

The role of semantic features in the acquisition of English articles by Russian and Korean speakers

Tania Ionin¹, Heejeong Ko² and Ken Wexler²

¹University of Southern California ²Massachusetts Institute of Technology

1. INTRODUCTION

Much recent work on second language (L2) acquisition has examined the role of *features* in L2-grammar. For example, the features related to Tense and Agreement have been the subject of much L2-literature (see, among many others, Eubank 1993/94; Prévost and White 2000; Lardiere 2000). Other features that have been studied in L2-acquisition include Aspectual features (see Gavrusseva 2003, among others) and features underlying the use of pronominal and null subjects (see Sorace 2000, among others).

In this paper, we are particularly concerned with features that determine article choice inside the DP. There have been a number of studies concerning article choice in L2-English (see Huebner 1983; Master 1987; Parrish 1987; Thomas 1989; Young 1996; Murphy 1997; Robertson 2000; and Leung 2001, among others). Many studies of L2-English article choice (e.g., Huebner 1983, Master 1987, Parrish 1987 and Thomas 1989) have found that L2-English learners overused the definite article *the* in contexts where the indefinite article *a* was required. (One exception was the study of Leung 2001, which found high overuse of *a* in definite contexts). To our knowledge, there is at present no consensus about the causes underlying article misuse in L2-English.

In the present work, we will provide a feature-based account of article errors in L2-English. We will examine two semantic features: the feature [+definite] and the feature [+specific], and investigate the role that these features play in the acquisition of English articles by adult speakers of article-less languages, Russian and Korean. We will argue that L2-learners have access to both of these features. We will show that both features play a role in article specifications cross-linguistically, and argue that the patterns of article use and misuse in L2-English reflect direct access to universal semantic features in L2-acquisition.

This paper is organized as follows. Section 2 gives the necessary theoretical background on definiteness and specificity. In Section 3, we make a hypothesis for article choice in L2-English, and provide a brief summary of three formal written elicitation studies which provide support for this hypothesis. Section 4 reports our findings of article use in L2-English written narrative data. Finally, Section 5 concludes the paper. Section 6 contains an appendix describing coding procedures.

2. THEORETICAL BACKGROUND: DEFINITENESS AND SPECIFICITY

Articles cross-linguistically can encode different semantic features. In this paper, we are particularly concerned with the *definiteness* and *specificity* features. While the term *specificity* has received multiple definitions in the literature, we use it throughout this paper in a very particular sense, *specificity as speaker intent to refer* (cf. Fodor and Sag 1982).

2.1. Definiteness and specificity: the definitions

The features [+definite] and [+specific] are both discourse-related: they are related to the knowledge/mind state of the speaker and/or the hearer in the discourse. This is shown by the informal definitions in (1): the feature [+definite] reflects the state of knowledge of *both speaker and hearer*, while the feature [+specific] reflects the state of knowledge of *the speaker only*.

1. Definiteness and Specificity: informal definitions

If a Determiner Phrase (DP) of the form [D NP] is...

- a) [+definite], then the speaker assumes that the hearer shares the speaker's presupposition of the existence of a unique individual in the set denoted by the NP
- b) [+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

The more formal definitions of definiteness and specificity are given in (2). We adopt the standard presuppositional (Fregean) analysis of definites.¹ Our definition of specificity is based on Fodor and Sag's (1982) proposal concerning *speaker intent to refer*. The definition that we use involves the additional concept of *noteworthy property*, as developed in Ionin 2003 (see also footnote 4 below for more discussion).

2. Definiteness and Specificity: formal definitions

a) Definiteness (Fregean analysis)

[*the* ζ] ξ expresses that proposition which is

- true at index i , if there is exactly one ζ at i , and it is ξ at i ,
- false at an index i , if there is exactly one ζ at i , and it is not ξ at i ,
- truth-valueless at an index i , if there isn't exactly one ζ at i .

(from Heim 1991:9)

b) Specificity

A sentence of the form [sp α] ζ expresses a proposition only in those utterance contexts c where the following felicity condition is fulfilled: the speaker of c intends to refer to exactly one individual x_c in c , and there exists a property φ which the speaker considers noteworthy in c , and x_c is both α and φ in c . When this condition is fulfilled, [sp α] ζ expresses that proposition which is true at an index i if x_c is ζ at i and false otherwise.

(based on Fodor and Sag 1982, with modifications: see Ionin 2003:56)

While the definitions in (2) are for singular DPs only, plural DPs can also be definite or specific (on definite plurals, see Heim 1991 and the references cited therein). If the Fregean analysis of definites is extended to plurals, a definite plural like *the cats* is presupposed to contain a "greatest" element, where "a

¹ The definition of definiteness in (2a) does not actually involve speaker or hearer knowledge. However, this definition does involve a *presupposition* of uniqueness. We treat presuppositions, including the presupposition of uniqueness, as accessing the *common ground* – i.e., knowledge shared by the speaker and hearer (for the relationship between presuppositions and the common ground, see Stalnaker 1974 and much subsequent work). In contrast, the definition of specificity in (2b) bears only a felicity condition, which accesses speaker knowledge only, and *not* the common ground – see Ionin (2003) for more discussion.

greatest element of a set M is an element of M which has all other elements of M as parts” (Heim 1991:22). The analysis of specificity can also be made to apply to plural as well as singular DPs, as long as instead of “exactly one individual x_c ” we talk about a set of individuals which the speaker intends to refer to, and the property φ , as well as the property denoted by the restrictor NP, hold for the maximal member of this set.

2.2. Definiteness and specificity cross-linguistically

In this section, we will show that the features [+definite] and [+specific] both play a role in article specifications cross-linguistically. We discuss how these features are instantiated in two languages: English and Samoan. To clarify: a *context* that is [+definite] or [+specific] is a context that has satisfied the conditions on definiteness (2a) or the conditions on specificity (2b), respectively. When we say that a context is [-definite] or [-specific], we mean that these conditions have not been satisfied. Below, we explain what it means for *lexical items* (as opposed to contexts) to be [+definite] or [+specific].

2.2.1. Definiteness in English. Standard English has two articles, *the* and *a*, which are used in [+definite] and [-definite] contexts, respectively. While *the* is used in both singular and plural [+definite] contexts, *a* is used only in singular [-definite] contexts. For the purposes of this paper, we will take bare plurals to be the plural counterpart of singular *a*-indefinites.² (In the case of mass nouns, we will similarly take lack of an overt article to be the counterpart of *a*.)

Following Heim (1991), we treat *the* as being specified [+definite], and *a* (and by extension, bare plurals and bare mass nouns) as being underspecified for definiteness. This means that *the* can be used felicitously only when the conditions on definiteness (2a) have been satisfied, while *a* has no such requirement. As a result, whenever the conditions on definiteness have been met, *the* is used.³ If these conditions have not been met, then *the* cannot be used, and *a* (or no article, in the case of plurals and mass nouns) is used instead.

An illustration of the use of *the* vs. *a* is given in (3). Upon the first mention of a cat, there is no presupposition that a unique cat exists, so the conditions on definiteness given in (2a) have not been met. As a result, the indefinite article *a* is used. In contrast, upon second mention of the same cat, the existence of a particular, unique cat (the one that has just been mentioned) has been established. The conditions on definiteness have been met, so *the* is used.

3. I saw a cat. I gave the cat some milk.

Previous discourse is not always necessary for establishing uniqueness. The uniqueness presupposition can also be satisfied as a result of mutual world knowledge, as in *the president*, *the winner of this race*, etc. For more discussion of the conditions on definiteness, see Heim (1991). We will also come back to a more detailed discussion of different types of definites in Section 4.4.1.

2.2.2. Specificity in English. Crucially, standard English has no marker of specificity. Colloquial (spoken) English, on the other hand, does have a marker of specificity: the demonstrative *this* on its indefinite referential use. This is illustrated in examples (4), from Maclaran 1982. In (4a), the speaker is intending to refer to a particular telephone, which has the noteworthy property of being a weird purple telephone. In (4b), the identity of John’s telephone is irrelevant, and the speaker is not intending to refer to

² Bare plurals are in fact known to behave differently from *a*-indefinites – e.g., by obligatorily taking narrow scope under an operator (see Carlson 1977; see Chierchia 1998 on the conditions under which bare plurals *can* take wide scope). For the purposes of this paper, we will ignore the differences in the behavior of *a*-indefinites and bare plurals in English, since these differences are irrelevant for our discussion: what is relevant is under what conditions L2-English learners overuse *the* with singular as well as plural indefinites.

³ In principle, both *the* (marked [+definite]) and *a* (underspecified for definiteness) are compatible with a [+definite] context; however, only *the* is in fact used in [+definite] contexts. Heim (1991) proposes that the choice of *the* over *a* is due to the rule “Make your contribution presuppose as much as possible” (Heim 1991:28), which has become known as the “Maximize Presupposition” principle. Whenever two lexical items (such as *the* and *a*) are both compatible with a given context, the item with the most presuppositions (*the*) wins out, as the most informative.

a particular telephone. The conditions on specificity are thus satisfied in (4a) only; only in this case can referential *this* be used. As Maclaran (1982:90) put it, referential *this* “draws attention to the fact that the speaker has a particular referent in mind, about which further information may be given.”

4. a) John has a/*this* weird purple telephone.
- b) John has a/*#this* telephone, so you can reach me there.

(from Maclaran 1982:88, ex. 85)

We conclude that referential *this* in colloquial English bears the feature [+specific], and can therefore be used only in specific indefinite contexts. The indefinite article *a*, on the other hand, is underspecified for specificity, and can therefore be used both in [+specific] indefinite contexts (4a) and in [-specific] indefinite contexts (4b). For more discussion of the nature of referential *this* see Prince 1981 and Maclaran 1982; for the relationship between *this* and *noteworthiness*, see Ionin 2003.⁴

While referential *this* is incompatible with definites, the conditions on specificity (2b) can be satisfied by definite as well as indefinite contexts (cf. the referential/attributive distinction of Donnellan 1966). This is illustrated in (5).

5. a) I'd like to talk to the winner of today's race – she is my best friend!
- b) I'd like to talk to the winner of today's race – whoever that is; I'm writing a story about this race for the newspaper.

In (5a), the speaker is intending to refer to a particular individual, who is the winner of today's race, and who has the noteworthy property of being the speaker's best friend. In (5b), the speaker is not intending to refer to a particular individual, but simply wants to talk to whoever happens to be the winner of today's race. Thus, the conditions on specificity are satisfied in (5a) but not in (5b). We conclude that the specificity distinction cross-cuts the definiteness distinction: the conditions on specificity can be satisfied, or not satisfied, in both definite and indefinite contexts.⁵

It is important to note that the feature [+specific] on our definition is not identical to wide scope. In both sentences in (4), for instance, the DP takes widest scope in the sentence; both sentences assert the existence of a particular telephone in the actual world (a telephone owned by John). Only (4a), however, has the additional property of “noteworthiness” attached to the DP, and only this sentence licenses the use of referential *this*. On our analysis, a [+specific] DP necessarily takes widest scope over intensional verbs or modals: if the speaker intends to refer to a particular individual in the actual world, that individual must exist in the actual world. The reverse does not hold, however: a wide-scope DP can be [-specific], as illustrated by (4b). The situation is similar for definites: a definite DP which takes scope under an operator is obligatorily [-specific], as in (5b). A definite DP which takes scope over an operator may be [+specific] or [-specific].⁶

⁴ The view of specificity as involving a *noteworthy property*, developed in Ionin (2003), is based on data such as (i), which show that use of indefinite *this* in English requires that the speaker say something *noteworthy* about the referent: e.g., that *it's a movie that my friends have been recommending to me for ages*. These data show that *speaker intent to refer* is not enough: the speaker intends to refer to a particular new movie in all of the sentences in (i).

- i. a) #I want to see this new movie.
- b) I want to see this new movie that my friends have been recommending to me for ages.
- c) I want to see this new movie - it's one that my friends have been recommending to me for ages.

⁵ See Ionin 2003 (Ch. 2) for a discussion of why referential *this* is incompatible with definites. On Ionin's analysis, *the* carries a presupposition while *this* carries only a felicity condition. Since presuppositions are more informative than felicity conditions, *the* rather than *this* is used whenever the presupposition on *the* is satisfied (i.e., in all definite contexts). This is derived via the “Maximize Presupposition” principle of Heim (1991) (see footnote 3). Note that in [+specific] indefinite contexts, both *this* and *a* are possible (as in (4a)): “Maximize Presupposition” does not force the choice of *this* over *a* because neither lexical item bears a presupposition (*this* has only a felicity condition).

⁶ Our definition of the feature [+specific] is thus crucially different from the view of specificity adopted by much L2-literature. Starting with Huebner (1983), much work on articles in L2-English (e.g., Thomas 1989a, Robertson 2000, among others), has treated specificity (the [+specific referent (SR)] feature) as essentially corresponding to *existence in the actual world*. Our view of specificity is more

2.2.3. *Specificity in Samoan*. There is evidence from the work of Mosel and Hovdhaugen (1992) that articles in Samoan encode specificity rather than definiteness (see also Lyons 1999). According to Mosel and Hovdhaugen, Samoan uses one article (*le*) with specific DPs, and another article (*se*) with non-specific DPs. As Mosel and Hovdhaugen (1992:259) state, “[t]he specific article singular *le/l=* ART indicates that the noun phrase refers to one particular entity regardless of whether it is definite or indefinite.” Use of *le* in a [-definite, +specific] context is illustrated in (6a): the speaker is beginning to tell a story, introducing new characters who will be important later on in the story. This is arguably a specific indefinite use of *le*: in English, referential *this* could be used in this context. On the other hand, “[t]he nonspecific singular article *se/s=*ART(nsp.sg.) expresses the fact that the noun phrase does not refer to a particular, specified item, but to any member of the conceptual category denoted by the nucleus of the noun phrase and its adjuncts” (Mosel and Hovdhaugen 1992:261). Use of *se* in a [-definite, -specific] context is illustrated in (6b): here, the lady’s identity is irrelevant, so the non-specific article is used.

6. a) ‘O *le* *ulugāli’i*, *fānau* *l=a* *lā* *tama* ‘o *le*
 PRES ART couple give birth ART=Poss 3.du. child PRES ART
teine ‘o *Sina*
 girl PRES Sina
 “There was a couple who had a child, a girl called Sina.”

(Mosel and Hovdhaugen:259, ex. 6.37)

b) *Sa* *fesili* *mai se* *tamaitai* *po=o* *ai* *l=o*
 PAST ask DIRART(nsp.sg.) lady Q-PRES who ART=Poss
ma *tama*.
 1.exc.du. father
 “A lady asked us who our father was.”

(Mosel and Hovdhaugen 1992:261, ex. 6.46, 6.50)

Mosel and Hovdhaugen provide additional examples which suggest that Samoan also distinguishes definites on the basis of specificity: *le* is used in [+definite, +specific] contexts, and *se* is used in [+definite, -specific] contexts. Thus, there is reason to believe that Samoan articles encode specificity rather than definiteness (see Ionin 2003 for more discussion).

3. ARTICLE CHOICE IN L2-ACQUISITION

The above discussion shows that the features [+definite] and [+specific] both receive morphological expression cross-linguistically. The [+definite] feature is expressed by *the* in English (as well as by definite articles in a variety of other languages, e.g., French, Spanish, German, etc.). The [+specific] feature is expressed by *le* in Samoan as well as by referential *this* in English. A natural question to ask, then, is how L2-learners come to know which of the two features is encoded by articles in their L2. We explore this next.

3.1. Hypothesis and predictions

Suppose that L2-English learners are exposed to a more or less standard variety of English, without much exposure to the referential *this*. They are thus learning a language in which the feature [+definite] receives morphological expression through *the* while the feature [+specific] does not receive morphological expression at all. Since, as shown in the previous section, both the [+definite] and [+specific] features are part of the cross-linguistic inventory of article specifications, how will L2-English learners come to know that *the* is [+definite] and not [+specific]?⁷

restrictive: it involves *speaker intent to refer* to an individual who exists in the actual world. See Ionin and Wexler (2003), Ionin, Ko and Wexler (to appear) for evidence that *existence in the actual world* without *speaker intent to refer* is insufficient to account for article choice in L2-English.

⁷ We are assuming that the adult L2-English learners in our studies do not receive much exposure to referential *this*, since their exposure to English takes place largely in the classroom, while referential *this*

One answer to this question has to do with L1-transfer. For instance, if speakers of Spanish, which also encodes the feature [+definite] in its article system, are learning English as an L2, they may transfer article specifications from Spanish to English and correctly treat *the* as [+definite]. In contrast, speakers of Samoan learning English as an L2 may mistakenly treat *the* as [+specific]. The role of L1-transfer in L2-English article choice is an interesting question awaiting further study.

In this paper, however, our focus is on L2-English learners whose native languages, Russian and Korean, lack articles, and who are therefore unlikely to be influenced by L1-transfer.⁸ These learners do not, initially, have any way of knowing that *the* and *a* are distinguished on the basis of definiteness rather than on the basis of specificity.

We hypothesize that these learners will start out accessing both possibilities: the possibility that English articles are distinguished on the basis of the definiteness feature (Option (a) in Table 1); and the possibility that English articles are distinguished on the basis of the specificity feature (Option (b) in Table 1).

Table 1. Article grouping cross-linguistically

a) Article grouping by definiteness (e.g., English)			b) Article grouping by specificity (e.g., Samoan)		
	+definite	-definite		+definite	-definite
+specific			+specific		
-specific			-specific		

We propose that, until they have received sufficient input, L2-learners should fluctuate between the two possibilities in Table 1, as stated in our *Fluctuation Hypothesis* in (7). (See Ionin (2003, Ch. 1) for evidence that similar fluctuation exists in other domains, e.g., verb-raising.)

7. The Fluctuation Hypothesis (FH) for L2-English article choice:
- 1) L2-learners have full access to the features that can underlie article choice cross-linguistically: the features [+definite] and [+specific].
 - 2) L2-learners fluctuate between dividing English articles on the basis of the [+definite] feature vs. the [+specific] feature, until the input leads them to choose the [+definite] feature.

According to the FH, L2-learners' interlanguage grammar is at all times constrained by the universal semantic features of definiteness and specificity. The learners essentially entertain multiple possible grammars (an English-like grammar and a Samoan-like grammar) at the same time. With sufficient input, learners should eventually be able to converge on the right option – i.e., to choose the English grammar, in which articles are divided on the basis of the [+definite] feature. The question of what constitutes

occurs mainly in casual, informal speech. Hence, the learners are acquiring a two-article language, with only *the* and *a*. If L2-English learners were truly acquiring a three-article language, considering referential *this* on par with *the* and *a*, we would expect them to incorporate both definiteness (marked by *the*) and specificity (marked by *this*) into their article system (see Ionin 2003, Ch. 2 for more discussion, and speculation that acquiring a three-article language may be easier than acquiring a two-article language, since the learners would not have to “choose” between encoding definiteness and specificity in their article system).

⁸ Russian and Korean have no articles, and no way of obligatorily marking either definiteness or specificity, so no direct L1-transfer is expected. (See Ionin (2003, Ch. 3) for more discussion of the nominal systems of Russian and Korean, and evidence that they do not directly encode either definiteness or specificity). Russian and Korean do, however, have non-obligatory markers of definiteness (such as demonstratives) and/or of specificity (the numeral ‘one’ in Russian appears to have a specific indefinite reading – see Ionin 2003, Ch. 3, for more discussion). In Russian, definiteness and/or specificity may also interact with word order. Thus, some aspects of Russian or Korean grammar could in principle affect L2-English article choice through L1-transfer. However, it is very unlikely that L1-transfer from these two very different languages would result in similar patterns of L2-English article use. Thus, similar patterns of behavior on the part of L1-Russian and L1-Korean learners of English can be taken as evidence that universal processes rather than L1-transfer are at work. Our formal written elicitation data (summarized in Section 3.2) show that L1-Russian and L1-Korean learners of English do in fact exhibit similar patterns of article use, a finding which rules out the L1-transfer explanation in favor of the *Fluctuation Hypothesis* discussed in this section.

“sufficient input” is addressed in Ionin et al. (to appear), where we show that individual contexts of *the* use are frequently both [+definite] and [+specific], resulting in ambiguous input. In that paper, we suggest that in order to choose the “English” option in Table 1, L2-learners would need to generalize across multiple instances of article use. For the sake of space, we will not be addressing the question of input in the present paper. Our focus here is on describing the *state* of L2-learners’ grammar, rather than on the *process* by which L2-English learners achieve target-like article use.

An important property of the state of L2-learners’ grammar under the FH is *non-random article use*. The FH makes explicit predictions for where errors of L2-English article use should occur. Whether learners adopt the division (a) or (b) in Table 1, they will use one article (*the*) in [+definite, +specific] contexts and a different article (*a*) in [-definite, -specific] contexts: there should be few or no errors in these two context types.⁹

On the other hand, as long as the learners are fluctuating between the two possibilities in Table 1, they will show interchangeable use of *the* vs. *a* in [+definite, -specific] as well as [-definite, +specific] contexts. Whenever the L2-learners adopt Possibility (b) in Table 1, they will associate *the* with the [+specific] feature, which will result in overuse of *the* in [-definite, +specific] contexts, and, conversely, lack of *the* (hence overuse of *a*, for singulars) in [+definite, -specific] contexts.

This prediction is captured in Table 2. The two cells in bold are where errors of article use are expected.

Table 2. Predictions for article choice in L2-English: singular contexts

	+definite (target: <i>the</i>)	-definite (target: <i>a</i>)
+specific	correct use of <i>the</i>	overuse of <i>the</i>
-specific	overuse of <i>a</i>	correct use of <i>a</i>

3.2. Previous findings: formal written elicitation studies

We tested the predictions in Table 2 in a series of formal written elicitation studies with adult L2-English learners. Since these studies are reported in detail elsewhere (see the papers cited below, as well as Ionin 2003), we summarize them only briefly here.

3.2.1. Study design and participants. The first of our series of formal written elicitation studies (reported in Ionin and Wexler 2003) tested article use among adult L1-Russian learners of English. The second and third studies (reported in Ionin, Ko and Wexler 2004 and *to appear*, respectively) compared article use among adult L1-Russian and adult L1-Korean learners of English. Russian and Korean both lack articles, so direct L1-transfer was not expected. Since the two languages are highly dissimilar, similar patterns of performance among the two groups of L2-learners are highly unlikely to be due to some indirect form of L1-transfer (see also footnote 8).

All of the studies used the format of formal written elicitation. The participants were presented with a series of dialogues between two people, such as the definite and indefinite contexts in (8) and (9) (from Ionin and Wexler 2003).¹⁰ The target article (which is underlined here) was missing. The L2-learners’ task in the first study was to fill in the blank with *the*, *a*, or a dash (to represent that no article was needed). In the last two studies, these three options were provided, and the learners were instructed to choose one of them.

⁹ In principle, article misuse could occur in these contexts due to other factors besides definiteness and specificity. We are at present investigating another semantic factor that may in fact be at work in L2-English article choice, namely *partitivity*, or membership in a previously established set (see Ko, Ionin and Wexler 2004). Partitivity is not relevant for the results reported in the present paper, since none of the contexts discussed here (either in the formal elicitation or the written narrative tasks) were partitive.

¹⁰ In the first two studies, the context of the dialogue was given in the learners’ L1, with only the target sentence in English. In the third study, the entire context was given in English. Since similar patterns of article choice were obtained across all three studies, the language of the context did not appear to make a difference.

8. A: I know that Betsy went to the bookstore yesterday and bought a novel and a magazine. Do you happen to know which one she read first?
B: She read the magazine first.
9. A: Nick just went to the living room. Do you know what he is doing there?
B: He is watching a cartoon on television.

The first two studies tested article use in three of the four contexts in Table 2; the missing context was the [+definite, -specific] one. All four contexts in Table 2 were tested in the last study, reported in Ionin et al. (to appear). In all studies, contexts intended to be [+specific] made it clear that the speaker had knowledge of the referent and considered the referent's identity important. Contexts intended to be [-specific] did not establish such speaker knowledge, and often included explicit denial of speaker knowledge and/or narrow scope of the (in)definite DP under an operator. All three studies focused on singular contexts, although the second study (Ionin et al. 2004) tested plural contexts as well. In all studies, the formal written elicitation tasks were piloted with adult native English speakers.

3.2.2. Summary of results. The three formal written elicitation studies showed that L2-English learners do not misuse English articles randomly: rather, their errors are due to an association of *the* with the feature [+specific]. Across all three studies, learners showed largely correct use of *a* in singular [-definite, -specific] contexts while overusing *the* in singular [-definite, +specific] contexts. (In the plural contexts tested in the second study, overuse of *the* was likewise largely confined to [-definite, +specific] contexts). The last study further showed that L2-learners showed largely correct use of *the* in [+definite, +specific] contexts, while overusing *a* in [+definite, -specific] contexts. Since the feature [+specific] plays a role in article specifications cross-linguistically (e.g., in Samoan *le* as well as in English referential *this*), L2-learners' errors thus reflect the learners' access to the universal semantic distinctions which underlie article choice.

Before this conclusion can be definitively reached, however, it is important to obtain convergent results using different methodologies. A drawback of the formal written elicitation studies is that they test the learners' *conscious* knowledge of English articles. It is important to find out whether L2-learners would show a similar pattern of performance when their *unconscious* knowledge of articles is tested. We turn to this in the next section.

4. ARTICLES IN L2-ENGLISH: WRITTEN NARRATIVE DATA

In order to find out what L2-learners' unconscious knowledge of English articles is like, we supplemented our last formal written elicitation study with a collection of written narrative data, a task which did not explicitly target English articles. The main findings obtained from the written narrative data are briefly summarized in Ionin et al. (2004, *to appear*). In this section, we report the written narrative study in greater detail.

4.1. Predictions for L2-English written narrative data

The main advantage of written narrative data is that they give us a better indication than the formal written elicitation studies of how L2-English learners use articles in daily life. While the formal written elicitation tasks focus the learners' attention on article use, collection of written narrative data allows us to examine whether such errors as overuse of *the* with indefinites also occur when the learners are not focusing on article choice. Additionally, written narrative data allow us to explore an aspect of article use that was not investigated in any detail in the formal written elicitation studies: article omission.

4.1.1. Article choice, definiteness, and specificity in written narrative data. Our first prediction is that the patterns of article use in written narrative data will be similar to article use in the formal written elicitation tasks reported above. The Fluctuation Hypothesis for article choice proposed in (7) predicts that we should see overuse of *the* primarily in [-definite, +specific] contexts and overuse of *a* primarily in [+definite, -specific] contexts in written narrative data as well as in formal written elicitation. Article use in

other contexts should be more accurate: crucially, L2-learners' article errors will be non-random, but will be constrained by definiteness and specificity.¹¹

Given this hypothesis, as well as the findings of the formal written elicitation tasks reported in Section 3, we can therefore make Prediction A in Table 3. While the formal written elicitation studies focused on article use with singular count nouns, we expect L2-learners to use singular, plural, and mass nouns in the written narrative data. Thus, Table 3 includes predictions for use of *the* and *a* as well as for article omission, which is appropriate with indefinite plural and mass nouns. The symbol '--' in Table 3 stands for article omission.

Table 3. Prediction A: article choice, definiteness, and specificity in written narrative data

	+definite	-definite
	Target: <i>the</i>	Target: <i>a</i> (sg), -- (pl, ms)
+specific	correct use of <i>the</i>	overuse of <i>the</i>
-specific	overuse of <i>a</i> , --	correct use of <i>a</i> , --

sg = singular count noun

pl = plural count noun

ms = mass noun

4.1.2. *Determining specificity in written narrative data.* The main disadvantage of written narrative data collection, as compared to formal written elicitation tasks, is that it does not allow the investigators to control the contexts in which articles are produced. It is possible to code each context in the L2-English data as [+definite] or [-definite] by asking native English speakers whether the context should receive *the* or *a*. However, it is quite difficult to code a context in L2-English written narrative data as [+specific] or [-specific]: as seen in the definition in (2b), specificity is largely in the mind of the speaker, and whether an L2-learner will use a specific DP depends on whether she intends to single out a particular individual via some noteworthy property.

While for most contexts, it is impossible to determine whether they are [+specific] or [-specific], there are some cues that can be used. Let's consider indefinites first. Scope is a cue: indefinites which take narrow scope under an intensional verb or modal are obligatorily [-specific], as discussed in Section 2.2.2. Wide-scope DPs may be specific or non-specific, on the other hand. Next, consider indefinites in *there* and *have* constructions. One way to find out whether an indefinite is [+specific] is to check whether it is compatible with the use of referential *this*, a marker of specificity in English. As illustrated in (10a) vs. (10b), [+specific] indefinites headed by referential *this* are incompatible with *there*-constructions that simply list a series of descriptions. On the other hand, it is quite possible to have a *this*-indefinite in a *there*-construction when the speaker intends to attract attention to a particular individual, as in (10c). The same is shown for indefinites in *have*-constructions (11).

10. a) There are a bird and a squirrel in the garden.
 b)# There are this bird and this squirrel in the garden.
 c) There is this peculiar bird in the garden – it doesn't look like anything I've ever seen!
11. a) In my kitchen, I have a stove, a refrigerator, and a large round table.
 b)# In my kitchen, I have this stove, this refrigerator, and this large round table.
 c) I have this really neat new coffeemaker in my kitchen – it has a timer and it turns itself off automatically.

Since expressions such as (10c) and (11c) are quite felicitous, it is not possible to argue that [+specific] indefinites are inherently incompatible with *there* or *have* constructions. However, the written narrative task administered in our study aimed at eliciting *there* and *have* constructions of the type in (10a) and (11a), which involve a simple listing of descriptions, with no importance attached to a particular referent. Such contexts are highly unlikely to contain [+specific] indefinites. Thus, for the purposes of this paper, we will treat *there* and *have* constructions as being biased in favor of containing [-specific] indefinites, and will analyze them separately from all other indefinite constructions.

¹¹ We should point out that our focus is on article choice in *non-generic* definite and indefinite contexts. Genericity adds a new semantic dimension, one which we will not discuss here. There were in fact very few definite or indefinite generics in the written narrative data, not enough for any meaningful conclusions (see Ionin 2003, Ch. 7 for the data).

Given the above discussion, we can make specific predictions for article errors with indefinites in written narrative data, as stated in (12):

12. *Prediction A1: article misuse with indefinites*

Overuse of *the* with indefinites will be limited to wide-scope indefinite contexts, and will not occur with narrow-scope indefinites, or in *there* and *have* constructions.

Fewer cues are available in the case of definites. Definites in *previous mention* environments are obligatorily [+specific]. Whenever the speaker mentions a referent for a second time, she is intending to refer to a particular individual: thus, in (13), the speaker is referring to a very particular cat, which has the noteworthy property of being *the cat that the speaker saw*.

13. I saw a cat and a dog. I picked up the cat.

Definites which take narrow scope with respect to intensional verbs or modals must be [-specific] (see Section 2.2.2). Other types of definites (not previous mention, not narrow-scope) may in principle be either [+specific] or [-specific]. The only clear-cut instances of [-specific] definites (besides narrow-scope definites) would be instances where the speaker's knowledge of the referent's identity is explicitly denied. We can thus make the prediction in (14).

14. *Prediction A2: article misuse with definites*

Overuse of *a* with definites will occur primarily in narrow-scope definite contexts, and other definite contexts in which speaker knowledge of the referent is denied.

A more detailed consideration of different types of definites is given in Section 4.4.1.

4.1.3. *Article omission in written narrative data.* In our formal written elicitation studies, the L2-learners almost never omitted articles with singular DPs. However, it is well-known, both anecdotally and from previous studies of L2-English (e.g., Huebner 1983, Robertson 2000), that L2-English learners do omit articles to a fairly large extent in more spontaneous production. The very low rates of article omission in our formal elicitation studies may be traceable to these studies' focus on article use: the learners' awareness that they were being tested on articles may have increased the rate of article suppliance. In written narrative data, on the other hand, where there is no overt focus on articles, we may see rates of omission more comparable to those found in previous literature on L2-English article use. Thus, written narrative data have an advantage over formal written elicitation data: they allow us to examine both article use and article omission in L2-English, and to see whether article omission corresponds to a particular semantic distinction.

Article omission may occur in written narrative data due to a variety of non-linguistic factors, such as retrieval difficulties and performance pressures; examining the various non-linguistic factors that may contribute to article omission is beyond the scope of this work. Of particular interest to us is whether article omission will follow any particular linguistic pattern: whether articles are omitted more in some semantic environments than in others. (See also Robertson 2000 for a discussion of some discourse-related factors contributing to article omission in conversation).

A possible prediction concerning article omission is expressed in (15).

15. *Prediction B:*

L2-English learners omit articles more in singular [-definite] than in singular [+definite] environments in written narrative data.

The rationale for this prediction is that the [+definite] article is more informative than the [-definite] article, since it carries a presupposition of uniqueness. It is reasonable to expect that if L2-learners omit articles under performance pressure, they do so when the article conveys the least amount of information, and hence omit *a* more than *the*: while *a* carries only the information that the DP is singular (information that can also be obtained from the form of the head noun), *the* carries the information that the DP has a unique referent in the discourse. Prediction B in (15) is inapplicable to plurals, since article omission with indefinite plurals is licensed in English.¹²

¹² A caveat is called for here. The hypothesis in (15) assumes that L2-learners at least sometimes correctly treat *the* as [+definite] (rather than [+specific]): our formal written elicitation studies found that

4.2. Methods

The collection of the written narrative data was part of the same study as the last formal written elicitation task, reported in Ionin et al. (to appear), where the major findings of article use with singular DPs in the written narrative data were also summarized. In the present paper, we discuss article use with singular, plural, and mass nouns, and go over article use as well as article omission.

4.2.1. Participants. The participants in this study were 70 adult L2-English learners: 30 L1-Russian speakers and 40 L1-Korean speakers. The characteristics of the two groups of participants are given in Table 4. For more details concerning individual participants, see Ionin (2003) and Ionin et al. (to appear).

Table 4. Characteristics of L2-learners:

	L1-Russian participants	L1-Korean participants
<i>Number</i>	30 (18 female, 12 male)	40 (22 female, 18 male)
<i>Age range</i>	19;2 to 56;7 (mean = 38; median = 35)	19;11 to 40;0 (mean = 31; median = 31)
<i>Age of first exposure to English</i>	8 to 53 (mean = 14; median = 11)	9 to 14 (mean = 12; median = 12)
<i>Age at arrival in the US (start of intensive exposure to English)</i>	19;1 to 55;10 (mean = 36; median = 33)	16;0 to 35;1 (mean = 28, median = 29)
<i>Time in the US</i>	<0;1 to 10;4 (mean = 1;9, median = 1;2)	0;1 to 8;7 (mean = 2;6, median = 1;8)
<i>Proficiency</i>	4 beginner, 11 intermediate, 15 advanced	1 beginner, 6 intermediate, 33 advanced

Most of the learners (28 of the L1-Russian speakers and all 40 of the L1-Korean speakers) had received English instruction before arriving to the U.S. However, intensive exposure to English did not begin for the learners until arrival in the U.S., and all of the L2-learners arrived in the U.S. as late adolescents or adults.

The L1-Russian participants included immigrants, international students and foreign workers. All of them resided in the greater Boston area at the time of the study, and were recruited through advertisement in the Russian community. The L1-Korean participants were primarily international students and their spouses, as well as foreign workers and their spouses. All of them resided in Gainesville, FL at the time of the study and were recruited through advertisement.

The L2-learners' proficiency in English was measured by a portion of the Michigan test of L2-proficiency, which consists of 30 multiple-choice items testing various aspects of English grammar.

4.2.2. Procedure. The participants completed three tasks: a formal written elicitation task; a written narrative task; and the proficiency test described above. The format and results of the formal written elicitation task are reported in Ionin et al. (to appear) (see also the brief summary in Section 3.2 above). Here, we focus on the written narrative task.

The order of presentation between the formal written elicitation and written narrative tasks was randomized, and the Michigan proficiency test was always given last. The tests were administered to the learners singly or in small groups, in a laboratory or classroom environment. Average length of participation was 1.5 hours. The learners were reimbursed monetarily for their participation.

this was indeed the case. On the other hand, we have seen that learners also often treat *the* as encoding the feature [+specific]. Suppose that we see an instance of omission in a context where an L1-English speaker would put *the*. We do not know whether the L2-learner who omitted the article construed the context as [+definite] or [+specific]. The same is true for omission of *a* in a context that could be construed as either [-definite] or [-specific] by the L2-learner. Thus, the results reported in this paper should not be taken as a definitive answer on the relationship between semantic features and article omission.

4.2.3. *The written narrative task.* The L2-English learners were asked to provide written answers to five English questions. A written rather than oral format was chosen to facilitate data analysis: given the low phonological status of English articles, especially of *a*, it would not always be possible to tell, in oral production, whether a learner was producing an article or omitting one.

The questions that the L2-learners responded to are given in (16), in the order that they were presented to the L2-learners. The participants were instructed to provide between three and six sentences as an answer to each question, and to not worry about grammar or spelling. The L2-learners were not aware that the written narrative task tested their article use.

16. a) Talk about some valuable object that you own or owned in the past: either (1) talk about something that you received as a gift, and tell about how you received it; or (2) talk about something valuable that you lost, and tell about how you lost it.
- b) Talk about the day when you first arrived in the U.S. Describe your experiences of that day – what you did, where you went, to whom you talked, etc.
- c) Describe your room – talk about what objects you have in your room, and describe them.
- d) Talk about what you did on one of your recent vacations (for example, winter vacation, Thanksgiving week-end, or summer vacation). Talk about where you went and what you did.
- e) Imagine that you get \$1000 as a gift, and you have to spend it right away (you can't put it in the bank). Talk about how you would spend this money.

The first question (16a) was aimed at eliciting [-definite, +specific] DPs, since the learners would write about an object that was well-known and important to them, but unknown to their reader. The question in (16c) was aimed at eliciting indefinites in *there* and *have* constructions, i.e., indefinites which are typically [-specific]: answers to this question were typically of the form *My room has...* or *In my room there is...*, followed by a list of indefinite descriptions, with no great importance attached to any individual object. The question in (16e) was aimed at eliciting narrow-scope indefinites (and possibly also narrow-scope definites), since the learners would talk about the possible world in which they would win \$1000, rather than the actual world. The questions in (16b) and (16d) did not focus on a particular DP type, and were designed to elicit both indefinites and definites as the learners described their past experience. All of the questions in (16), with the possible exception of (16e), were likely to elicit definites in previous-mention contexts.

4.2.4. *Coding procedure.* Once the written narrative data had been collected from the L2-English learners, they were typed and organized into two sets: a set of narratives by L1-Russian speakers, and a set of narratives by L1-Korean speakers. A version of the L2-learners' narratives was then created for coding by native English speakers – henceforth, *coders*.

The coding procedure was as follows. The coders were given versions of the L2-learners' narratives in which each NP was preceded by a blank, regardless of whether or not it would require an article in L1-English, and regardless of whether the L2-learner had put in an article. The coders were then asked to fill the blank with *the*, *a*, a dash (to represent that no article was needed), or, when necessary, another word (such as *some*, *my*, *this*, etc.). The coders also had the option of putting in a question mark if they could not understand the context well enough to decide which word goes in the blank. Each narrative was coded by four native English speakers. The full details of how the narratives were prepared for coding, and what instructions were given to the coders, are given in the Appendix (Section 6).

The rationale for the above procedure was as follows: since we were interested in L2-English article use in definite vs. indefinite contexts, we needed to evaluate L2-English article choice in contexts that were *unambiguously* definite or indefinite. An unambiguously definite context is one in which L1-English speakers consistently put *the*. An unambiguously indefinite context is one in which L1-English speakers consistently put *a* (for singulars) or omit articles and/or use *some* (for plurals). By asking the L1-English coders to fill in blanks with articles, we could learn which contexts are definite and which are indefinite in L1-English, and we could exclude from the count contexts which are potentially ambiguous. Crucially, the coders were unaware which article had actually been used by the L2-English learners, so the coders' responses were unbiased.

After the responses of all the coders had been collected and entered into a computer file, a post-coding procedure took place. Each context was classified as [+definite] vs. [-definite], and, furthermore, as count noun singular vs. count noun plural vs. mass noun. A context was classified as *indefinite count noun singular* if all four coders put in *a* in that context. A context was classified as *indefinite mass noun* or as

indefinite count noun plural if each of the four coders put in -- (“no article”) or *some*. The distinction between plural and mass nouns was done based on the form of the noun. A context was classified as *definite* if all four coders who examined that context put in *the*. The classification of the definite into *count noun singular* / *count noun plural* / *mass noun* was then done on the basis of the form of the noun.

Contexts on which the coders did not agree were not included in the analysis: this means that contexts where some coders put *the* and others *a*, or where some put *a* and others --, etc., were excluded. Also excluded were the (very few) contexts in which at least one of the coders put in a possessive pronoun or a demonstrative. These fairly strict criteria of exclusion ensured that the final analysis looked only at contexts that are unambiguously definite or indefinite in L1-English. Contexts in which the L1-English coders did not agree are ambiguous, so we cannot tell whether L2-English article use in those contexts is correct or incorrect.¹³

All of the clearly definite and indefinite contexts were then evaluated for type, as discussed in the next section. The next section also reports the data on article use and omission in all of the categories. Results for all L1-Russian speakers are grouped together, and results for all L1-Korean speakers are grouped together. The total numbers are reported. There were not enough contexts in the data to allow individual analyses.

4.3. Article use with indefinites

In this section, we discuss L2-English article use with indefinites, and the relationship between L2-English article use and specificity. We consider singular and plural/mass indefinites separately.

4.3.1. *Types of singular indefinites*. The contexts which native English speakers unambiguously coded with *a* (i.e., singular indefinite contexts) were classified into several different types; the types of particular interest to us are given in (17).¹⁴ Detailed explanations and illustrations follow; illustrations are contexts in which L2-learners either correctly put *a* or omitted the article.¹⁵ For each indefinite context, we will also discuss whether it is more likely to be specific or non-specific.

17. Types of singular indefinite DPs

wide scope
narrow scope
there construction
object of have
postcopular position

Most of the DPs classified as *wide scope indefinites* occurred in sentences with no intensional operators, such as (18a-c). There were a few cases of indefinites clearly scoping over an intensional verb or modal (18d-e). A few cases in which the indefinite was ambiguous between a wide-scope and a narrow-scope reading were excluded from the analysis.

18. *singular wide-scope indefinites: correct use of ‘a’*

- a) I live in a small room in a big private house in Brookline.
- b) In the airport one old lady with a dog greeted us with a real warm smile.
- c) On last Christmas, I got a new leather coat from my relatives in Korea.

¹³ Contexts in which coders put a question mark were evaluated as follows. If only one coder put a question mark in a given context, while the other three coders agreed on the appropriate article, the response of the three coders was taken as definitive. If the other three coders disagreed, or if more than one coder put in a question mark, the context was discarded from further analysis. The same procedure was employed for the few contexts which a coder accidentally skipped. As long as only one coder skipped a given context, and there was agreement among the other three coders, the three coders’ response was taken as definitive.

¹⁴ Not reported here are indefinites in generic, formulaic, and other special contexts. See Ionin (2003, Ch. 7) for more discussion

¹⁵ In all examples of L2-learners’ responses throughout this paper, the original responses are given, complete with errors of spelling and grammar. The relevant DPs are underlined.

- d) Recently, we need \$1000 to fill out \$13,000 to buy a used car that we have negotiated.
- e) Next day I visited UT at Austin and fortunately could meet an american student who helped me all day. I still miss her so far.

Wide-scope indefinites may be [+specific] or [-specific]. In some cases, such as (18c-e), the speaker seems to be attaching some importance to a particular individual from the set denoted by the restrictor NP, be it coat, car or student. For instance, in (18e), the speaker may have in mind the noteworthy property *x gave me a lot of help and I still miss her*; the DP is therefore likely to be [+specific]. In other cases, such as (18a-b), no importance is attached to a particular member of the set denoted by the restrictor NP, be it the set of rooms, dogs or smiles; the DP is then more likely to be [-specific]. Since it is impossible to tell for certain whether a given wide-scope indefinite context is [+specific] or [-specific], we will treat all of the instances of this category as potentially, but not necessarily, [+specific].

The category of *narrow scope indefinites* consists of indefinites that take scope with respect to another scope-bearing element. Most of these occurred in answer to the question in (16e), which asked the L2-learners to speculate what they would do with a thousand dollars (e.g., (19a-c)).

19. *singular narrow scope indefinites: correct use of 'a'*

- a) My answer will be very simple – I will buy a car. In America you can't do anything without a car, even to do a shopping.
- b) I would buy a computer for me. I needed a computer, but couldn't afford it so far. So extra \$1000 would be just right.
- c) I will go to a fancy restaurant with my wife and order an expensive dinner and win that I couldn't afford.
- d) For a long time, I really wanted to have an English bible, but I could not get which I satisfied.
- e) Every night, we saw a movie with my new DVD player.

In most cases, the narrow scope indefinites had *de dicto* readings, as a result of scoping under an intensional verb or modal (19a-d). There were also a few cases of indefinites taking narrow scope with respect to a higher quantifier, as in (19e). As previously discussed, narrow-scope indefinites are obligatorily [-specific].

Indefinites in *there*-constructions (20a-b) and as objects of *have* (20c-d) occurred primarily as descriptions of the narrator's room in response to (16c). Also included in this category were a few indefinites that occurred in locative copular constructions (20e); as part of a sentence fragment that followed a *there*-construction in a previous sentence (20f); and in the *consist of* construction, which, like the *have*-construction, denotes possession (20g).

20. *singular indefinites in there-constructions and as objects of 'have': correct use of 'a'*

- a) There are a bed, a table, four chairs, a TV, and some pictures on the walls.
- b) When I got at the Gainesville airport, there was a man from Korean church.
- c) My room is not big but not small. It has a table for studying and a chair.
- d) Here in the US I have a room which has a middle size. In my room I have a bed, desk, chairs, bookshelf and computer.
- e) In my room are a big bed, a table, TV, and three cheas.
- f) There are nothing much in my room. A computer, bookself and chair. Those are all I put in my room. Sometimes, a drying pole for the laundry was there.
- g) My bedroom is consist of queen size bed, hanger, closet, dresser and big mirro.

As we argued in Section 4.1.2, indefinites in *there*-constructions and *have*-constructions in a "listing" context are very likely to be [-specific]. The same holds for indefinites in locative constructions: for instance, in (20d), no particular importance is attached to a particular bed, table, or TV.

Postcopular indefinites were indefinites in the postcopular positions of copular constructions. We do not distinguish between different types of copular constructions, such as predicational and identificational (Higgins 1979).¹⁶ Some examples of indefinites in postcopular position are given in (21).

¹⁶ In Higgins's (1979) taxonomy, copular constructions can be predicational (*She is a teacher*), equative/identity (*Mark Twain is Samuel Clemens*), identificational (*This is Kim*), or specificational (*The winner is Chris*). The postcopular NP in predicational copular constructions is semantically a predicate and

21. *singular indefinites in postcopular position: correct use of 'a'*
- As I am a great traveler, I will spend more than half of this amount to discover a new place in the U.S...
 - Last Christmas, I visited parents' house with my husband. It's a white Christmas as everyone always wishes.
 - During my last vacation I got a chance to travel to Florida. That was a place I heard a lot about.

Indefinites in predicational copular constructions (21a) are necessarily [-specific] while indefinites in identificational copular constructions (21b-c) may be [+specific]: for (21c), for instance, the speaker may have in mind the noteworthy property *x is a beautiful place that I've wanted to visit for a long time* which singles out a particular individual from the set denoted by the NP *place I heard a lot about*.

4.3.2. *Types of plural/mass indefinites*. Plural and mass indefinites were similarly divided into several types; of particular interest to us are the four types in (22).¹⁷ These largely correspond to the types for singular indefinites, except that there were no mass or plural indefinites in postcopular position. All illustrations below are cases where L2-learners did not put any articles, as is appropriate. Most of the examples are with plurals, since there were very few mass nouns in these categories.

22. **Types of plural/mass indefinite DPs:**

- wide scope
- narrow scope
- there*-construction
- object of *have*

As was the case with singulars, wide-scope plural/mass indefinites in most cases took wide scope by default, in the absence of other operators, as in (23). Just like singular wide-scope indefinites, plural/mass wide-scope indefinites may be [+specific] or [-specific].

23. *plural/mass wide-scope indefinites: correct article omission*
- From airport (here in Boston) I was taken by representatives of my University and driven to a hotel.
 - The object that is of value for me is my father's gift: beautiful diamond earrings.

Narrow-scope plural/mass indefinites, which are obligatorily [-specific], always took scope under an intensional operator, modal, or negation, as in (24).

24. *plural/mass narrow-scope indefinites: correct article omission*
- For the rest of the money, I will buy souvenirs of the visited places and gifts for my Mom and nephews.
 - There's a few dogs for inspection to look for weapons or drugs.
 - [description of a skiing experience] Friend of mine show(n) me how to deal with speed, how to manage yourself without injury.

Examples of plural/mass indefinites in *there*-constructions and as objects of *have* are given in (25). As already discussed, indefinites in *there* and *have* constructions are likely to be [-specific].

25. *plural/mass indefinites in there-constructions and as objects of 'have': correct article omission*
- There only bed, iron table, some chairs, mirror and closets.
 - On the floor, there are always toys here and there.

thus must be [-specific]. The postcopular NP in the other types of copular constructions, on the other hand, has been analyzed as semantically an individual (see Mikkelsen 2004, Geist 2003), which means that it is in principle compatible with the [+specific] reading (the lexical entry in (2b) requires that specific DPs are semantically individuals rather than predicates). Thus, on our proposal, overuse of *the* for postcopular indefinites is expected in all postcopular constructions except predicational. However, it is difficult to test this prediction, since the different types of copular constructions cannot always be easily distinguished in naturalistic data.

¹⁷ As in the case of singular indefinites, various special cases of plural/mass indefinites, such as generics, are not reported here – see Ionin (2003, Ch. 7) for more discussion.

- c) The room where I moved recently is very nice. It has hardwood floors, closet, large windows in it.
 d) We went to several springs, and had BBQ parties and fishing there.

4.3.3. Results: articles in indefinite contexts. Tables 5 and 6 report article use for singular and plural/mass indefinites, respectively, in the main types of indefinite contexts discussed above.

As predicted, most overuse of *the* is restricted to wide-scope contexts – there is very little *the* overuse in narrow-scope contexts or in *there/have* constructions.

Table 5. Article use across all singular indefinite contexts; target: *a*

article used by L2-learners	wide scope	narrow scope	<i>there</i> -construction	object of <i>have</i>	postcopular position
L1-Russian					
the	13	2	3	0	1
<i>a</i>	17	18	17	32	18
null	16	4	22	50	13
L1-Korean					
the	8	4	0	0	1
<i>a</i>	37	35	42	42	28
null	15	27	22	27	20

Table 6. Article use across all plural/mass indefinite contexts; target: -- (null)

article used by L2-learners	wide scope	narrow scope	<i>there</i> -construction	object of <i>have</i>
L1-Russian				
the	4	1	1	0
<i>a</i>	1	3	0	2
null	34	16	5	13
L1-Korean				
the	3	4	0	0
<i>a</i>	2	1	0	2
null	3	20	2	4

Examples of *the* overuse in wide-scope indefinite contexts, singular as well as plural, are given in (26) and (27) for L1-Russian and L1-Korean participants, respectively. A rare case of *the* overuse in a *there*-construction is given in (26g).¹⁸ In all cases, there is no previous mention of the underlined DP's referent.

26. overuse of 'the' with wide-scope indefinites & in a *there*-construction (L1-Russian participants)

- a) When I was living in Ulan-Ude yet unmarried my friends presented me the small seamese kitten.
 b) I lost the health tooth, and I have realized after some time how it was valuable for me. It happened unexpectedly – I bit off the solid sweet and that's all: my nice – facial! – tooth was fractured.
 c) My husband met us in airport and drove us to our new home. Then we went to our neighbours house for the small party.

¹⁸ Milsark (1974) notes that definites can in fact occur in *there*-constructions in list contexts (thanks to an anonymous reviewer for pointing this out to us). It is possible that the apparent overuse of *the* in (26g) is in fact appropriate use of *the* in a list context. However, we note that the L1-English coders never used *the* in (26g), using instead *a* for singulars, and no article or *some* for plurals.

- d) When I was a boy, I found a mine (I mean, an armour, from the World War Two). I liked this kind of things, so I kept it initially in the secret place in our yard, and then at home.
- e) On Thanksgiving week-end we went to NY for the first time. We took the room in the New-Yorker Hotel and went outside to see the town.
- f) First I arrived in the US at the end of June. It was in New York. I have met a lot of people. I had to stay at the long line in order to get through the custom.
- g) I have beautiful room. There are the big sofa, the table, the chairs, TV and shelves with the books.
27. *overuse of 'the' with wide-scope indefinites (L1-Korean participants)*
- a) I received the frame with picture for Christmas from my roommate.
- b) [discussion of how the author's husband gave her a bible] When he gave me the bible, he attached the memo which was written about the his love about me.
- c) The most valuable object that I have received is the ball and the signature of the famous baseball player is signed on it.
- d) I went to the cruise to Bahamas with my mother.
- e) In New Orleans, we visited a couple of well-known places and had a traditional cuisine at the restaurant in French quarter.
- f) We have been looking for a really good used car. Eventually we got the information, but the problems came out.
- g) Through the window from airplanes, I saw African-American men working with the packages.

Many of the above contexts are arguably [+specific]: in most cases, the narrator can single out a particular individual by naming something noteworthy about it. For instance, in (26a), the kitten was the author's much-loved pet; in (26d), the author has a very particular secret place in his back yard in mind; in (27f), the author knows exactly what the problems were; and so on. Contexts which are less likely to be specific include (26e) and (26g), where the identity of a particular hotel room, or a particular set of books, are quite irrelevant.¹⁹

Next, we examine overuse of *the* in contexts where it is completely unexpected – narrow-scope environments, which are obligatorily [-specific]. The complete lists of the instances of *the* overuse in narrow-scope contexts for L1-Russian and L1-Korean speakers are given in (28) and (29), respectively. Most of the answers are the L2-learners' descriptions of how they would spend a thousand dollars.

28. *overuse of 'the' with narrow-scope indefinites: L1-Russian participants*
- a) If it is happen I'll spend money for the trip to California or Florida. I'm tired for winter this year.
- b) I'd spend some money for travelling anywhere. And another part I'd spend on the gifts for my sons and daughter-in-law.
- c) I would go for example to the Grand Canyon and use the chance to fly over it in the helicopter.
29. *overuse of 'the' with narrow-scope indefinites: L1-Korean participants*
- a) And I will spend the rest of money (\$200) for my daughter. like buying her the bike and dolls.
- b) After get that money, I'll buy a new car and buy the clothes for my child.
- c) First of all, I will donate \$100 as an offering to my church. With the rest of it, I will take a trip to the famous place or city and some historical cities, which especially were cited in the Bible.
- d) I remeber that first day I was very tired. it's first time I spent over twenty hours in the plane.

¹⁹ In a few cases of *the* overuse, use of *the* does not appear to be completely infelicitous: in (27g), for instance, *the packages* might be accommodated if one imagines packages as being an obligatory item in an airport (e.g., if it taken to mean *the luggage*). In the case of (26e), an anonymous reviewer suggests that use of *the room* could be similar to use of *the telly*: every hotel is assumed to have a room in which one stays, just as every family is assumed to own a television set, making *the room* a definite by associative use.

However, in most cases of *the* overuse in the data, use of *the* is truly infelicitous: for instance, in (26a), there is no way that the reader can presuppose the existence of a unique kitten, in (26b), the narrator has multiple healthy teeth and has undoubtedly eaten multiple sweets, and so on.

- e) I really want to visit the famous city in a European country which has old history. The exact place will be depended on the time which I have or the flight price.
- f) I wanna give my church \$1000 as a donation for the missionaries. I hope the money would be spent to the valuable work or people.
- g) First of all, I will donate a tenth of \$1000 to the church. Then, I will buy the gifts for my son and wife.
- h) It feels always fresh to walk around the new places.

In some cases, even though the context looks like a narrow-scope (and hence [-specific]) indefinite context, it may not be: for instance, in (29a), the narrator may have in mind a particular bike that he'd like to buy (e.g., the one his daughter has been asking for); similarly, in (29b), the narrator may need to buy his child a particular set of clothes (e.g., a school uniform). Hence, these contexts may in fact be wide scope and [+specific] (although we cannot know for sure). Similarly, in (29c), the author may be intending to use a plural DP (*the famous places or cities*) which would in fact be [+specific] (having the noteworthy property of being *the places cited in the Bible*). A couple of other cases, such as (28c) and (29d), contain what look like possible definite generics, referring to the *class* of helicopters or planes. Only a few cases ((28a-b), (29e-h)) are unambiguously narrow-scope and hence [-specific].

There were two instances of *the* overuse in postcopular position. These are given in (30a) and (30b), from an L1-Russian speaker and an L1-Korean speaker, respectively. Both may be cases of [+specific] indefinites in identificational copular constructions (see footnote 16): in (30a), the narrator may have in mind the noteworthy property *x is the city that I had been planning to visit for a long time*, and in (30b), the narrator may have in mind the noteworthy property stated in the non-restrictive relative clause: *x is the watch that I always wanted to have but couldn't afford*. As with wide-scope indefinites, we cannot know for sure whether the DP is specific without having access to what the narrator had in mind.

30. *overuse of 'the' in postcopular position*

- a) On the last summer I went to the Prague. It is the very beautiful, ancient city.
- b) On my last birthday, I receive a special gift from one of my best friends. It was the watch made by Swiss army, which I always wanted to have but haven't because of its high price.

4.3.4. *Statistical comparisons*. In this section, we will look at whether *the* was overused significantly more often with [+specific] than with [-specific] indefinites. We will treat wide-scope indefinites as (potentially) [+specific], and narrow-scope indefinites, as well as indefinites in *there* and *have* constructions, as [-specific]. Postcopular indefinites are not included.

First, we looked at singular indefinites and computed the χ^2 statistic²⁰ on use of *the* vs. use of *a*. Table 7 reports the χ^2 distribution in *the* vs. *a* use for wide-scope indefinites vs. indefinites in *there* and *have* constructions. The distribution of articles was highly significant for both L1 groups: both groups used *the* more in wide-scope contexts than in *there/have* constructions. Table 8 reports use of *the* vs. *a* for wide scope vs. narrow scope contexts. This distribution was significant only for the L1-Russian speakers.

Next, we grouped singular and plural/mass contexts together and looked at whether the distribution in use of *the* vs. use of *a*/article omission was significant. Table 9 compares use of *the* vs. use of *a*/omission for wide scope vs. *there/have* contexts; for both L1 groups, the distribution is highly significant. Table 10 compare wide scope and narrow scope contexts; the distribution is significant only for L1-Russian speakers.

Table 7. Article use with wide scope indefinites vs. in *there* and *have* constructions: singular contexts only

L1-Russian speakers		
	wide scope	there/have
the	13	3
a	17	49

$\chi^2 = 17.09^{***}$

L1-Korean speakers		
	wide scope	there/have
the	8	0
a	37	84

$\chi^2 = 15.92^{***}$

***p<.001

²⁰ Since only group results are analyzed, the χ^2 test is used only to get an approximation of whether the distribution is significant. There were not enough contexts to allow individual analyses.

Table 8. Article use with wide scope vs. narrow scope indefinites: singular contexts only

L1-Russian speakers		
	wide scope	narrow scope
the	13	2
a	17	18

$$\chi^2 = 6.35^*$$

L1-Korean speakers		
	wide scope	narrow scope
the	8	4
a	37	35

$$\chi^2 = 0.97$$

*p<.05

Table 9. Article use with wide scope indefinites vs. in *there* and *have* constructions: all contexts

L1-Russian speakers		
	wide scope	there/have
the	17	4
a / omission	68	141

$$\chi^2 = 19.20^{***}$$

L1-Korean speakers		
	wide scope	there/have
the	11	0
a / omission	57	141

$$\chi^2 = 24.08^{***}$$

***p<.001

Table 10. Article use with wide scope indefinites vs. narrow scope indefinites: all contexts

L1-Russian speakers		
	wide scope	narrow scope
the	17	3
a / omission	68	41

$$\chi^2 = 3.85^*$$

L1-Korean speakers		
	wide scope	narrow scope
the	11	8
a / omission	57	83

$$\chi^2 = 2.02$$

*p<.05

4.4. Article use with definites

We now turn to article use in definite contexts in the data. We consider both singular and plural/mass definites.

4.4.1. *Types of definites.* Definite DPs (singular, plural, and mass) were classified into a number of different types, the most relevant of which are reported in (31).²¹ This was done in order to determine whether there was any relationship between the type of definite and L2-English article choice.

The contexts are described in more detail below, with illustrations from the L2-learners' narratives. Illustrations always include cases where the L2-learners put in the target article, *the*.

31. Types of definite DPs:

- anaphoric
- unique by entailment
- associative use
- obligatorily unique
- narrow-scope

The *anaphoric* use of a definite (32) is the basic previous-mention use – any case where the referent of the DP was explicitly mentioned in the previous discourse. As discussed previously, anaphoric definites are obligatorily [+specific]: in (32a), for instance, the speaker has in mind the noteworthy property *x is the sweater that my mom made and which I wear and love* which distinguishes the referent of *the sweater*.

32. *anaphoric definites: correct use of 'the'*

- a) I have *a beautiful sweater* that my mom made [description of how the narrator's mother made it follows] And I still wear the sweater and I love it.
- b) We live in *half-basement room*. It's small, but enough for us. In the room we have full-size bed, a large dresser and smaller dresser, and 2 tables – one is for my husband and the other is mine.

²¹ Not reported here are various special cases such as generics, superlatives, and proper names – see Ionin (2003, Ch. 7) for more discussion.

The *unique by entailment* definites are characterized by the presence of a complement phrase. These definites were classified in Hawkins's (1978) traditional classification of definites as a subtype of *unfamiliar definites*. There is no previous mention of the referent in these contexts, or any knowledge of the referent on the part of the hearer. The presupposition of uniqueness (or, for plurals, maximality) is satisfied because the complement phrase narrows down the domain of the discourse sufficiently to allow the hearer to share the speaker's knowledge of uniqueness. Thus, in (33a), the speaker refers to the maximal set of doctors at a particular hospital; in (33b), the speaker refers to the fame of a particular city; and so on for the other cases in (33).

33. *definites that are unique by entailment: correct use of 'the'*
- a) I'm really thankful to the doctors of Moscow Hospital #20.
 - b) I visited the New York city with my wife by my car during the winter vacation. I've just heard the fame of the great city.
 - c) I spent my last winter break in Puerto-Rico and Miami. First, I flew to San-Juan, the capital of Puerto-Rico.
 - d) First of all, I had to go to the housing office in the school to get my room to stay in.
 - e) But Key West is the best place. Besides the beautiful scene there.

Definites which are unique by entailment may be [+specific] or [-specific], depending on whether or not the speaker has in mind some noteworthy property. In such cases as (33a), for instance, the speaker could have in mind a noteworthy property such as *x are doctors who cured me*, which distinguishes the referent of the underlined DP. Similarly, in all of the other cases in (33), the speaker is able to provide a noteworthy property that holds of the referent of the definite description. Definites that are unique by entailment are thus quite likely to be [+specific], although we cannot know for sure. One exception would be cases where all speaker knowledge is explicitly denied (e.g., *I want to speak to the winner of this race, whoever that is*). However, such contexts are extremely unlikely to be produced in L2-data; in fact, we did not find a single definite (or, for that matter, indefinite) context in the data which involved explicit denial of speaker knowledge.

The *associative use* of definites is similar to the *associative anaphoric use* of definites often described in the literature (see Poesio and Vieira 1997 for an overview). When such definites are used, "Speaker and hearer may have (shared) knowledge of the relations between certain objects (the triggers) and their components or attributes (the associates): associative anaphoric uses of definite descriptions exploit this knowledge" (Poesio and Vieira 1997:6). The hearer/reader frequently has to accommodate the uniqueness/maximality presupposition: thus, in (34a), the reader is led to understand that the relevant sunset view is the one at Sarasota; similarly, in (34b), the reader understands that the relevant windows and yard are the ones that go with the narrator's room. This category included DPs that are regularly used with the definite article despite lack of previous mention, such as *airport*, *ocean*, and *beach* (as in 34c-d). The reader typically understands that there is a unique and contextually salient airport, beach, etc., under discussion.

34. *associative use of definites: correct use of 'the'*
- a) Last Christmas, my friends who live in Indiana came here for vacation. We went to Sarasota to see the sunset view, and go fishing.
 - b) Our room is pretty big [a description of the room's furniture follows] Unfortunately the windows looks to the yard, therefore we don't have enough light.
 - c) I went to the Georgia state where I finished near the beach but I caught nothing at night.
 - d) I remember I came to the US at night. I was so tired and sleepy because of the tiring, long-lasting flight that I wanted to sleep. I was met by a representative of ESL program at the airport who was actually the first person I practiced my English in the U.S.

Definites on their *associative use* are more likely to be [+specific] than [-specific], since the referent of the definite DP can usually be singled out via some noteworthy property – arguably the same property that the hearer has to accommodate in order to consider the use of the definite article felicitous. For instance, in (34a), the noteworthy property might be *x is the view at Sarasota*, and in (34c) it might be *x is the beach in Georgia where I traveled to*.

Definites which fall into the category *obligatorily unique* (35) are, as their name suggests, definites with referents that are obligatorily unique in the actual world. Definites in this class are likely to be [+specific], since the narrator is referring to a very particular world or sun: the only relevant one in

existence, from the perspective of Earth's inhabitants (however, it is in principle possible for these contexts to be [-specific] if the identity of the sun, world, etc., is not in any way noteworthy).

35. *definites with obligatorily unique referents: correct use of 'the'*

- a) There so many places in the US and across the world worth seeing.
- b) It was already too cold to swim, but still quite nice to enjoy the sun.
- c) Note-book is placed on my desk which makes me and my wife surf the internet.

Narrow-scope definites are definites which take scope under an intensional verb or modal. In these cases, the speaker does not have in mind a particular individual in the actual world which is the referent of the definite. In (36a), the pictures and tape of the narrator's baby boy exist only in the possible future world in which the narrator succeeds in video-taping her son. Similarly, in (36b), the stealing of the car has not taken place in the actual world – in fact, it is an event that the narrator's friend wants to prevent from ever taking place. Narrow-scope definites are obligatorily [-specific].

36. *narrow-scope definites: correct use of 'the'*

- a) I have a really cute baby boy. I want to record all of the pictures of him and watch the tape after he grows up.
- b) My friend locked the wheel to prevent the steal of car.

To sum up, narrow-scope definites are the only definites appearing in the data which are clearly [-specific]. All other categories of definites are most likely to be used in [+specific] contexts.

4.4.2. *Results: articles in definite contexts.* Table 11 reports article use in singular contexts for the five categories of definites discussed above, and Table 12 reports article use in plural/mass contexts. Since *a* was never used with plural/mass definites, only use of *the* vs. article omission are reported for plurals.

Table 11. Article use with different types of definites: singular definite contexts

<i>article used by L2-learners</i>	anaphoric	entailment	associative use	obligatory uniqueness	narrow scope
<i>L1-Russian</i>					
the	18	18	28	2	2
<i>a</i>		1			
null	6	3	12	1	
<i>L1-Korean</i>					
the	52	23	33	4	4
<i>a</i>	1	1	5		1
null	6	4	8	1	

Table 12. Article use with different types of definites: plural/mass definite contexts

<i>article used by L2-learners</i>	anaphoric	entailment	associative use	obligatory uniqueness	narrow scope
<i>L1-Russian</i>					
the	3	2	5		
null		1	2		
<i>L1-Korean</i>					
the	9	9	13	3	2
null	3	2	2	1	

These tables do not suggest any particular pattern of article use or omission. Cases of *a* overuse occur in different categories, primarily in the *associative use* category in the case of the L1-Korean speakers. Article omission also occurs across categories, and is clearly not tied to lack of specificity – in fact, there is no article omission (and very little article use) in the clearly [-specific] category of narrow-scope definites.

We will now look at *a* overuse in the data. The cases of *a* overuse for L1-Russian and L1-Korean speakers are given in (37) and (38), respectively. In most of these cases, there is no obvious link between use of *a* and the [-specific] feature. The only DP that is obviously [-specific] is the narrow-scope definite in (38h): since the speaker is not talking about any particular full collection of toys, she cannot have any noteworthy property in mind that would single out a unique collection of toys. The previous-mention definite in (38a), on the other hand, is clearly [+specific]. The other contexts below may be [+specific] or [-specific] – we cannot know one way or the other without knowing exactly what the speaker had in mind. In (38c), it is likely that the author did not consider the nurse’s identity important; on the other hand, in (38f), the identity of the landscape view seems fairly important: it is *the landscape view of Florida*. Finally, in some cases, the underlined DP, which the L1-English coders treated as definite, may have been intended by the narrator to be indefinite: this may be the case for (38b-d), in particular.

37. *overuse of ‘a’ with definites: L1-Russian speakers*

entailment context: may be [-specific]

Before I got here I have lost my keys from my apartment. [description of the loss follows] After I had to change a lock of front door.

38. *overuse of ‘a’ with definites: L1-Korean speakers:*

a) *anaphoric context: must be [+specific]*

There’s a bed for me and my wife and crib for my baby. Also there’s a chain on which my wife feed baby. I have a humidifier and air purifier to keep pleasence for a baby.

b) *entailment context: may be [-specific]*

First impressive thing was that people are gentle and generous. Later, I came to know that there is a reason of that generosity. It is welfare of this country. In my country people should compete each other to live well. It might be a reason why people in Korea are less generous.

associative contexts: may be [-specific]

c) My daughter is very precious to me. She was born about 16 months ago. At that time I was in great pain and almost lost my conscience. But when a nurse show me my daughter, I thought she was an angel. Now I raise an angel.

d) The first person whom I talked to was a lady in the airplane. [description of the lady follows] Her husband was supposed to pick her up at the airport. After I got off the airplane, they were kind enough to drive me and my heavy luggages to a dormitory.

e) 2 days before thanksgiving, my baby was bourn. We have to spend a holiday at the hospital.

f) Florida is very flat territory. There is no mountain that’s different from Korean. Also a landscape view is wide to make me feel very fresh.

g) When I got the Syracuse airport, NY, I was very nervous about a new environment, even I couldn’t able to speak English well.

h) *narrow scope context: must be [-specific]*

I have a daughter, 15 month old. I keep several toys for her in my mind. So, if I got \$1000, I will buy her those toys. But unfortunately, \$1000 is not enough to buy many toys. For example, a Brio wooden train costs almost half of \$1000. It is very tough to raise a kid with a full collection that he needs.

To sum up, there was very little overuse of *a* with definites in the data, and the few cases of *a* overuse were not clearly tied to lack of specificity. On the other hand, there were also very few contexts that are unambiguously [+definite, -specific]. Such contexts are narrow-scope definites. As shown in Table 11, there were only two singular narrow-scope definites in the Russian speakers’ data (both of which were used with *the*) and five singular narrow-scope definites in the Korean speakers’ data (one of which was used with *a*).

Finally, we will take a look at article use vs. omission in singular contexts, definite as well as indefinite. We computed the χ^2 statistic for article use (*the+a*) vs. article omission with definites vs. indefinites in all singular contexts. As shown in Table 13, the χ^2 distribution was highly significant for both L1 groups. L2-learners were significantly more likely to omit articles with indefinites than with definites in singular contexts.

Table 13. Article use vs. omission with definites vs. indefinites in singular contexts

L1-Russian speakers			L1-Korean speakers		
	definite	indefinite		definite	indefinite
<i>the+a</i>	115	132	<i>the+a</i>	171	208
omission	26	109	omission	34	114

$\chi^2 = 27.94^{***}$

$\chi^2 = 21.96^{***}$

***p<.001

4.5. Discussion

We will now examine the predictions made in Section 4.1, looking at whether they are supported by the results of the written narrative data.

4.5.1. *L2-English article use and Prediction A*. First, we look at whether L2-English article use in the written narrative data followed the patterns established by the formal written elicitation studies – whether Prediction A, repeated in Table 3, was supported.

Table 3. Prediction A for article choice in written narrative data in L2-English

	+definite	-definite
	Target: <i>the</i>	Target: <i>a</i> (sg), -- (pl, ms)
+specific	correct use of <i>the</i>	overuse of <i>the</i>
-specific	overuse of <i>a</i> , --	correct use of <i>a</i> , --

sg = singular count noun

pl = plural count noun

ms = mass noun

First, we consider the particular prediction for article use with indefinites, repeated in (39).

39. *Prediction A1: article misuse with indefinites*

Overuse of *the* with indefinites will be limited to [wide-scope indefinite contexts](#), and will not occur with narrow-scope indefinites, or in *there* and *have* constructions.

These prediction were largely supported. Overuse of *the* was in fact largely limited to wide-scope indefinites. There was practically no overuse of *the* in *there* and *have* constructions. While some overuse of *the* occurred with narrow-scope indefinites, contrary to (39), it was fairly low; moreover, some narrow-scope indefinites in which *the* overuse occurred may be more appropriately classified as wide-scope indefinites or generics. We furthermore saw a couple of cases of *the* overuse in the postcopular position; in both cases, the indefinite was likely to have a [+specific] reading.

Next, we consider the prediction for article use with definites, repeated in (40).

40. *Prediction A2: article misuse with definites*

Overuse of *a* with definites will occur primarily in narrow-scope definite contexts, and other definite contexts in which speaker knowledge of the referent is denied.

We saw that L2-English learners overwhelmingly used *the* across all definite contexts in written narrative data: anaphoric definites, definites which are unique by entailment, definites on their associative use, definites with obligatorily unique referents. All of these definites may (or even must) be [+specific]: the speaker typically had knowledge of the definite DP's referent and could state something noteworthy about it.

Narrow-scope definites and definites with a denial of speaker knowledge were the only categories on which overuse of *a* was predicted. However, narrow-scope definites were extremely rare in the data, while definites with a denial of speaker knowledge were absent altogether. Therefore, the prediction in (40) cannot be tested.

To summarize, we see that Prediction A was partially supported by the written narrative data. L2-English learners did indeed overuse *the* in [-definite, +specific] contexts to a much greater degree than in [-definite, -specific] contexts. They also used *the* correctly in [+definite, +specific] contexts. Unfortunately, [+definite, -specific] contexts were largely absent from the data, so we do not know whether overuse of *a* (as well as bare plurals/mass nouns) is in fact prevalent in these contexts.

4.5.2. *Article omission in written narrative data.* Next, we examine the predictions concerning article omission in L2-English, and the corresponding results. Prediction B, which concerns article omission, is repeated in (41).

41. Prediction B:

L2-English learners omit articles more in singular [-definite] than in singular [+definite] environments in written narrative data.

This prediction was fully supported. We saw that both L1-Russian and L1-Korean speakers were in fact significantly more likely to omit articles with indefinites than with definites. This suggests that article omission in L2-English, like article use, is not random. While factors such as performance pressure may cause L2-learners to omit articles more in written narrative data than in a controlled formal written elicitation study, the learners do not omit articles to the same extent across all categories. When the uniqueness presupposition is satisfied, L2-learners use *the*, rarely allowing article omission. They are much more likely to omit *a*, which carries little semantic information. Thus, article omission, like article use, shows that L2-English learners are aware of article semantics.

5. CONCLUSION

In this paper, we examined L2-English article use in light of the features [+definite] and [+specific]. We reviewed our three formal written elicitation studies, which provide evidence that errors in L2-English article use are not random, but are due to an association of *the* and *a* with the features [+specific] and [-specific], respectively. We then reported written narrative data that largely supported the conclusions drawn on the basis of the formal written elicitation task. The written narrative data showed that overuse of *the* with specific indefinites holds in written narrative data as well as formal written elicitation. While overuse of *a* with definites was not replicated in the written narrative data, this quite likely stemmed from the fact that [-specific] definite contexts are rare in written narrative data.

We conclude that L2-English learners have access to the universal semantic features [+definite] and [+specific], but do not know which of these features underlies article choice in English. As a result, they fluctuate between dividing English articles on the basis of definiteness vs. on the basis of specificity, until the input leads them to choose the [+definite] feature as the feature that underlies English article choice. Our findings provide evidence for direct access to universal semantic features in L2-acquisition, as well as for the reality of the feature [+specific].

6. APPENDIX: CODING PROCEDURE

6.1. Details of the coding procedure

The coders' versions of the written narratives were prepared as follows. Every article used by an L2-learner was replaced with a blank. The only exceptions were articles in formulaic contexts, such as *a few* and *a lot of*. These were left in place. Furthermore, a blank was inserted before each NP which lacked an article in the L2-learners' data; this was done for NPs which did not require an article (e.g., indefinite plurals) as well as for those that did. The exceptions were proper or geographic names which do not require an article, and indefinite plural or mass nouns following expressions such as *a lot of*, since such expressions may be treated by L2-learners as unanalyzed chunks: for instance, an L2-learner may treat *a lot of* as a determiner that is incompatible with an article (cf. **I have many the books*). Expressions such as *a lot of* were nearly always followed by bare plurals in the L2-data. Finally, no blanks were inserted before DPs which contained numerals or quantifiers.

Another change made for the coders' version of the narratives was correction of spelling errors; this was done so that the coders could concentrate on the grammar and not be distracted by spelling mistakes (which occurred very frequently in the narratives of less advanced L2-learners). Only obvious spelling mistakes were corrected – e.g., words like *coutry*, *appartements*, and *spetial* were changed to *country*, *apartments*, and *special*, respectively. When it was not clear what word the L2-learner had intended, no correction was made: for instance, one L2-learner invented the word *strangeous*: it was not clear whether the learner meant *strange*, *dangerous*, or both at the same time; the word was left uncorrected.

No errors that were grammatical in nature, such as errors in plurality or person marking, were corrected. Capitalization was corrected in many cases (e.g., many L2-learners failed to capitalize the first word in a sentence). Punctuation was corrected only to a very small extent (see Ionin 2003, Ch. 7 for more discussion).

As an illustration of the changes that an L2-learner's narrative underwent before being given to a coder, consider (42). The passage in (42a) is an L1-Russian speaker's response to the question in (16b). The passage in (42b) is the version that was given to the coders. The spelling error in *buatiful* has been corrected, but the error in *haved* has not, since this error involves incorrect formation of a past tense form. Each article in the original passage has been replaced a blank, and blanks have also been inserted before the phrase *very beautiful impressions*.

42. a) I first arrived in the U.S 5 years ago, in September, 97 as a tourist. The first person whom I talked was officer of I.N.S. I stayed at my parents, who lived in Newton M.A. I haved two trips, to New York and Washington (district Columbia). That was very buatiful impressions of the trips.
- b) I first arrived in _____ U.S 5 years ago, in September, 97 as _____ tourist. _____ first person whom I talked was _____ officer of I.N.S. I stayed at my parents, who lived in Newton M.A. I haved two trips, to New York and Washington (district Columbia). That was _____ very beautiful impressions of _____ trips.

6.2. Coding by L1-English speakers

The resulting versions of L2-learners' narratives were given to native English speakers to code, with written instructions. The coders were told that they were looking at actual narratives of L2-English learners, with all the articles removed. They were asked to read the narratives carefully and to insert, in each blank, the article that they considered most appropriate: the possible choices were *the*, *a*, and -- (the last to be used when no article was required). The coders were asked to use one of these three options whenever possible; they were told that if none of the three choices sounded right, they could insert other words, such as possessives, numerals, demonstratives, or *some*. The coders were encouraged to pay attention to the plurality of the noun: if it looked like the L2-learner was clearly intending to use a plural form but put in a morphologically singular form (as in, *I read one of the book*), the coders could treat the entire NP as a plural²².

Finally, the coders were told that if they could not understand a given context at all (because of grammatical errors on the part of L2-English speakers) and did not know what article would be appropriate, they were to write in a question mark. The coders were asked to ignore grammatical errors in the narratives as much as they could, and to focus on the meaning that the L2-learner was trying to convey, since the meaning of the context determined what article would be most appropriate.

Since the task of reading the narratives and inserting the appropriate article required much concentration on the part of the coders, there was a worry that once the coders became tired, they would start disregarding the context and putting in articles more or less randomly. To avoid this, we set a time

²² Omission of plural marking was actually fairly low: in unambiguously plural contexts (e.g. *two NP*, *one of the NP*, *many NP*, etc.), L1-Russian speakers omitted the -s morpheme 9.8% of the time (4 instances out of 41 obligatory contexts), and L1-Korean speakers omitted it 13.6% of the time (12 instances out of 88 obligatory contexts). Since Korean (unlike Russian) lacks plurality marking, the higher -s omission on the part of the Korean speakers may be due to L1-transfer. However, the overall relatively low rates of omission suggest that both groups of L2-learners had mastered or nearly mastered the rules of plural formation in English.

limit of 40 minutes on the task, and instructed the coders to code as much as they could in 40 minutes, focusing on accuracy rather than speed. The rates of coding varied from one coder to another, so that some could code as many as 15 sets of narratives (i.e., responses from 15 L2-learners, each consisting of answers to five questions) in 40 minutes, while others got through only half as many. The ultimate goal was to have every context coded by exactly four native English speakers. This goal was achieved after 11 coders had been recruited to code the narratives of the Korean speakers, and seven coders – to code the narratives of the L1-Russian speakers.²³

7. ACKNOWLEDGEMENTS

We would like to thank Suzanne Flynn, Irene Heim, and David Pesetsky for comments and suggestions. We are grateful to the audiences of GALA 2003, BUCLD 28, and the UCLA Psychobabble Seminar, where parts of this paper were presented, to the editors of this volume, and to two anonymous reviewers. Any remaining errors are our own.

8. REFERENCES

- Carlson, Greg. 1977. A unified analysis of the English bare plural. *Linguistics and Philosophy*, 1, 413-457.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics*, 6, 339-405.
- Donnellan, Keith. 1966. Reference and definite descriptions. *The Philosophical Review*, 75, 281-304.
Reprinted in D. Steinberg and L. Jakobovits (eds.), 1971, *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, Cambridge: Cambridge University Press, 100-114.
- Eubank, Lynn. 1993/1994. On the transfer of parametric values in L2 development. *Language Acquisition* 3, 183-208.
- Fodor, Janet and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy*, 5, 355–398.
- Gavruseva, Elena. 2003. Aktionsart, aspect, and the acquisition of finiteness in early child grammar, *Linguistics*, 41 723-755.
- Geist, Ljudmila. 2003. Predication and equation in copular clauses: The syntax-semantics interface. Ms., ZAS, to appear in a Kluwer volume in the SLAP series, edited by Ileana Comorovski.
- Hawkins, John A. 1978. *Definiteness and indefiniteness*. London: Croom Helm.
- Heim, Irene. 1991. Articles and definiteness. Published in German as “Artikel und Definitheit,” in A. v. Stechow & D. Wunderlich (eds.), *Semantics: An international handbook of contemporary research*. Berlin: de Gruyter.
- Higgins, Roger. 1979. *The pseudo-cleft construction in English*. New York: Garland Publications.
- Huebner, T. 1983. *A longitudinal analysis of the acquisition of English*. Ann Arbor: Karoma.
- Ionin, Tania. 2003. Article semantics in second language acquisition. Ph.D. thesis, MIT. Distributed by MITWPL.
- Ionin, Tania and Ken Wexler. 2003. The certain uses of *the* in L2-English. In J. Liceras et al. (eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*, Cascadilla Press, 150-160.
- Ionin, Tania, Heejeong Ko and Ken Wexler. 2004. Specificity as a grammatical notion: Evidence from L2-English article use. In G. Garding and M. Tsujimura (eds.), *Proceedings of WCCFL 22*, Cascadilla Press, 245-258.
- Ionin, Tania, Heejeong Ko and Ken Wexler. To appear. Article semantics in L2-acquisition: the role of specificity. *Language Acquisition*.

²³ Two additional coders were excluded due to not being native English speakers. One coder of the Korean narratives was a native speaker of Hindi, and one coder of the Russian narratives was bilingual in English and Urdu; it was not clear which of the two languages was primary for this coder. The coding of these two coders diverged noticeably from the coding of the native English-speaking coders.

- Ko, Heejeong, Tania Ionin, and Ken Wexler. 2004. Adult L2-learners lack *the* maximality presupposition, too. Paper presented at the 1st GALANA Conference, Honolulu, Hawaii.
- Lardiere, Donna. 2000. Mapping features to forms in second language acquisition. In J. Archibald (ed.), *Second language acquisition and linguistic theory*. Oxford: Blackwell, 102-129.
- Leung, Yan-kit Ingrid. 2001. The initial state of L3A: Full Transfer and Failed Features? In X. Bonch-Bruевич, W.J. Crawford, J. Hellermann, C. Higgins & H. Nguyen (eds.), *The past, present and future of second language research: Selected proceedings of the 2000 Second Language Research Forum*. Somerville, MA: Cascadilla Press, 55-75.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge, UK: Cambridge University Press.
- Maclaran, Rose. 1982. The semantics and pragmatics of the English demonstratives. Ph.D. dissertation, Cornell University.
- Master, P. 1987. A cross-linguistic interlanguage analysis of the acquisition of the English article system. Ph.D. dissertation, UCLA.
- Mikkelsen, Line. 2004. Specifying who: on the structure, meaning, and use of specificational copular clauses. Ph.D. dissertation, UCSC. Distributed by SLUG Publications.
- Milsark, Gary. 1974. *Existential sentences in English*. Ph.D. dissertation, MIT.
- Mosel, Ulrike and Even Hovdhaugen. 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Murphy, Susan. 1997. Knowledge and production of English articles by advanced second language learners. Ph.D. dissertation, University of Texas at Austin.
- Parrish, Betsy. 1987. A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37, 361-383.
- Poesio, Massimo and Renata Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics*, 24, 183-216.
- Prévost, Philippe and Lydia White. 2000. Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103-133.
- Prince, Ellen. 1981. On the inferencing of indefinite-*this* NPs. In A.K. Joshi, B.L. Webber, and I.A. Sag, (eds.) *Elements of discourse understanding*. Cambridge: Cambridge Univ. Press, 231-50.
- Robertson, Daniel. 2000. Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16, 135-172.
- Sorace, Antonella. 2000. Syntactic optionality in non-native grammars. *Second Language Research*, 16, 93-102.
- Stalnaker, Robert. 1974. Pragmatic Presuppositions, in M.K. Munitz and P.K. Unger (eds.), *Semantics and Philosophy*, New York University, New York, 197-213.
- Thomas, Margaret. 1989. The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10, 335-355.
- Young, R. 1996. Form-function relations in articles in English interlanguage. In R. Bayley and D.R. Preston (eds.), *Second language acquisition and linguistic variation*. Amsterdam: John Benjamins, 135-175.