8-2019

# The L2 English Non-Generic Definite Article Use of L1 Azerbaijani Speakers

David J. Giles
*St. Cloud State University*

# The L2 English Non-Generic Definite Article Use of L1 Azerbaijani Speakers

by

David Giles

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree

Master of Arts in

English: Teaching English as a Second Language

August, 2019

Thesis Committee:
Edward Sadrai, Chairperson
James Robinson
Matthew Barton

**Abstract**

Researchers agree that article acquisition is one of the most difficult challenges facing learners of English (Chrabaszcz & Jiang, 2014; Lee, Park, & Heinz, 2018; Liu & Gleason, 2002). Research suggests that not all of the various uses of the article are equally difficult for all learners, and that the L1 plays a role in the type of difficulty experienced by learners (Chrabaszcz & Jiang, 2014; Lee et al., 2018; Liu & Gleason, 2002). This thesis proposes a hierarchy of difficulty of written L2 English non-generic definite article usage by L1 Azerbaijani speakers using the four categories of *missed obligatory usage* presented by Liu and Gleason (2002) (cultural, situation, structural, and textual), along with their four categories of *overuse* (cultural, structural, general reference, and ungrammatical), building on previous studies of definite article use by speakers of various L1s. I used a shortened version of Liu and Gleason's (2002) testing instrument consisting of sentences into which participants insert *the* wherever they deem necessary. Participants were also asked to write a brief narrative, which was analyzed for vocabulary size using Tom Cobb's *Compleat Lexical Tutor*, in order to determine their relative proficiency levels. In missed obligatory uses of the non-generic definite article, cultural and structural use presented equal difficulty for learners. In overuse of *the*, a clearly greater level of difficulty for the cultural category of use was demonstrated. The finding that the cultural use is problematic for learners supports the findings of Liu and Gleason (2002), Wong and Quek (2007), and Dikilitaş and Altay (2011), among others, while the finding that the structural use is also highly difficult for learners differs from the findings of these researchers. Results of this study also indicated that in missed obligatory usage, the situation category of use presented the least difficulty, while in overuse the ungrammatical category presented the least difficulty, supporting the findings of Liu and Gleason (2002).

# Table of Contents

## List of Tables

## List of Figures

# Chapter 1: Introduction

## 1.1. Difficulty of the English Article System

Researchers agree that English learners usually struggle a great deal to master English article usage (Chrabaszcz & Jiang, 2014; Lee, Park & Heinz, 2018). For learners, regardless of L1 background, the correct choice of article often continues to confound attempts at acquiring native-like accuracy and fluency even after years of practice and exposure. The complex article system of English is often considered to be one of the most difficult aspects of English structure for learners to acquire and for teachers to teach (Liu & Gleason, 2002).

One of the primary reasons for the difficulty of the English article system lies not in its form (there is, for example, only one definite article form in English, as compared to four in Spanish), but rather in the lack of a simple one-to-one correspondence between the article and a specific meaning or usage. According to Master (2002), the English article "stacks multiple functions onto a single morpheme" (p. 332), causing difficulty to the learner who "generally looks for a one-form-one-function correspondence" (p. 332). Instead of a simple, concrete meaning, the learner encounters a situation not limited to English: many different meanings and usages governed by rules which may seem vague and contradictory to learners and intertwined with the context and/or cultural references which may be unfamiliar. Master (2002) states two additional reasons for the difficulty experienced by foreign learners of English article use: articles occur very frequently, "making continuous conscious rule application difficult over an extended stretch of discourse" (p. 332), and, like other function words, are "normally unstressed and consequently very difficult if not impossible for a NNS [non-native speaker] to discern, thus affecting the availability of input in the spoken mode" (p. 332).

Research has shown that not all English article usages are equally difficult for learners, nor are all acquired at the same stage of learning (Liu & Gleason, 2002). In focusing on the definite article, for example, Liu and Gleason (2002) have delineated four different types of usage for the so-called *non-generic* definite article (as opposed to the generic usage, in which *the* refers to a general category as in *The Germans are known for their love of beer*). They found that some of these categories of usage are far more difficult for learners than others, and some continue to present problems even into the advanced stages of proficiency.

The English article system is explored in more detail in Sections 2.3 and 2.4. More background on the issue of L2 English article acquisition is provided in Section 2.5.

**1.2. Studies of L1 Influence on Article Acquisition**

Additionally, a large number of studies have examined the effects of L1 influence on L2 English article use (Chrabaszcz & Jiang, 2014; Ionin, Ko, & Wexler, 2004; Lee et al., 2018, among others). The general conclusion we may reach from such studies is that not all L2 English speakers make the same types or number of errors with articles. There does seem to be a good deal of influence from a learner's L1 in this regard (Lee et al., 2018, p. 80). As one might expect, learners whose L1 has an article system more or less similar to that of English experience less difficulty and greater accuracy in their L2 English article use, whereas those from article-less L1 backgrounds experience greater difficulty and less accuracy. A number of studies have been done using Liu and Gleason's (2002) model of four categories of non-generic definite article usage focusing on subjects with various L1s, with somewhat varying results (Dikilitaş & Altay, 2011; Garcia-Mayo, 2008; Koç, 2015, and Wong & Quek, 2007, among others). Such studies

have generally found varying levels of difficulty for each category of article use for learners at different proficiency levels.

L2 acquisition of English articles is explored further in Sections 2.5.

### 1.3. The Azerbaijani Language

Azerbaijani, also known as Azeri, is a member of the Turkic language family, and is closely related to standard Turkish. Thus, much of the literature relating to the grammatical structure of Turkish is valid for Azerbaijani as well. Azerbaijani is spoken by approximately 23 million people in two major dialects which differ somewhat in phonology, morphology and syntax (Dehghani, 2018). The northern variety is used in the Republic of Azerbaijan; it is currently written in a version of the Latin alphabet which is similar to that of Turkish. Formerly it was written in the Cyrillic alphabet and has been influenced by Russian, especially in its vocabulary. The southern variety is spoken in Iran, is written in the Persian alphabet, and has been more heavily influenced by the Persian language. As a whole, the Azerbaijani language makes use of many loan words from Arabic and/or Persian; additionally, the northern variety, that used in the Republic of Azerbaijan, has borrowed a significant number of Russian terms. Many Azerbaijanis speak some Russian, as the Republic of Azerbaijan was, until 1991, a republic of the Soviet Union.

Azerbaijani, like other Turkic languages, is an agglutinative language[1], making heavy use of suffixes to express grammatical relationships, and lacks grammatical gender. The rules of vowel harmony, in which the vowels of suffixes must harmonize with the backness or roundness

[1] According to www.glossary.sil.org, "An agglutinative language is a language in which words are made up of a linear sequence of distinct morphemes and each component of meaning is represented by its own morpheme."

of the preceding vowel in the root word, affect the phonetic form of suffixes. Azerbaijani lacks a definite article; definiteness or indefiniteness is expressed in other ways, such as word order and the use of case suffixes.

Further details on the structure of Azerbaijani and its expressions of definiteness vs. indefiniteness will be given in Section 2.6.

**1.4. The Aim of This Thesis**

This thesis aims to continue this line of investigation and, building on the categories set forth by Liu and Gleason (2002) and subsequent studies of hierarchies of difficulty, order of acquisition, and effect of L1 on article acquisition, determine levels of difficulty and order of acquisition of non-generic definite articles specifically for L1 Azerbaijani speakers. In order to accomplish this, a testing instrument (adapted from that of Liu and Gleason, 2002) utilizing 47 sentences into which participants insert *the* wherever they deem necessary, was administered to a group of English learners from the Republic of Azerbaijan (see Sections 3.3.1 and 3.4.1 for discussion of the testing instrument, and Section 3.2 for discussion of the participants in this study). Additionally, participants wrote a brief narrative which was used to determine their vocabulary size, and thus their relative English proficiency levels, using Tom Cobb's *Compleat Lexical Tutor* (see Sections 3.3.2 and 3.4.2 for discussion of the writing task). The resulting data was analyzed according to various categories of article usage. To date no such studies have been carried out for L1 Azerbaijani speakers to my knowledge, and studies of article use of speakers of Turkic languages other than Azerbaijani have yielded somewhat inconsistent results. In this study, comparison is made (see Section 2.5.3) to findings of other researchers, especially those

investigating article acquisition by L1 Turkish speakers, with an eye towards improving the effectiveness of instruction in L2 English definite article use in the Republic of Azerbaijan.

### 1.5. The Structure of This Thesis

In Chapter 2 (literature review), we will look at the theoretical background of the definite article, touching upon definiteness vs. specificity (Section 2.2), the use of articles (definite, indefinite, and zero) in English (Sections 2.3 and 2.4.1), different taxonomies of the non-generic definite article in English (Sections 2.4.2 and 2.4.3), and studies of L2 acquisition of the English non-generic definite article (Section 2.5). Additionally, the literature review will look in detail at the background and structure of the Azerbaijani language, including how it expresses definiteness (Section 2.6). The primary research question under investigation in this thesis is posed in Section 2.7.

In Chapter 3, we will look at the participants in this study and the methods and instruments used to gather data. Chapter 4 will discuss the results of the research, while chapters 5 and 6 discuss the implications and limitations of the results and conclude the thesis.

# Chapter 2: Literature Review

## 2.1. Introduction

To provide a background for understanding the dynamics when L1 Azerbaijani speakers attempt to acquire L2 English non-generic definite articles, it is necessary to look at the English system of articles, the acquisition of English articles by NNSs (non-native speakers), and the Azerbaijani language. In this chapter we will begin by looking at two semantic features that a language may encode in its articles: definiteness and specificity (see Section 2.2). We then examine the English article system (including *the, a/an*, and zero article), which encodes definiteness (not specificity), more closely (see Section 2.3). We will explore the expression of definiteness in English (see Section 2.4), including the non-generic definite article taxonomies of Hawkins (1978) (see Section 2.4.2) and Liu and Gleason (2002) (see Section 2.4.3), which will provide the framework for analyzing our data.

After a discussion of the taxonomy of Liu and Gleason (2002), we will look at some general issues in L2 acquisition of the English articles (see Section 2.5.1), and examine specifically the study of Liu and Gleason (2002) on the acquisition of non-generic definite articles (see Section 2.5.2), as well as a series of studies by other researchers on the acquisition of non-generic definite articles by speakers of various L1s (see Section 2.5.3).

Finally, we look in some detail at the Azerbaijani language, including its place in the Turkic language family (see Section 2.6.1), the general background of the language and its two varieties (see Section 2.6.2), a brief outline of its grammatical structure (see Section 2.6.3), and two ways in which Azerbaijani, which lacks a definite article, expresses definiteness (see Section 2.6.4).

The chapter concludes with the research question to be addressed in this thesis (see Section 2.7).

**2.2. Definiteness and Specificity**

According to Ionin et al. (2004), "articles cross-linguistically can encode different semantic features" (p. 5). Two such features are definiteness and specificity. Both of these features are "discourse related: They are related to the knowledge/mind state of the speaker and/or the hearer in the discourse" (Ionin et al., 2004, p. 5). If an article preceding an NP (noun phrase) is [+definite], then "the speaker and hearer presuppose the existence of a unique individual in the set denoted by the NP." In English, the feature [+definite] is encoded in the article *the*, and the feature [-definite] is encoded in the English indefinite article *a/an*:

(1) I saw a horse. The horse was an odd color.

If, however, an article is [+specific], then "the speaker intends to refer to a unique individual in the set denoted by the NP and considers this individual to possess some noteworthy property" (Ionin et al., 2004, p. 5). The feature [+specific] is not morphologically expressed in standard English; the articles *the* and *a/an* are used without regard to specificity. The feature [+specific] is expressed in the nonstandard language, however, in "the demonstrative *this* in its indefinite referential use" (Ionin et al., 2004, p. 7). For example:

(2) Dave has this really weird-looking hat.

Other languages do encode the feature [+specificity] morphologically, as Ionin et al. (2004) demonstrate in the case of Samoan, which has an article *le/l* referring to a particular entity without regard for its definiteness or indefiniteness. The difference between the two possible features of articles is that "the feature [+definite] reflects the state of knowledge of both speaker

and hearer, whereas the feature [+specific] reflects the state of knowledge of the speaker only" (Ionin et al., 2004, p. 5).

**2.3. The English Article System**

According to Ionin, Zubizarreta, and Maldonado (2008), the difference between the definite article *the* and the indefinite article *a/an* "is that of presupposition: use of *the* is felicitous when uniqueness of the referent has been established…while use of *a* is felicitous when this is not the case" (p. 556). Huebner (1983) concurs in stating that English articles mark "referentiality of noun phrases" (p. 13).

According to Liu and Gleason (2002), the most important and enlightening work in the analysis of the use of English articles was that of Bickerton (2016), originally published in 1981 in *Roots of Language*, as "it renders a new and unique systematic approach to the analysis of the use of the English article system" (p. 2). Bickerton (2016) proposed that the use of *the, a/an*, and zero article is governed by the semantic function of the NP (noun phrase) in question, and that the uses of these articles could be classified by reference to two binary discourse features: whether the noun is a specific referent (S), and whether the referent is "presupposed" (P) referring to "information presumed shared by speaker and listener" (p. 214). Analyzed in this way, any NP will fall into one of four types (Bickerton, 2016):

1. [+P, -S], "Generic," including *the, a/an*, and zero article.
2. [+P, +S], "Definite," including *the* only.
3. [-P, +S], "Indefinite," including *a/an* only.
4. [-P, -S], "Other," including *a/an* and zero article only.

Huebner (1983) renamed Bickerton's (2016) "presupposed" [P] discourse feature as [HK], meaning that the referent is assumed known to the hearer, and renamed Bickerton's (2016) [S] "specific referent" as [SR]. Huebner (1983) elaborates on Bickerton's (2016) four categories of article use as follows (p. 133, examples mine):

1. [-SR, +HK]. This is the generic category, which could include a definite, indefinite, or (if the NP is plural) a zero article (e.g., *The dog/A dog is a faithful friend*, or *Dogs are faithful friends*).
2. [+SR, +HK]. This category requires the use of the definite article. It includes four subcategories of use: (a) unique referent (e.g., *the Sun*) or conventionally assumed unique referent (e.g., *the King*); (b) physically present referent (e.g., *Please pass me the salt*); (c) previously mentioned referent; and (d) specific referent assumed known to the hearer (e.g., *I'm going to the store*, when there is only one store in the village).
3. [+SR, -HK]. This category requires the indefinite article, or the zero article if the NP is plural. It includes two subcategories: (a) first mention of an NP assumed unknown to the hearer (e.g., *I bought a truck*); and (b) first mention of an NP following "existential have" and assumed unknown to the hearer (e.g., *My car has a new engine*).
4. [-SR -HK]. This category also requires the indefinite article, or zero article if the NP is plural. It comprises four subcategories: (a) NPs in equative statements (e.g., *I am a doctor*); (b) NPs in negative statements (e.g., *She doesn't have a computer*); (c) NPs in interrogative contexts (e.g., *Do you have a pencil?*); and (d) NPs in hypothetical contexts (e.g., *If I were rich, I would buy an airplane*).

## 2.4. Definiteness in English

**2.4.1. Overview of the English definite article**. Concerning the definite article, Huebner (1983), referencing the work of Bickerton (2016) (see type [+P, +S]/[+HK, +SR] in Section 2.3 above), explains that it "is used to mark noun phrases which have specific referents assumed known to the hearer. These include noun phrases with unique referents (i.e., the Pope), and those with referents conventionally assumed to be unique (the President), as well as noun phrases whose referents are physically present, and noun phrases the referents of which the speaker assumes the hearer knows" (p. 131). Reid (1991) claims that the definite article in English gives the hearer necessary information to interpret a word appropriately by signaling to the hearer that he or she has sufficient information to identify or differentiate a referent.

The definite article may be divided into generic and non-generic usages. Generic usage of *the* refers to cases in which it refers to "a species, a race, or people of a nation" (Liu & Gleason, 2002), for example:

(3) Speaker A: I met a man from Turkey who spent twenty minutes telling me how beautiful his country is.

Speaker B: Yes, the Turks are usually quite proud of their country.

This usage corresponds to Bickerton's (2016) [+P, -S] category or Huebner's (1983) [+HK, -SR] category, in which the definite article is possible to express a generic meaning.

The non-generic usage is much more common than the generic usage, and includes all other uses of the definite article (Liu & Gleason, 2002, p. 5). According to Liu and Gleason (2002), "the nongeneric use of *the* is much more complex and hence more problematic for ESL students than the generic use" (p. 6). Hawkins (1978) proposed his so-called location theory to

explain the various uses of non-generic *the*, and specified eight types of non-generic use to illustrate how his theory operates. In drawing upon the taxonomy of Hawkins (1978), Liu and Gleason (2002) argue that these eight categories can be consolidated into four different categories of non-generic use of the definite article. In Section 2.4.2. the taxonomy of Hawkins (1978), and in Section 2.4.3 the taxonomy of Liu and Gleason (2002), will be explored more fully, as this thesis is based upon the four categories of non-generic definite article use proposed by Liu and Gleason (2002).

**2.4.2. Hawkins' (1978) location theory and uses of the non-generic definite article**. According to Hawkins' (1978) Location Theory, a speaker performs the following when using the definite article (p. 167):

(a) introduces a referent to the hearer

(b) instructs the hearer to locate the referent in some shared set of objects

(c) refers to the totality of the objects or mass within this set which satisfy the referring expression.

Hawkins (1978) gives eight categories of non-generic definite article use (examples from Hawkins):

1. Anaphoric use (pp. 107-110). When *the* is used in mentioning something a second time and subsequently (e.g., *Fred was wearing trousers. The pants had a big patch on them*.). In this example, *trousers* and *pants* are nearly synonymous and "the hearer has no difficulty in understanding what the definite NP refers back to" (p. 107).
2. Visible situation use (pp. 110-111). When *the* is used to refer to something both speaker and hearer can see (e.g., *Pass me the bucket, please*.).

3. Immediate situation use (pp. 111-115). When *the* is used to refer to something that is present but not visible. Similar to visible situation use above, except that the referent is not visible (e.g., *Beware of the dog*.).
4. Larger situation use based on specific knowledge (pp. 115-123). When *the* is used with a first-mention noun known to the community (e.g., people in the same village talking about *the church* or *the pub*).
5. Larger situation use based on general knowledge (pp. 117-123). When *the* is used with a referent that it is assumed people from a particular country or from around the world would know (e.g., *The White House, the moon*).
6. Associative anaphoric use (pp. 123-127). When *the* is used with a first-mention noun which refers back to a previously mentioned noun, but is not the same noun (e.g., *The man drove past our house in a car. The exhaust fumes were terrible*.). Differs from anaphoric use in that the noun is not the same. Hawkins (1978) refers to the previously-mentioned NP as the "trigger," which "must conjure up a set of objects which are generally known to be part of some larger object or situation. Thus, having mentioned *a car* I can enumerate all the parts of this object using a first-mention *the*" (pp. 123-124).
7. Unfamiliar use in NPs with explanatory modifiers (pp. 130-147). When *the* is used with a first-mention noun accompanied by an explanatory or identifying modifier (e.g., *I can't stand the name Algernon.*). The modifier may appear as a clause, a prepositional phrase, or a noun.

8. Unfamiliar use in NPs with nonexplanatory modifiers. When *the* is used with a first-mention noun accompanied by a modifier that does not provide explanatory information (pp. 148-149). There are a small number of modifiers (such as *same*) which require *the* but do not "introduce the unknown, definite referent to the hearer" (p. 148) (e.g., *My wife and I share the same secrets*.)

Liu and Gleason (2002) summarize Hawkins' theory: "…when an individual uses *the*, he or she invites the listener or reader to locate the referent by using provided or assumed known cultural, situation, structural, or textual information" (p. 7).

**2.4.3. Liu and Gleason's (2002) taxonomy of non-generic definite article use**. According to Liu and Gleason (2002), some of the eight types of non-generic definite article use proposed by Hawkins (1978) can be naturally grouped together. The anaphoric and associative anaphoric categories (types 1 and 6) may be grouped together, as the listener is relying on textual information in both. Liu and Gleason (2002) call this newly formed category the "textual" category. In both the visible and immediate situation uses (types 2 and 3), the listener "makes use of information readily available within his or her sensory reach" (p. 7); Liu and Gleason (2002) name this new category the "situation" category of use. In the two unfamiliar use categories (types 7 and 8), the listener "locates the referent using structural information" (p. 7); Liu and Gleason (2002) name this category the "structural" category.

The remaining two categories proposed by Hawkins (1978), the larger situation use based on specific knowledge and on general knowledge (types 4 and 5), are dealt with a bit differently by Liu and Gleason (2002). They retain type 5 (larger situation based on general knowledge), and rename it the "cultural" category, as it involves "information considered unique and shared

by all the people who speak the language" (p. 7). Type 4 (larger situation use based on specific knowledge) is placed into their new "situation" category along with types 2 and 3.

Following are the four categories of non-generic definite article use proposed by Liu and Gleason (2002), with examples of each (examples mine). These four categories form the framework for this thesis:

1. **Cultural**: The cultural use of *the* refers to cases in which it is used "with a noun that is a unique and well-known referent in a speech community" (p. 7). This use of *the* can also be "determined, to a large extent, by conventional practice" (p. 7). For example, *the* is used with some disease names, but not with others; or preceding some geographical terms, but not others.

    (4) (a) The Pacific Ocean is immense.

    (b) The White House is in Washington.

    (c) He caught the measles.

2. **Situation**: The situation use of *the* refers to its use "when the referent of a first-mention noun can be sensed directly or indirectly by the interlocutors or the referent is known by the members of a local community, such as the only dog in a family or the only bookstore in town" (p. 7).

    (5) (a) "Would you open the door for me please?"

    (b) The grocery store was robbed yesterday.

    (c) He already fed the dog.

3. **Structural**: The structural use of *the* is with "a first-mention noun that has a modifier" (p. 7).

(6) (a) The books that are sold here are mostly about sports.

(b) The teacher from the US speaks Azerbaijani well.

(c) I'm sitting next to the older man.

4. **Textual**: Textual use is when *the* is "used with a noun that has been previously referred to or is related to a previously mentioned noun" (p. 7).

(7) (a) I had a green shirt and a red shirt. The red shirt was nicer.

(b) He was exiled to an island. He had to live on the island for a long time.

These four categories of non-generic definite article use are similar to Huebner's (1983) four subcategories of use of the definite article in his [+SR, +HK] category of article use (see Section 2.3), with some differences. Huebner's (1983) first subcategory, that of unique referents, corresponds to Liu and Gleason's (2002) cultural category, since "what is assumed to be unique is often culturally based" (Liu & Gleason, 2002, p. 8). Huebner's (1983) second subcategory, that of physically present referents, corresponds closely to the situation category. His third subcategory, that of a previously mentioned referent, corresponds to the textual category. Huebner's (1983) fourth and final category, that of a specific referent known to the hearer, falls into Liu and Gleason's (2002) situation category. Huebner's (1983) taxonomy does not include Liu and Gleason's (2002) structural category, however, leading Liu and Gleason (2002) to claim that "our classification of *the* is both more complete and concise" (p. 8) than that of Huebner (1983).

Liu and Gleason's (2002) taxonomy of non-generic definite article use was chosen as the framework for this thesis. It is more concise than that of Hawkins (1978) and deals more

appropriately (in my opinion) with the categorization of the cultural and situation uses than does that of Hawkins (1978). Liu and Gleason (2002) mention that Hawkins (1978) included the cultural category in the situation category in his work, but that differences between them warrant separate categories. Specifically, the cultural use differs from the situation use in that cultural use does not rely on specific knowledge. Instead, in cultural use the speakers "have to assume and resort to a shared knowledge in the entire language community to make such use functional" (p. 8). Additionally, the cultural use of *the* as guided by convention makes use of rules that are much "more complex and numerous" (p. 8) than the simpler rules guiding the other uses of *the*, thus justifying a separate category.

A further reason for choosing Liu and Gleason's (2002) taxonomy is that a number of studies of L2 English definite article acquisition have been done using their taxonomy as a framework (see Section 2.5.3 for a discussion of studies using their taxonomy) and of these studies some additionally use their testing instrument (see Section 2.5.2 for a description of their testing instrument and Section 2.5.3 for a discussion of studies using their testing instrument). This thesis is intended to provide some continuity with those studies; using a common taxonomy of article use facilitates comparison with these previous studies of speakers of various L1s.

## 2.5. L2 English Definite Article Acquisition

**2.5.1. Overview of L2 English definite article acquisition**. In the course of acquiring the English article system, learners may omit articles where needed, "substitute one article for another" or "oversupply articles" (Chrabaszcz & Jiang, 2014, p. 354). A learner's L1 seems to have a significant impact on L2 English article acquisition. Some languages, such as Chinese, Japanese, Korean, and Russian, do not grammatically encode definiteness or indefiniteness and

lack articles entirely; L1 speakers of article-less languages "have been found to experience more difficulties in learning to use articles appropriately" (Lee et al., 2018, p. 80). Learners of a dual-article language (such as English with its definite and indefinite articles) who are L1 speakers of an article-less language "have to discover that there is (i) a functional category D which hosts articles, (ii) that in certain contexts D must be filled, and (iii) that definite and indefinite articles have different meanings" (Schönenberger, 2014). Lesniewska (2016) found that for learners from an article-less L1 background such as Polish "failing to provide an article is much more common than providing an incorrect one" (p. 216). However, even learners whose L1 uses an article system similar to that of English may experience difficulties as their L1 may use the articles differently. For example, in Spanish, a language with a more complex article system than that of English, but one that resembles English in many respects, the definite article is required with mass nouns, unlike in English where a zero article is required with mass nouns (Snape, Garcia-Mayo, & Güreli, 2013, p. 6-7):

(8) El oro es caro.
The gold is-PRES-SG expensive.
'Gold is expensive.'

Research shows that L2 learners do not acquire all articles or the various uses of articles at the same time, and in some cases may vary in their ability to use articles correctly according to their proficiency levels. Chrabaszcz and Jiang (2014) state that in general definite articles are acquired earlier and indefinite articles later in the L2 acquisition process, and that learners may tend to omit both types of articles in the earlier stages of learning, especially if their L1 lacks articles (pp. 353-354). Lee et al. (2018) state that "the tendency to omit articles does not seem to

improve along with proficiency" (p. 85), while overuse of articles does seem to improve as proficiency increases (Liu & Gleason, 2002). Article use, as Lesniewska (2016) points out, continues to be a problem even for advanced learners and may become "a marker of nonnativeness in otherwise proficient output" (p. 208).

Analysis of various types of L2 English article use and acquisition, according to Liu and Gleason (2002), entered a new era with the work of Huebner in 1983, drawing on Bickerton's (2016) taxonomy (which originally was published in 1981): "Huebner (1983) opened up a new avenue of research on ESL article acquisition by employing Bickerton's (1981) noun classification system" in analyzing the NPs in his data (Liu & Gleason, p. 4). Thus, he was able to go beyond the traditional focus only on "the presence or absence of articles in an obligatory context" (p. 4) and look in more detail at the various semantic functions of the articles in different contexts, and learners' acquisition of them. Following this same line of research, Liu and Gleason (2002), using their four-category classification system of non-generic definite article use based on the work of Hawkins (1978), analyzed the acquisition of L2 English non-generic definite articles.

**2.5.2. Liu and Gleason's (2002) study**. Liu and Gleason's (2002) study examined definite article use by ESL learners and whether different uses of the non-generic definite article presented differing levels of difficulty and were acquired at different stages by learners. Groups of low, intermediate, and advanced students in intensive English programs at universities in the Oklahoma City area were the participants. The majority of the students were East Asian. A 91-question testing instrument was used in which participants are given sentences which are missing the definite article and asked to supply definite articles as they see fit. 51 of the sentences

featured missing obligatory uses of *the*, while the remaining 40 were distractors. Both types of questions specifically targeted the four different non-generic uses of the definite article (discussed above). No blanks were included in the questions to avoid biasing the results. Because of the relatively open-ended nature of the instrument, two types of errors could be analyzed: omission of obligatory *the* and overuse of *the*.

Liu and Gleason's (2002) study shows that the four different non-generic uses of the definite article "present different levels of difficulty for ESL students and do not appear to be acquired at the same time" (p. 18). The differences between the four types of definite article use "suggest a hierarchy of difficulty" with "cultural use being the most difficult followed in order by textual use, structural use, and situation use" (p. 15). The situation use of *the* was acquired first, while the cultural use was acquired last (p. 18) and continued to be a problem for even advanced students (p. 14).

Liu and Gleason (2002) also discovered that the underuse of *the* (e.g., **Please give me book)* where it is required, in all four categories of usage, decreases significantly as learners progress from low to intermediate levels. Moving from intermediate to advanced levels, however, significantly lower underuse was reported only in the structural and textual categories, indicating that learners' "grasp of the cultural usage and the situation usage appears to have ceased improving significantly" (pp. 13-14). Cultural usage, they found, is nevertheless a much greater struggle for advanced ESL students than situation usage, representing a much larger share of the errors made (p. 14).

The overuses of *the* were classified by Liu and Gleason (2002) into four overuse categories (separate from the four categories of non-generic definite article use). Following are the four categories of overuse:

1. Cultural. If the reason a noun does not use a definite article is cultural, the occurrence of overuse is classified as cultural (e.g., **The Lake Michigan is a large lake in North America.*).
2. General reference. If the occurrence of overuse involves a general reference noun, which does not allow a definite article, the overuse is classified as general reference (e.g, *At the zoo, I saw several tigers. *I think that the tigers are beautiful animals*.).
3. Structural. If the article is used in the structural distractors, the overuse is classified as structural (e.g., **The people from around the world are meeting here today.*).
4. Ungrammatical. If the overuse of *the* is grammatically unacceptable, it is classified as ungrammatical (e.g., **We have two the new shoes*.)

Liu and Gleason (2002) found that the overuse of *the* increases in the cultural, general reference and structural categories as students move from low to intermediate levels and then drops as they move to more advanced levels (pp. 18-19). Nevertheless, "…results seem to suggest that overuse of *the* still remains a problem for advanced ESL students" (p. 17). The overuse of *the* in the ungrammatical category, on the other hand, "presents a very different picture. It decreased continuously, and understandably so, as the participants' English proficiency increased. In fact, very few intermediate subjects and no advanced subjects placed *the* in structurally unacceptable places" (p. 17).

### 2.5.3. Other studies based on Liu and Gleason's (2002) taxonomy.

***Wong and Quek (2007).*** In Wong and Quek's (2007) study of 50 Chinese L1 and 50 Malay L1 (both of which are article-less languages) secondary school students of various proficiency levels, an adapted version of Liu and Gleason's (2002) 91-question testing instrument was used to determine the order of acquisition of the four non-generic uses of the definite article, as well as the effect of proficiency level on the order of acquisition. They found that across the groups of both Chinese and Malay L1 learners, the acquisition order was first situation use, then structural use, then textual use, with cultural use acquired last. This confirms the findings of Liu and Gleason (2002), as does the ensuing hierarchy of difficulty: cultural>textual>structural>situational. In both studies, cultural use was clearly found to be the most difficult.

Wong and Quek (2007) also found across the groups that the overuse of *the* increased as one moved from low to intermediate level, and then dropped with advanced level learners, mirroring the findings of Liu and Gleason (2002) with regard to the cultural, general reference and structural categories. They note also that "…Chinese and Malay ESL learners tended to overuse the definite article *the* in the situation, structural, and textual contexts but underused this definite article in the cultural context" (p. 230). Interestingly, L1 Chinese learners overall had higher mean scores in structural and textual uses, while L1 Malay students had higher mean scores in the situation and cultural uses, indicating possible L1 influence.

***Garcia-Mayo (2008).*** This study of groups of elementary, low-intermediate, and advanced students in Spain used Liu and Gleason's (2002) instrument and sought to address some perceived shortcomings in Liu and Gleason's (2002) method, especially their inclusion of

mostly speakers of article-less languages and some problems in categorizing students' proficiency levels.

Similar to what Liu and Gleason (2002) found, missed obligatory uses of *the* decreased as students' proficiency increased. The hierarchy of difficulty found by Garcia-Mayo (2008) is the same as that found by Liu and Gleason (2002), with cultural being the most difficult, followed by textual, structural and situational.

Garcia-Mayo (2008) also analyzed the instances of overuse of *the*, categorizing them into the four categories used by Liu and Gleason (2002): cultural, structural, general reference, and ungrammatical. They found an across-the-board decrease in overuse of the article as proficiency increased, contrasting with the results found by Liu and Gleason (2002), who only observed such a decrease in the ungrammatical category. Garcia-Mayo (2008) attributes this different finding to the different L1 of the participants in the Spanish study (pp.560-561).

***Chrabaszcz and Jiang (2014).*** This study looked at article use among L1 Spanish and Russian speakers, seeking to address the fact that in the Liu and Gleason (2002) study, "the role of the native language was addressed in a post-hoc manner" (p. 356) and the only distinction made was between non-Indo-European and Indo-European, which ignores that fact that there is wide variation within the Indo-European family on the issue of article use. Recognizing also that ensuing studies of non-generic definite article use, such as those of Wong and Quek (2007) and Garcia-Mayo (2008) used different testing and proficiency measures, making comparison difficult, Chrabaszcz and Jiang (2014) attempted to "single out the L1 transfer effect by comparing two groups of L2 learners (L1 Russian and L1 Spanish) who are closely matched on

L2 proficiency and language learning experience" (p. 357). Russian is an article-less language while Spanish, as noted above, has an article system.

Chrabaszcz and Jiang (2014) used an elicited imitation procedure, rather than using the Liu and Gleason (2002) instrument used by other studies, although the stimulus materials were adapted in part from Liu and Gleason's (2002) instrument. Participants were divided into groups based on L1, but not on proficiency level; the researchers state that "both groups of L2 English speakers were comparable in terms of their L2 proficiency" (p. 361). They found that Russian speakers underused the article (omitted *the* in obligatory contexts) "significantly more often" than the Spanish speakers (p. 369). For the Spanish speakers, they found a similar hierarchy of difficulty as reported earlier by Liu and Gleason (2002); however, for L1 Russian speakers, "conventional and cultural uses of the nongeneric *the* also presented difficulty, though no more so than situational uses. Textual and structural uses of *the* seemed to present less difficulty for Russian speakers and yielded a higher EI score" (p. 374).

***Dikilitaş and Altay (2011).*** In this study, which is of particular interest because of the L1 of the participants, groups of upper-intermediate, advanced, and proficient learners in Turkey were tested using Liu and Gleason's (2002) instrument to determine difficulty hierarchy of the four categories of non-generic definite article usage and the effect of proficiency level on article usage. They found that at the upper-intermediate and advanced levels, participants showed an identical hierarchy of difficulty as found by Liu and Gleason (2002): cultural>textual>structural >situational, with cultural being clearly the most difficult. Among proficient participants the order was cultural>situational>textual>structural, with little difference found between the latter three categories, and much poorer performance in the cultural category (p. 193). Dikilitaş and

Altay (2011) further concluded that the structural use category was "the one which not only shows the most improvement by proficient speakers but has also been the most successfully acquired by the participants in this study" (p. 192). This finding differs from that of Liu and Gleason (2002), who found that participants were most successful in the situational use of the article. Dikilitaş and Altay (2011) also discovered a drop in overall mean scores from upper-intermediate to advanced level learners.

***Koç (2015).*** In another study of Turkish L1 students, Koç (2015) examined the non-generic definite article usage of 100 (50 intermediate and 50 advanced) English learners in a one-year intensive program in Turkey. While Koç (2015) used Liu and Gleason's (2002) classification system for non-generic definite article usage, the data was acquired using a written task, as Liu and Gleason's (2002) instrument was felt to be limited. Errors were classified in the same way as in Liu and Gleason's (2002) study: article omission errors were classified as cultural, textual, situational, or structural, following the four areas of article use, while errors of overuse were categorized as cultural, general reference, structural or ungrammatical. A general hierarchy of non-generic definite article usage was not presented.

Koç (2015) found that all groups of learners omitted more definite articles than they overused them. The intermediate group omitted more definite articles than the low-advanced group; the difference was significant only in structural and textual catgories. This is in line with Liu and Gleason's (2002) conclusion that learners' usage of *the* in the structural and textual categories improves as proficiency levels increase, while cultural and situational usages do not show much improvement.

Koç (2015) found that in overuse of the definite article, for general use contexts, intermediate students overused more than advanced. For ungrammatical use contexts, intermediate students also overused more than advanced. However, in the cultural and structural categories, advanced students overused articles more than intermediate students. This differs from Liu and Gleason's (2002) finding that cultural and structural overuse falls as students reach advanced levels. One possible explanation for the discrepancies may be differing measures of proficiency.

***Yıldırım (2015).*** In another study of Turkish L1 students, groups of first-year and fourth-year students in the ELT program at a Turkish university were tested using Liu and Gleason's (2002) 91-question instrument to determine order of difficulty and relation to proficiency level in participants' use of the four categories of non-generic definite article. Yıldırım (2015) discovered no significant differences in usage between the groups, and found a difficulty hierarchy of cultural>situational>structural>textual, which, apart from the clear difficulty of the cultural category, differs significantly from Liu and Gleason's (2002) findings.

## 2.6. The Azerbaijani Language

**2.6.1. Place in the Turkic language family**. Azerbaijani belongs to the Southwestern sub-branch of the Turkic language family. There are at least 35 Turkic languages spoken across Eurasia. The total number of speakers of Turkic languages (including L2 speakers) is at least 200 million; the best-known member of the Turkic family is Turkish, accounting for about 40% of all Turkic speakers. The Turkic languages are commonly divided into two main branches, Common Turkic and Oghur Turkic. Common Turkic is further subdivided into five sub-branches,

including the Southwestern sub-branch, into which Azerbaijani fall. The Oghur branch contains only one language.

**-Common Turkic**

-**Southwestern** (Oghuz); this is the largest sub-branch and includes Turkish, Azerbaijani, and Turkmen.

-**Northwestern** (Kipchak); includes Tatar, Kazakh, and Kyrgyz, among others.

-**Southeastern** (Karluk); includes Uzbek and Uyghur, among others

-**Northeastern** (Siberian); includes a number of minor languages spoken primarily in Siberia such as Yakut

-**Arghu** (Khalaj); includes only one divergent Turkic language of Iran

**-Oghur Turkic**; includes only the Chuvash language spoken in European Russia

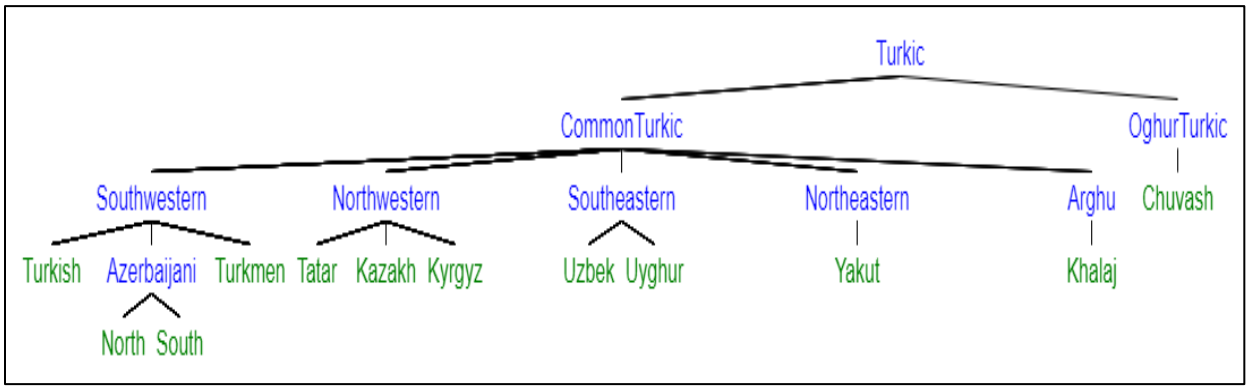


*Figure 1.* This Turkic family tree represents the relation between the various sub-branches of the Turkic language family, with some major languages indicated under each sub-branch.

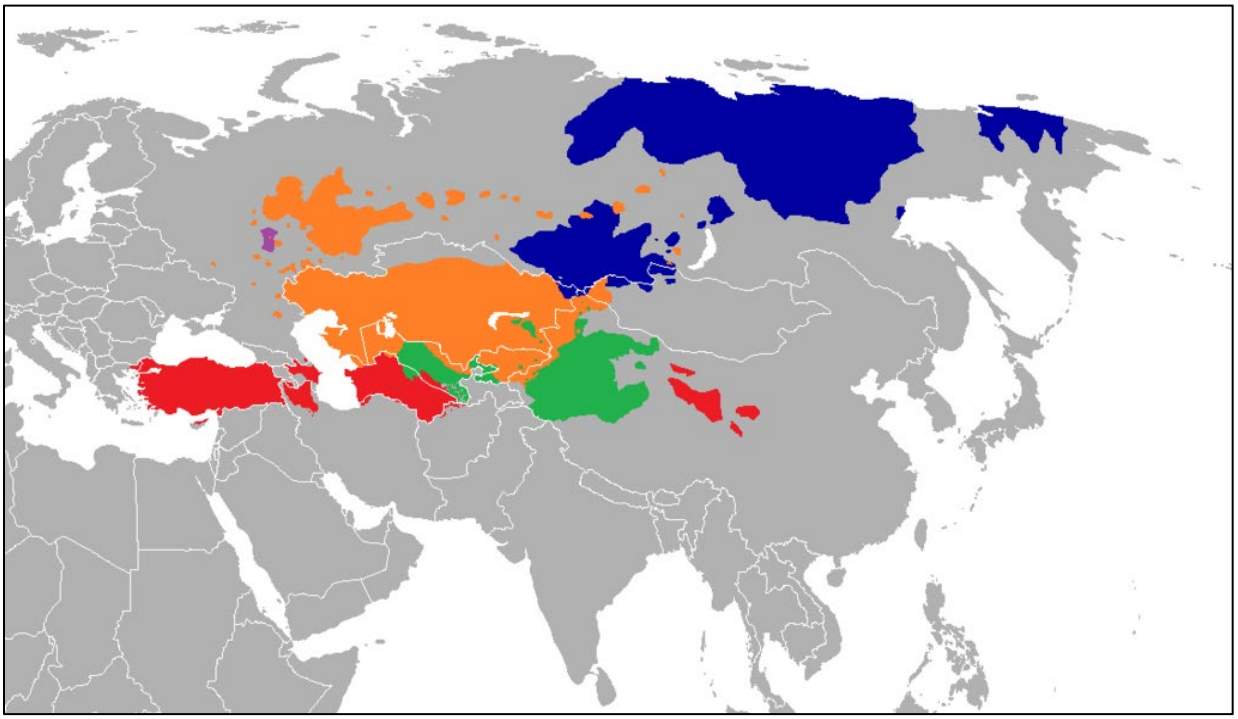

*Figure 2.* A map showing the distribution of the Turkic sub-branches across Eurasia.
*Map key*: Southwestern (red); Southeastern (green); Northwestern (orange); Oghur (purple); Northeastern (blue)
*Note: The Arghu family is not indicated, being represented only by one small language in Iran.

**2.6.2. General background of the Azerbaijani language**. The Azerbaijani language exists in two major variants. According to ethnologue.com, North Azerbaijani and South Azerbaijani are both members of the "macrolanguage" known as Azerbaijani, but have "significant differences…in phonology, lexicon, morphology, syntax and loanwords." In this thesis I will focus primarily on the North Azerbaijani variant, as the participants in my research are all located in the Republic of Azerbaijan.

North Azerbaijani is spoken primarily in the Republic of Azerbaijan, where it is the national language, by over 9 million people. It has numerous loanwords from Arabic and Persian, and, more recently, from Russian, as the Republic of Azerbaijan was formerly part of

the USSR, being known as the Azerbaijan Soviet Socialist Republic. Its speakers have had extensive contact with the Russian language and many L1 North Azerbaijani speakers also speak Russian. The Arabic script was used for North Azerbaijani until the 1920s, the Cyrillic script was used until 1992, and since 1992 a version of the Latin script closely resembling that of Turkish has been in use. Following is the Azerbaijani Latin alphabet:

Aa, Bb, Cc, Çç, Dd, Ee, Əə, Ff, Gg, Ğğ, Hh, Xx, Iı, İi, Jj, Kk, Qq, Ll, Mm, Nn, Oo, Öö, Pp, Rr, Ss, Şş, Tt, Uu, Üü, Vv, Yy, Zz

All letters are the same as those used in Turkish, with the exceptions of Əə, Xx, and Qq, which indicated sounds which are not separate phonemes in Turkish.

South Azerbaijani is spoken primarily the Iranian Azerbaijan, and in some regions of Iraq and Turkey, by over 13 million people. Although South Azerbaijani is the second most-spoken language in Iran after Farsi (Persian), it is not an official language. In Iran, South Azerbaijani is written in the Persian alphabet and is even more heavily influenced by Persian than North Azerbaijani. Most South Azerbaijani speakers in Iran are bilingual in Persian.

**2.6.3. A sketch of Azerbaijani grammar**. All of the Turkic languages share a core set of characteristics, the most prominent of which are agglutination, vowel harmony, a lack of grammatical gender, and SOV word order.

According to Dehghani (2018), "Azeri morphology is primarily agglutinative and suffixing" (p. 11), making extensive use of both derivational suffixes and inflectional suffixes (including case, possessive, plural, tense, mood, and aspect suffixes, among others). Azerbaijani nouns are inflected for case and number; there are six cases: the nominative (which requires no suffix), the accusative, the dative, the genitive, the locative, and the ablative. Suffixes must

follow a strict order; for example, plural suffixes always come first, followed by possessive suffixes, followed by case suffixes. For example:

(9) kitɑb-lɑr-ɯm-dɑn

book – PL – my – ABL

'from my books'

The rules of vowel harmony in Azerbaijani affect the form that a suffix will take, and all the suffixes must "harmonize" with the preceding root vowel (and with each other). Each suffix, then, may have either two or four possible forms based on the backness and roundness of the vowels in question (Dehghani, 2018). There are two types of suffix in Azerbaijani; suffixes of the so-called *-æ* type version have two possible forms, while suffixes of the so-called *-i* type version have four possible forms (Allazov, 2018). The following chart outlines which suffix vowels must follow which root vowels.

Table 1

*Vowel Harmony Chart 1*

| Last vowel of the root word: | *-æ* type suffixes must use: | *-i* type suffixes must use: |
|---|---|---|
| -ɑ, -ɯ | -ɑ | -ɯ |
| -o, -u | -ɑ | -u |
| -e, -æ, -i | -æ | -i |
| -ø, -y | -æ | -y |

For example, the suffix indicating a meaning of "without" is an *-i* type suffix and may have four forms (*-suz, -siz, -suz, -syz)*, with the vowel of the suffix agreeing with the preceding root vowel (examples from Dehghani, 2018, p.13):

Table 2

*Vowel Harmony Chart 2*

| | | | |
|---|---|---|---|
| ɑd | "name" | ɑd-sɯz | "unknown" |
| ev | "home" | ev-siz | "homeless" |
| gydʒ | "power" | gydʒ-syz | "powerless" |
| pul | "money | pul-suz | "poor" |

Azerbaijani does not mark grammatical gender in any way. A single pronoun, *o*, may mean *he*, *she* or *it*.

Word order in Azerbaijani (as in all Turkic languages) is Subject-Object-Verb (SOV):

(10) Mæn sæn-i øp-yr-æm.

I you-ACC kiss-PRES-1PS.

'I kiss you.'

**2.6.4. Definiteness in Azerbaijani**. Azerbaijani lacks a definite article, as does Turkish. Definiteness, however, can be expressed in Azerbaijani in ways other than via an overt article. In the subject position a bare noun such as *kitab* (book) carries a definite meaning:

(11) Kitɑb tʃox kitʃik-diɾ.

Book very small-is.

'The book is very small.'

In the direct object position, definiteness is expressed by means of an accusative suffix (expressed as *-ɯ, -i, -u, -y* after a consonant, or as *-nɯ, -ni, -nu, -ny* after a vowel) while indefiniteness is expressed by the lack of a suffix:

(12) a. Mæn kitɑb oxu-muʃɑm
I book read-PAST.
'I read a book.'

b. Mæn kitɑb-ɯ oxu-muʃɑm
I book-ACC read-PAST.
'I read the book.'

(13) a. Annɑ ʃyʃæ gøtyr-yr
Anna bottle take-PRESENT.
'Anna is taking a bottle.'

b. Annɑ ʃyʃæ-ni gøtyr-yr
Anna bottle-ACC take-PRESENT.
'Anna is taking the bottle.'

As Snape et al. (2013) explain in regard to Turkish, a bare nominal is "interpreted as definite in subject position" but "interpreted as indefinite in direct object position" (p. 7). Similarly, in Azerbaijani the accusative case "surfaces as zero when the object is indefinite" and "the non-zero realization of the accusative case marking in Azeri is restricted to definite noun phrases" (Dehghani, 2018, pp. 51-52).

## 2.7. Research Question

This study considers L2 English non-generic use of definite articles by L1 Azerbaijani speakers, using the Liu and Gleason (2002) taxonomy and an adapted (shortened) version of the testing instrument created by them. Only written data was collected, both because of feasibility concerns (the researcher is not located in the Republic of Azerbaijan) and to maintain continuity with previous studies (Dikilitaş & Altay, 2011; Garcia-Mayo, 2008; Wong & Quek, 2007; Yıldırım, 2015) which have utilized Liu and Gleason's (2002) written testing instrument, as opposed to oral data collection. The primary research question I investigated was:

> Are each of the four types of English non-generic definite article use (as proposed by Liu and Gleason, 2002) equally difficult for L1 Azerbaijani speakers, and if not, what is the hierarchy of difficulty?

## 2.8. Conclusion

Having looked at a number of different ways of approaching the English article system, including the distinction drawn by Ionin et al. (2004) between definiteness and specificity, and the taxonomies of Bickerton (2016), Huebner (1983), Hawkins (1978) and Liu and Gleason (2002), as well as some of the issues involved in L2 acquisition of English articles, we have settled on the taxonomy of Liu and Gleason (2002) as the framework of this thesis. In using an adapted version of Liu and Gleason's (2002) testing instrument, this thesis follows their line of research on L2 acquisition of English non-generic definite articles and that of other researchers who have used their taxonomy and testing instrument to study the acquisition of non-generic definite articles by speakers of various L1s.

We have seen that the L1 group on which this thesis is focused, L1 Azerbaijani speakers, shares a language which lacks definite articles, instead expressing definiteness through word order and accusative suffixes. L1 Azerbaijani speakers thus come to the task of acquiring English definite articles with a possible handicap, as do speakers of other article-less languages such as Russian or Korean, as opposed to speakers of languages such as Spanish which have a system of articles resembling that of English in some ways. This may especially be true in regard to the so-called "cultural" category of non-generic definite article use proposed by Liu and Gleason (2002), which has been shown to be significantly more difficult for NNSs to acquire than other categories of use. We turn now to a discussion of the methods used to gather data and the subjects from which the data was gathered.

# Chapter 3: Method

## 3.1. Introduction

Data for this study was collected from L1 Azerbaijani-speaking respondents via email, WhatsApp, and Facebook messenger. Two different instruments were used: a 47-question article-use test was given, along with a brief writing assignment, which was used to determine the relative proficiency levels of the respondents for data analysis purposes.

## 3. 2. Participants

There were 10 participants who are native Azerbaijani speakers. All but one currently live in the Republic of Azerbaijan; one of them has been living in the US for one year. The participants are intermediate and advanced adult EFL learners. Learners who are beginners were not selected for this study. The participants were recruited via online contacts with an English instructor and English learners in the Republic of Azerbaijan. Students under the age of 18 did not participate in this study. Potential participants were directly asked to specify their age before participating via a brief questionnaire at the beginning of the article-use test.

The instrument was also administered to a control group of three native speakers of English, in order to identify and eliminate unclear sentences or sentences in which more than one use of *the* could be acceptable. The English L1 speakers did not submit the writing activity for proficiency level determination. These participants were all over the age of 18 and have all been recruited from among my acquaintances in the United States. All have a minimum of one year of university-level education.

A pilot study was completed using three Azerbaijani L1 speakers over the age of 18 who are at the intermediate or advanced level. They were recruited from among my online

acquaintances in the Republic of Azerbaijan. Additionally, the pilot study included three English L1 speakers (who completed only the article-usage test) recruited from among my acquaintances in the United States. All of them had a minimum of one year of university-level education.

### 3.3. Materials

**3.3.1. Article usage test.** An article usage test adapted from the 91-question testing instrument created by Liu and Gleason (2002) was used. For my shortened version of the original testing instrument, I have chosen a total of 47 questions from the original 91. These have not been altered in any way from their original form. The original instrument was shortened in order to avoid excessive burden on the respondents and decrease attrition rates.

In the original testing instrument, Liu and Gleason (2002) included 51 sentences containing a total of 60 instances of missing obligatory *the*, with 15 instances from each category of non-generic definite article use. Additionally, they included 40 distractors, with 10 from each category, for a total of 91 questions.

My adapted version includes 27 sentences containing 32 missed obligatory uses of *the*, with 8 from each category. Additionally, I have included 20 distractors, with 5 from each category, for a total of 47 questions.

In this test, no blanks are used; as Liu and Gleason (2002) explain, "Our rationale for not including blanks was that if we did, some of the students, especially the low-level students, might fill every blank with the, making the data very unreliable…" (p. 10). Not including blanks also gives participants freedom to use articles in perhaps unexpected ways. Following are example items from each category, which are included in my shortened version (Liu & Gleason, 2002):

1. President of the United States lives in White House. (Cultural)
2. A man says to his wife at the breakfast table, “Can you pass me newspaper?” (Situational)
3. Man I met in New York later became my husband. (Structural)
4. Fred bought a car on Monday. On Wednesday, he crashed car. (Textual)

At the beginning of the testing instrument, students were asked to state their native language and the number of years they have studied English, as in the original Liu and Gleason (2002) test; this ensured that only L1 Azerbaijani subjects are participating and assisted in removing beginning learners from the study. Additionally, participants were asked to specify their age in order to verify that no one under 18 was included in the research. Directions were given, as in the original, as follows: “In some of the following sentences, the definite article *the* is missing. Please read the following sentences carefully and insert the article *the* wherever you believe necessary.”

I decided to use an adapted version of the Liu and Gleason (2002) testing instrument because my research question is very similar to theirs, excepting my focus on L1 Azerbaijani speakers. This also provides a point of comparison with the results of researchers who have used the Liu and Gleason (2002) testing instrument to study the non-generic definite article use of speakers of various L1s (e.g., Dikilitaş & Altay, 2011; Wong & Quek, 2007). I decided to shorten their testing instrument significantly, however, as I felt that a 91-question test may be considered excessively long and burdensome. I did retain similar proportions of questions in each of the four categories as well as distractors in each of the four categories.

A committee of three members, including one native speaker of British English and two of American English (different than the native speaker control group mentioned in section 3.2), were available to assist me in determining the grammaticality of any uncertain responses to the article usage test; no such uncertain responses were encountered, however.

The article usage test may be found in the Appendix.

**3.3.2. The writing assignment.** Participants were asked to complete a brief writing assignment to determine their general English proficiency levels relative to one another. This task could be completed before or after the article usage test, and was submitted via email, WhatsApp, or Facebook messenger. The participants wrote a description of their city, answering the prompt, "Please tell me about your city." The participants were asked to write a minimum of around 200 words and not to use any assistance from online helps, dictionaries, or other people.

The text generated by each participant were pasted into the *VocabProfiler – Compleat* on Tom Cobb's *Compleat Lexical Tutor* website (Simons & Fenning, 2018). The resulting percentages of tokens used by the participant which fell into the K1 (1-1000 most common English words) category versus the K2 (1001-2000 most common words) or higher categories, with proper nouns "classed offlist," were analyzed. Special attention was paid to the percentages of K2 or higher category words used, and all respondents ranked according to those percentages. The rankings were observed to fall naturally into two categories, separated by three full percentage points. Four respondents (who made more extensive use of K2 or higher words) were found to fall within the resulting relatively "advanced" category, and six (who made less extensive use of K2 or higher words) were found to fall within the relatively "intermediate" category.

### 3.4. Procedure

As the entire study was conducted online, participants were able to complete the two components on their own time in any order they chose. They were asked not to use any online reference helps such as dictionaries, translators, etc., or to seek help from other L1 or L2 English speakers.

**3.4.1. The article usage test.** Participants took the article usage test via email, WhatsApp, or Facebook messenger. There was no time limit. The test was done on the participant's own device, on their own time, and the results were returned to the researcher by email.

**3.4.2. The writing assignment.** Participants completed the written task via email, WhatsApp, or Facebook messenger. There was no time limit. Participants were requested to write a minimum of about 200 words and to answer the prompt, "Please tell me about your city." I then pasted the text into the *VocabProfiler – Compleat* on Tom Cobb's *Compleat Lexical Tutor* website (https://www.lextutor.ca/vp/comp/) and analyzed the resulting percentages of tokens in the K1 versus words in the K2 + categories.

### 3.5 Conclusion

In analyzing the results of the 10 L1 Azerbaijani respondents' efforts on the 47-question article usage test adapted from Liu and Gleason (2002), and placing them as relatively more or less advanced using the writing task (with the help of the *Compleat Lexical Tutor*), I was able to answer the research question (see Section 2.7) posed, compare my findings with those of other researchers such as Wong and Quek (2007), Garcia-Mayo (2008), Dikilitaş and Altay (2011),

and Yıldırım (2015), and gain some insight into the effect of proficiency level on L2 article difficulty.

# Chapter 4: Results

## 4.1. Introduction

An attempt was made to follow the procedure used by Liu and Gleason (2002) in the treatment of the data. Errors made by participants in the sentences of the testing instrument were tabulated. The data acquired was divided into a missed obligatory *the* category and an overuse of *the* category. For each participant, the number of obligatory uses of *the* which were missed was tabulated and subdivided into the four categories of non-generic definite article use proposed by Liu and Gleason (2002). The number of overuses of *the* was also tabulated for each participant, and then was subdivided into Liu and Gleason's (2002) four categories of overuse (see Section 2.5.2), in keeping with the procedure used in their study.

## 4.2. Tabulation of the Data

Before looking at individual tests, I created a master list of all sentences on the testing instrument which were missing obligatory uses of *the* and categorized each as cultural, situation, structural, or textual use. Then for each individual test, I first went through and noted all instances of missed obligatory usage. I then categorized each instance as either cultural, situation, structural, or textual, referring to the master list. For example, in one test, the following sentence was left uncorrected, with no *the* added (the number immediately preceding the first word of the sentence refers to the number of the sentence in the testing instrument):

(14) 35. Mississippi river runs through Louisiana.

This was marked as an instance of missed obligatory cultural usage, since the reason *the* is required with names of rivers is strictly cultural, determined by convention (see section 2.4.3).

In another test, the following sentence was uncorrected:

(15) 23. A man says to his wife at the breakfast table, “Can you pass me newspaper?”

This was recorded as an instance of missed obligatory situation use, since the use of *the* before *newspaper* is required due to the fact that the referent can be sensed directly or indirectly by the interlocutors (see section 2.4.3).

In another test, the following sentence was left uncorrected:

(16) 14. Blue car across the road is very suspicious.

This was determined to be a case of missed obligatory structural use, since the use of *the* before *blue car* is required due to the presence of the modifier *across the road* (see section 2.4.3).

Finally, in another test, the second sentence in the following was not corrected as required:

(17) 3. We rented a boat last summer at a lake. Unfortunately, boat hit another boat and sank.

This was marked as a case of missed obligatory textual use, since the use of *the* before the first instance of the word *boat* in the second sentence is required because the boat in question was mentioned in the first sentence.

Upon completion of the tabulation of results for missed obligatory uses, I examined the testing instrument in question again in order to record all instances of overuse of *the*. I categorized each case of overuse as one of the four categories of overuse proposed by Liu and Gleason (2002), which were cultural, structural, general reference, and ungrammatical. The categorization had to be done on a case-by-case basis, as overuse was unpredictable. For example, the following sentence was encountered in several tests:

(18) 5. The mother says to her children, “Come on, it’s time to go to the Grandma’s house.”

This was recorded as an instance of cultural overuse, since the reason *the* is not acceptable before *Grandma’s* here is that its use would go against convention (see section 2.5.2).

This sentence was encountered in one test:

(19) 17. I’ve heard of the parents who don’t give their children enough to eat.

This was considered to be an instance of structural overuse, since it occurred in one of the structural distractors, a sentence in which a modifier is present but where *the* is not in fact required (see section 2.5.2).

The following sentence was encountered:

(20) 19. The things of beauty always bring great joy.

This was tabulated as a case of general reference overuse; the error involves inappropriate use of *the* in reference to a general category, *things of beauty* (see section 2.5.2).

Finally, very occasionally sentences of this type were encountered:

(21) 3. We rented a boat last summer at a lake. Unfortunately, the boat hit another the boat and sank.

The overuse of *the* before the second occurrence of the word *boat* in the second sentence was categorized as ungrammatical overuse, since its use after *another* is grammatically unacceptable (see section 2.5.2).

The total number of each of the four types of missed obligatory use and each of the four types of overuse were entered into a spreadsheet for each participant in preparation for further analysis.

### 4.3. Types of Tests Conducted and Results

For both missed obligatory uses and overuses of *the*, I calculated the mean of the missed uses in each of the four categories, along with the standard deviation, and then conducted a MANOVA using the subject's proficiency level (as determined by the writing assignment) as an independent variable. In Liu and Gleason's (2002) study, the MANOVA uncovered significant differences in performance in the four categories of non-generic article use (both for missed obligatory use and overuse), supporting their hypothesis that the four categories were not equally difficult for learners.

The ten L1 Azerbaijani participants were found to naturally fall into intermediate and advanced groups, with six participants in the intermediate group and four in the advanced group, based on the analysis of their writing sample and the resulting percentage of non-K1 words used. The results for missed obligatory use demonstrate that the advanced group consistently made fewer mistakes across all four categories than the intermediate group; however, the results also show an identical order of difficulty for both the intermediate and the advanced groups.

Table 3

*Means and Standard Deviations for Missed Obligatory use of the for Both Groups*

| **Group** | ***n*** | **Cultural** | **Situation** | **Structural** | **Textual** |
|---|---|---|---|---|---|
| Intermediate | 6 | | | | |
| M | | 3.17 | 1.17 | 3.17 | 2.17 |
| SD | | 1.46 | 1.86 | 1.68 | 1.86 |
| Advanced | 4 | | | | |
| M | | 2.5 | 0.75 | 2.5 | 1.75 |
| SD | | 1.8 | 1.66 | 1.66 | 0.83 |

As can be seen in Table 3 above, the intermediate group showed identical means for cultural and structural use, and the advanced group showed identical means for cultural and structural as well. These two categories of use presented the largest number of errors for both groups, and appeared to be of equal difficulty for both groups. Textual use followed with a smaller number of errors for both groups, with situation use presenting the fewest problems for both groups. These results may be seen more clearly in Figure 3 below.

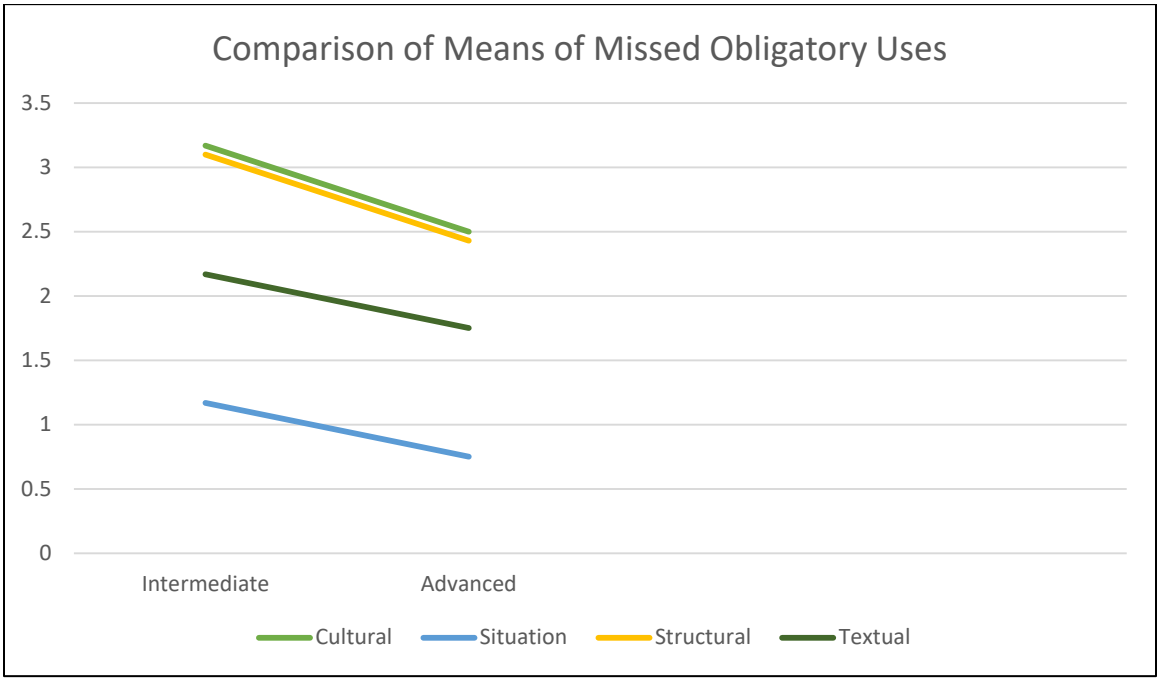


*Figure 3.* Comparison of means of missed obligatory use by category.

As seen in Figure 3 above, the cultural and structural use caused the most trouble for both intermediate and advanced participants, with textual presenting less of a problem and situation use the least difficult for both groups.

In considering the results of the study for overuse of *the*, our results present a different picture than we saw with missed obligatory use. Advanced participants did make fewer errors in all categories than intermediate participants; however, overuse in the cultural category presents a

significantly greater mean number of errors than any other category for both groups, unlike what we saw in the missed obligatory data. Structural and general reference overuse together comprise the next most difficult categories, with equal means for the intermediate participants and only a slight difference between their means for the advanced (with structural being slightly higher than general reference). Ungrammatical overuse proved to be only a small problem for some intermediate respondents, with no errors at all in this category for advanced respondents.

Table 4

*Means and Standard Deviations for Overuse of the for Both Groups*

| **Group** | ***n*** | **Cultural** | **General Ref.** | **Structural** | **Ungrammatical** |
|---|---|---|---|---|---|
| Intermediate | 6 | | | | |
| M | | 8.33 | 1.5 | 1.5 | 0.67 |
| SD | | 4.46 | 1.5 | 0.96 | 1.11 |
| Advanced | 4 | | | | |
| M | | 4.0 | 0.75 | 1.0 | 0 |
| SD | | 2.55 | 1.3 | 1.22 | 0 |

The relationship between the categories of overuse for intermediate and advanced respondents may be seen more clearly below in Figure 4. Again, as we saw in the missed obligatory usage, the order of difficulty is very similar for both intermediate and advanced respondents, with only a slight difference between the structural and general reference categories for the advanced learners, whereas for intermediate those two categories were identical.

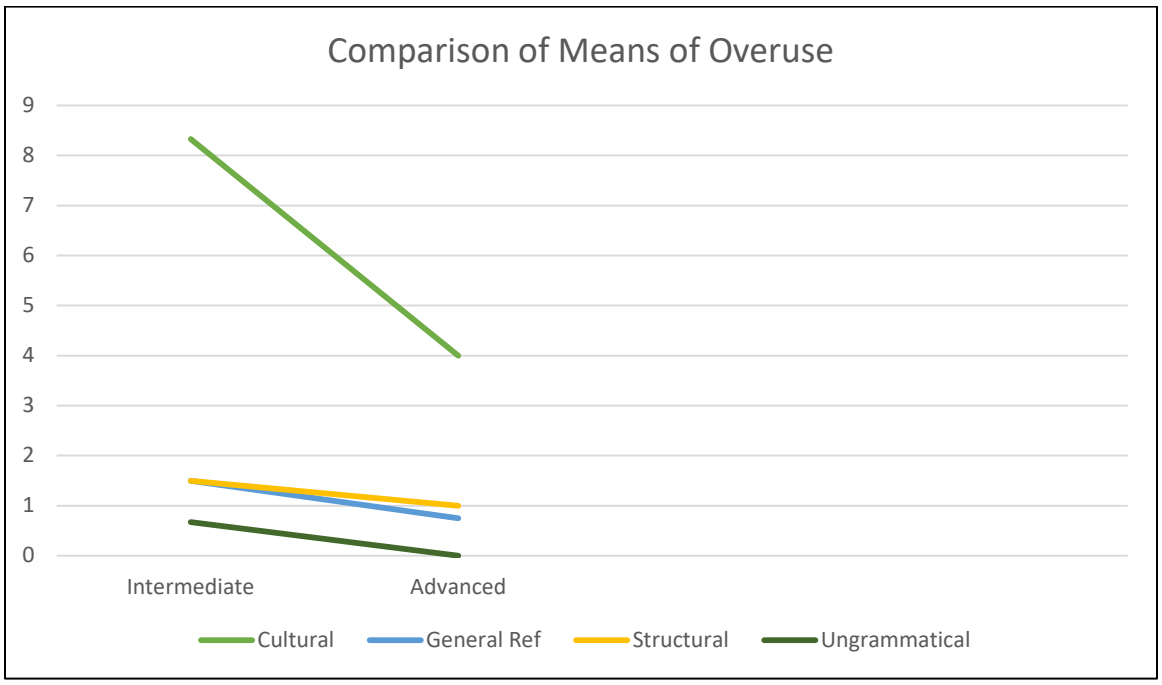


*Figure 4.* Comparison of means of overuse of *the* by category.

As may be clearly seen above in Figure 4, cultural overuse presented a significantly greater challenge to both groups of participants than the other categories, with general reference and structural following well behind, and ungrammatical usage presenting the least difficulty.

In my study, a multivariate MANOVA test was conducted across intermediate and advanced levels, comprising all four categories of missed obligatory usage and all four categories of overuse, but due to the relatively low number of participants in the study (four advanced and six intermediate), the results were statistically insignificant. A univariate MANOVA for each of the different categories was therefore not conducted.

Table 5

*Results of MANOVA Multivariate Tests across Two Levels in All Four Categories of Missed Obligatory Usage and All Four Categories of Overuse*

| Use Type | *df* | Λ | *F* |
|---|---|---|---|
| Missed obligatory | 4 | 0.936 | 0.086 |
| Overuse | 4 | 0.713 | 0.503 |

The post-hoc Tukey test and pairwise t-tests which followed the MANOVA tests in the Liu and Gleason (2002) study were likewise not conducted in my study due to the fact that the results would be deemed statistically insignificant due to the low number of participants.

In summary, a calculation of the mean number of errors made by the six intermediate and four advanced participants in both the missed obligatory categories and overuse categories revealed that

a) advanced participants made fewer errors across the board in all categories of missed obligatory use and overuse than intermediate participants;
b) intermediate and advanced participants showed the same order of difficulty for missed obligatory use, which was cultural and structural both ranking equally as most difficult, followed by textual, with situation following as least difficult;
c) intermediate and advanced participants showed nearly the same order of difficulty (with only very slight difference) for overuse, which was cultural being clearly the most difficult, followed by structural and general reference, which ranked nearly equally, with ungrammatical overuse as the least difficult.

The MANOVA testing was limited to multivariate due to the low number of respondents in this study.

**4.4. Answers to Research Questions**

In analyzing the data and attempting to answer my primary research question ("Are each of the four types of English non-generic definite article use (as proposed by Liu & Gleason, 2002) equally difficult for L1 Azerbaijani speakers?"), my data clearly shows that the answer is ***no***; not all categories of non-generic definite article use present the same difficulty to learners. As we see in Figures 3 and 4, for example, although advanced participants made fewer errors than intermediate participants across the board, cultural and structural use in the missed obligatory category and cultural use in the overuse category clearly presented greater difficulty than other uses of *the* for both intermediate and advanced participants.

In answering the secondary research question ("…and if not, what is the hierarchy of difficulty?"), we can consider the comparison of the means of the various categories of use for both the missed obligatory uses of the definite article and for overuses of the definite article. For the missed obligatory uses, cultural and structural use I found to be the most difficult (each with exactly the same mean) for both intermediate and advanced participants. An example of cultural missed obligatory use from my testing instrument was:

(22) 40. I'm sick. I've come down with flu.

In this case, the word *flu* requires a definite article for no reason other than convention; other diseases may not require a definite article. An example of structural missed obligatory use was:

(23) 12. Do you know pilot who flies this airplane?

This example requires a definite article before *pilot* because of the modifier which follows.

Cultural and structural were followed by textual missed obligatory use as the next most difficult category for both groups of participants. An example of textual missed obligatory use from my instrument was:

(24) 4. My mother has a white dog and a black dog. White dog is taller than black one.

In this example there are two textual missing obligatory uses; *the* is required before the words *white* and *black* in the second sentence, since the referents were introduced in the first sentence. Finally, situational use was found to be the easiest for learners in both the intermediate and advanced groups. An example of situation missed obligatory use from my instrument was:

(25) 47. In their living room at bedtime, the mother says to the children, "Turn off television."

In this example, *the* is required before television since the referent is known to both speaker and hearers.

For overuse, where the means of the four categories of overuse were calculated, cultural was by far the most difficult category for learners. An example of cultural overuse, using an example from my testing instrument, was:

(26) 35. The Mississippi river runs through the Louisiana.

In this example, *the* was unnecessarily used before the name of a state, going against cultural convention. Cultural overuse was followed by structural and general reference overuse, where the intermediate learners' mean was identical for the two categories, and the advanced learners' mean was only slightly higher (i.e., greater difficulty) for structural over general reference. An example of structural overuse from my testing instrument was:

(27) 18. The people from around the world are meeting here today.

In this example, *the* is incorrectly inserted before *people*; this is a structural distractor in which there is a modifier present but the definite article is nevertheless not used. An example of general reference overuse from the testing instrument was:

(28) 6. I watched several old movies last weekend. I enjoy watching the old movies.

In this example *the* is overused before *old movies*, a general reference NP. Finally, the ungrammatical overuse category presented the least difficulty for learners in both the intermediate and advanced groups. An example of ungrammatical overuse was:

(29) 22. Water in this the glass is dirty.

In the above example, *the* is inserted after *this* and before the noun in a violation of English grammar.

In summary, the order of difficulty suggested by my data was (separation by a slash indicates identical means):

| | |
|---|---|
| Missed Obligatory (Intermediate): | Cultural/Structural > Textual > Situational |
| Missed Obligatory (Advanced): | Cultural/Structural > Textual > Situational |
| Overuse (Intermediate): | Cultural > Structural/General Reference > |
| UngrammaticalOveruse (Advanced): | Cultural > Structural > General Reference > Ungrammatical |

The order of difficulty was identical for both groups for missed obligatory usage, and differed only slightly between groups for overuse.

# Chapter 5: Discussion

## 5.1. Introduction

The data show that not all types of L2 English non-generic definite article usage are equally difficult for L1 Azerbaijani speakers. For the missed obligatory uses, the cultural usage type did not emerge as clearly the most difficult, as expected; rather, cultural and structural showed exactly the same means for both intermediate and advanced groups. Textual and situational, two categories which perhaps could be more easily reduced to a few rules which may be memorized and mastered by foreign students, clearly presented less difficulty. The means of the errors made by advanced learners was found to be somewhat less than that of the intermediate learners for each of the four categories, as we would expect with increasing general English proficiency.

## 5.2. Comparison with Liu and Gleason's (2002) Findings

**5.2.1 Missed obligatory uses.** Liu and Gleason's (2002) order of difficulty was cultural>textual>structural>situational, a finding confirmed by Wong and Quek (2007) and Garcia-Mayo (2008). This finding differs from my results in that the order of the textual and structural categories is reversed in the Liu and Gleason (2002) study, and in that there is a significant difference in the difficulty of the cultural and structural categories in their study, whereas in mine the two were found to be identical. A clearly greater level of difficulty for the cultural category of use was demonstrated in the Liu and Gleason (2002) study. They found that this was generally the last type of non-generic definite article use to be mastered (p. 18) and continued to cause problems for advanced students (p. 14), a finding confirmed by other subsequent researchers using the Liu and Gleason (2002) classification system, such as Wong

and Quek (2007), Garcia-Mayo (2008), Dikilitaş and Altay (2011), and Yıldırım (2015). I also found that the cultural use continued to present difficulties to advanced students, but the structural use continued to be a problem as well for L1 Azerbaijani learners.

In Liu and Gleason's (2002) study, there was a significant decrease between intermediate and advanced groups only in structural and textual uses. This suggests that "their grasp of the cultural usage and situation usage appears to have ceased improving significantly" (pp. 13-14). My data, on the other hand, shows a proportionally similar drop in the number of errors made between the intermediate and advanced groups for all four categories (see Figure 3), indicating an improving ability to use all four categories correctly. Liu and Gleason (2002) did show, however, that the frequency of errors in the situation category was so low as to be of negligible importance. They concluded that "it is probably safe to say that the cultural use of *the* is still a difficult problem for advanced ESL students to wrestle with," but "it is perhaps not the case with its situation use" (p. 14). I similarly found that the number of errors made by both groups, and especially by the advanced group, was very low (see Figure 3).

**5.2.2. Overuse.** The overuse of *the* presents a slightly different picture. Liu and Gleason's (2002) results "seem to suggest that overuse of *the* still remains a problem for advanced ESL students" (p. 17). My data suggested a very much greater difficulty level for the cultural category in the overuse of *the*, as opposed to the other three categories of overuse. There was a significant drop in the mean number of errors made between the intermediate and advanced group in cultural overuse, but the number of errors remained significant even in the advanced group. I found that the mean of the other three categories of overuse were much less significant, and each dropped between the intermediate and advanced groups. The

ungrammatical category especially presented few problems to intermediate students and none at all for advanced learners in my study; similarly, Liu and Gleason (200) found that the patterns of overuse of *the* in the ungrammatical category behaved very differently from the other categories, in that there was a continuous decrease in ungrammatical uses of *the* as proficiency increased.

**5.3. Conclusion**

My data from L1 Azerbaijani speakers showed some similarities and some differences with the findings of Liu and Gleason (2002) and other researchers. Cultural usage did emerge as the most difficult (as in Liu & Gleason's findings), having by far the highest mean number of errors in the overuse category, and, along with structural use, as the most difficult in the missed obligatory category. One unexpected finding was that cultural use and structural use shared exactly the same means in the missed obligatory category. I did not expect that the cultural use category would share its position as most difficult category with another category of non-generic article usage. Furthermore, the fact that the structural use category presented such a high level of difficulty was unexpected, since in Liu and Gleason's (2002) findings structural usage ranked as third most difficult out of the four. Not unexpected was my finding that situational usage presented the least difficulty in the missed obligatory category, and ungrammatical usage the least difficulty in the overuse category, as those findings mirror those of Liu and Gleason (2002). It would seem that in terms of pedagogical approaches to the teaching of L2 English article use to L1 Azerbaijani speakers, it will be necessary to focus special attention not only on the cultural use but also on the structural use of the non-generic definite article, especially as students reach intermediate and even advanced stages of acquisition. Situational use, on the other hand, as well as ungrammatical overuse, can be expected to improve as students acquire L2 English without

special attention, likely due to the fact that rules regarding these types of usage are limited in number and relatively simple to follow.

## Chapter 6: Conclusion, Limitations, and Future Research

My results show an across-the-board decrease in number of errors made by advanced participants as compared with intermediate participants. The four categories of non-generic definite article use presented in Liu and Gleason's (2002) study did not present equal difficulty to L1 Azerbaijani participants; the same order of difficulty was visible in both the intermediate group and the advanced group: cultural/structural > textual > situation. Likewise, the four categories of non-generic definite article overuse utilized by Liu and Gleason (2002) in their study did not present equal difficulty to participants in my study, with the order of difficulty observed to be cultural > structural/general reference > ungrammatical, with only a slight different between the structural and general reference categories for the advanced participants, and no difference for the intermediate participants.

My results demonstrate the persistence of definite article difficulties even among relatively advanced learners, as well as the fact that not all types of article use present equal difficulty, thus justifying the breaking down of non-generic definite article use into the separate categories of use proposed by Liu and Gleason (2002). In the case of L1 Azerbaijani learners of English, several practical applications of the findings of this study to teaching of English in the Republic of Azerbaijan may be extracted. The first is the realization among teachers that there are in fact different categories of non-generic definite article use, some of which are more explicitly governed by general rules than others. For example, while situation use is generally governed by simple rules which are more easily memorized and applied, cultural use is not. The second application would be the realization that article use, while a very fundamental and seemingly elementary feature of English grammar, continues to present problems even to

advanced learners (including perhaps even non-native teachers of English). The third is the awareness that cultural and structural use of the non-generic definite article present the most difficulty to L1 Azerbaijani students, and therefore could be drilled more extensively in a focused way than the other less difficult uses.

My findings in this study confirm those of other researchers such as Liu and Gleason (2002), Wong and Quek (2007), Garcia-Mayo (2008), Chrabaszcz and Jiang (2014), Dikilitaş and Altay (2011), Koç (2015), and Yıldırım (2015), that not all types of non-generic definite article usage presents the same level of difficulty to learners. The fact that the order of difficulty which I found differs from those of some other researchers, and that the order varies to some degree among other researchers cited, would seem to point to the fact that the specific order of difficulty experienced by the learner would vary by L1, among other possible factors.

Several limitations of this study should be noted. First, the number of respondents was significantly lower than expected. It proved quite challenging to find sufficient numbers of Azerbaijani speakers willing to take the time and effort to complete a 47-question testing instrument and a 200-word essay, especially via internet from a stranger, in spite of much-appreciated assistance from a teacher friend in Azerbaijan. As a result, I was only able to gather ten complete responses; I had aimed for at least thirty. It would be valuable in the future to perform a similar study with a much larger number of respondents; this would most likely need to be conducted in a classroom setting in Azerbaijan.

A second limitation was the somewhat imprecise method of determining relative English proficiency levels used in this study. The decision to use an essay and examine the percentage of tokens falling beyond the K1 and proper noun categories was made in an effort to avoid giving

potential respondents an additional test to take; however the use of a standardized instrument such as the TOEFL to determine level, or using respondents who are all in different classes at the same school, for example, may result in a more exact measure of relative general English proficiency. This method of determining level also likely makes it more difficult to compare results with those of earlier researchers such as Liu and Gleason (2002). A related issue potentially complicating comparison to earlier studies is that I included only two proficiency levels in my study (intermediate and advanced), whereas Liu and Gleason (2002) included three (beginner, intermediate and advanced).

Future investigations on L1 Azerbaijani non-generic definite article use in L2 English would need to be done using significantly larger test groups, a task which most likely is best accomplished in a classroom setting in the Republic of Azerbaijan. With a larger number of participants, the statistical tests run by Liu and Gleason (2002), such as the univariate MANOVA, Tukey tests, and pairwise t-tests, would be statistically significant and would enable more accurate comparison to their data and that of subsequent researchers using their model of non-generic article use. It is also likely that a more precise hierarchy of difficulty for L1 Azerbaijani speakers could be established with a significantly larger number of participants; for example, the means of cultural and structural missed obligatory use, while identical in my findings, would likely be seen to differ if a larger number of participants were contributing data. Similarly, in my findings structural and general reference overuse means were identical for intermediate participants and differed only slightly for advanced; a larger number of respondents would likely enable a clearer distinction between and ranking of the structural and general reference categories in an order of difficulty.

## References

Allazov, E. (2018). *Azerbaijani for beginners*. Middletown, DE: Createspace.

Bickerton, D. (2016). *Roots of language*. Berlin: Language Science Press.

Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, *30*(3), 351-379.

Dehghani, Y. (2018). *A grammar of Azeri*. Middletown, DE: Createspace.

Dikilitaş, K., & Altay, M. (2011). Acquisition sequence of four categories of non-generic use of the English definite article *the* by Turkish speakers. *Novitas-Royal, 5*(2), 183-198.

Garcia-Mayo, M. P. (2008). The acquisition of four nongeneric uses of the article *the* by Spanish EFL learners. *System, 36*(4), 550-565.

Hawkins, J. (1978). *Definiteness and indefiniteness*. New York: Routledge.

Huebner, T. (1983). *The acquisition of English: A longitudinal analysis*. Ann Arbor, MI: Karoma Publishers, Inc.

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

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua, 118*(4), 554-576.

Koç, K. (2015). The non-generic use of the definite article *the* in writing by Turkish learners of English. *Hacettepe Eğitim Dergisi, 20*(2), 56-68.

Lee, J., Park, S., & Heinz, M. (2018). Exploring patterns of article use by advanced Korean learners of English and Spanish. *International Review of Applied Linguistics in Language Teaching, 41*(1), 79-101.

Lesniewska, J. (2016). The use of articles in L2 English: A phraseological perspective. *Studies in Second Language Learning and Teaching, 6*(2), 203-224.

Liu, D., & Gleason, J. (2002). Acquisition of the article *the* by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition, 24*(1), 1-26.

Master, P. (2002). Information structure and article pedagogy. *System, 30*(3), 331-348.

Reid, W. (1991). *Verb and noun number in English: A fundamental explanation*. New York: Longman.

Schönenberger, M. (2014). Article use in L2 English by L1 Russian and L1 German speakers. *Zeitschrift für Sprachwissenschaft, 33*(1), 77-105.

Simons, G. F., & Fenning, C. D. (Eds.). (2018). *Ethnologue: Languages of the world, 21st edition).* Retrieved from https://www.ethnologue.com/language/azj

Snape, N., Garcia-Mayo, M. P., & Güreli, A. (2013). L1 transfer in article selection for generic reference by Spanish, Turkish and Japanese L2 learners. *International Journal of English Studies, 13*(1), 1-28.

Wong, B. E., & Quek, S. T. (2007). Acquisition of the English definite article by Chinese and Malay ESL learners. *Electronic Journal of Foreign Language Teaching, 4*(2), 210-234.

Yıldırım, F. B. (2015). Turkish ELT students' acquisition of the definite article: An analysis of four non-generic uses. *Procedia, 199*, 748-755.

## Appendix: The Article Usage Test

What is your native language? _____

How many years have you studied English? _____

What is your age? _____

**Directions**: In some of the following sentences, the definite article "the" is missing. Please read the following sentences carefully and insert the article "the" wherever you believe necessary.

1. Fred bought a car on Monday. On Wednesday, he crashed car.
2. I look after a little girl and a little boy on Saturdays. Little girl is smart but boy isn't.
3. We rented a boat last summer at a lake. Unfortunately, boat hit another boat and sank.
4. My mother has a white dog and a black dog. White dog is taller than black one.
5. The mother says to her children, "Come on, it's time to go to Grandma's house."
6. I watched several old movies last weekend. I enjoy watching old movies.
7. I have read a few science fiction books this semester. Science fiction books are really interesting.
8. At the zoo I saw several tigers. I think that tigers are beautiful animals.
9. While driving in their car to work, the husband asks his wife, "Could you open window please?"
10. Our office got some new computers last week. Someday, I really think that computers will replace people everywhere.

11. I saw a man in a car across the street. At first I wasn't sure, but then I realized that man driving car was a friend of mine.
12. Do you know pilot who flies this airplane?
13. Man I met in New York later became my husband.
14. Blue car across the road is very suspicious.
15. Did you hear house we saw last week was burned down last night?
16. While driving in their car to work, the father says to his son, "Please turn on radio."
17. I've heard of parents who don't give their children enough to eat.
18. People from around the world are meeting here today.
19. Things of beauty always bring great joy.
20. Sally Ride was first woman in space.
21. Professor who teaches the physics class explains things very well.
22. Water in this glass is dirty.
23. A man says to his wife at the breakfast table, "Can you pass me newspaper?"
24. Shoes in department stores tend to be expensive.
25. The manager asks her secretary, "Could you please check schedule for me?"
26. Pacific Ocean is the largest in world.
27. Moon is full tonight.
28. Who is leader of your club?
29. President of the United States lives in White House.
30. My mother likes to have salads at dinner because salads are very healthy.
31. Bill caught Malaria (a disease) while traveling in Africa.

32. In a bright sunny room, the woman asks the man “Could you close curtains, it’s too bright in here.”
33. I like to watch movies that are black and white.
34. The teacher says to his pupils, “Read Chapter Twenty in your book.”
35. Mississippi river runs through Louisiana.
36. Lake Michigan is a large lake in North America.
37. I start back to work on Monday.
38. The teacher says to her students, “The meeting will not be held until next week.”
39. The mother asks the father, “Is baby sleeping?”
40. I’m sick. I’ve come down with flu.
41. England is part of United Kingdom.
42. John’s wife died of cancer in 1996.
43. The wife says to her husband, who is hanging a picture in the room, “Picture isn’t straight.”
44. Salt Lake City is in Utah.
45. The game show host says to the contestant, “What’s behind door number one?”
46. Mount Etna in Sicily is still an active volcano.
47. In their living room at bedtime, the mother says to the children, “Turn off television.”