to live in if (2DA) the advocates of (1ZA) human equality persistently worked on promoting (1IA) a better understanding of what it actually means in terms of (1ZA) social relationship and social justice rather than hypocritically shutting their eyes to (1ZA) natural inequality. (1ZA) People have to feel that they are (4ZA) valuable members of their society regardless of (2DA) the amount of their personal contribution to its development, let alone (2DA) the colour of their skin, (1ZA) ethnic origin, 'party membership' or other such privileges or disadvantages.